A Biologically Inspired Analysis of the Incidence of the Aperture Problem in Natural and Urban Scenes

Lucas Pinto
Disciplina: Visão Computacional
Laboratório de Neurodinâmica
Depto. Fisiologia e Biofísica
ICB - UFMG
Junho de 2008

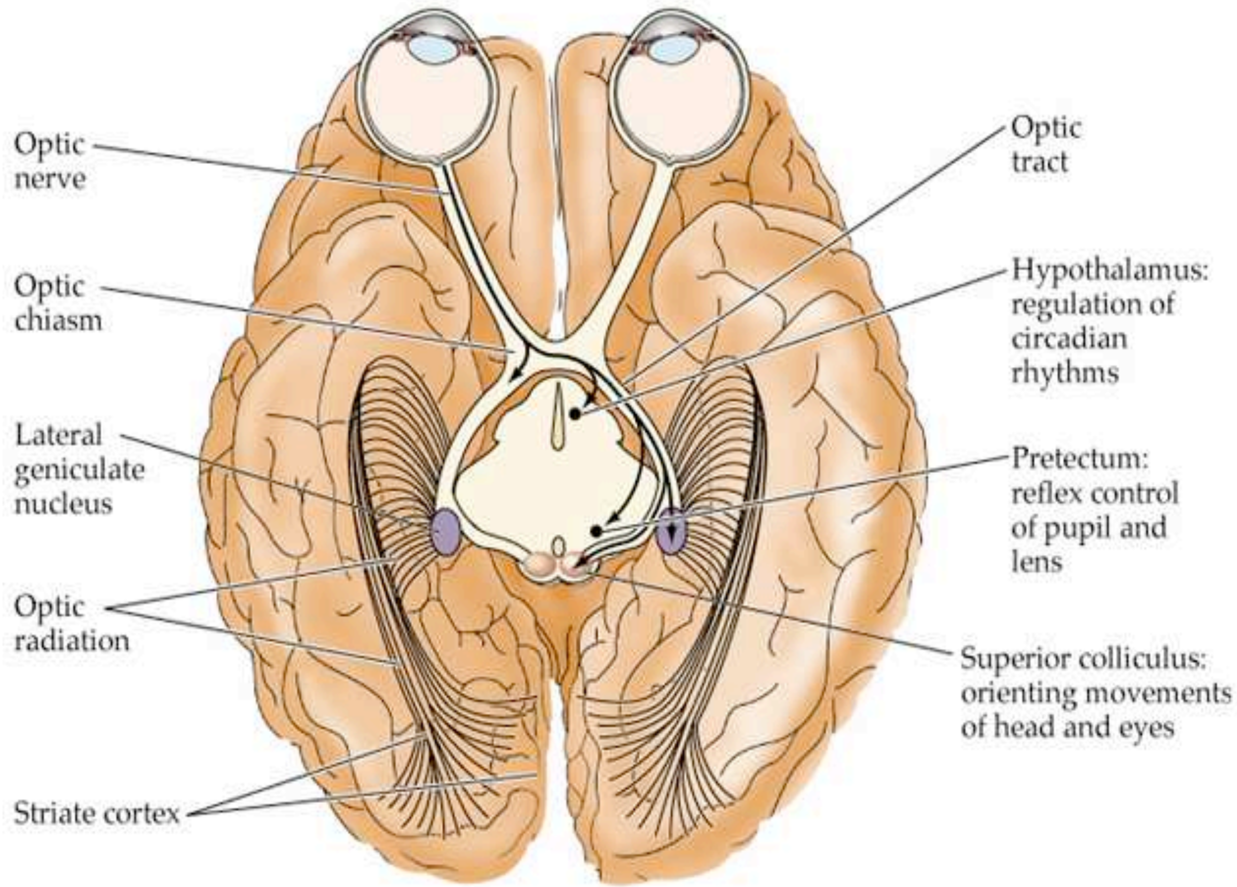
Introdução

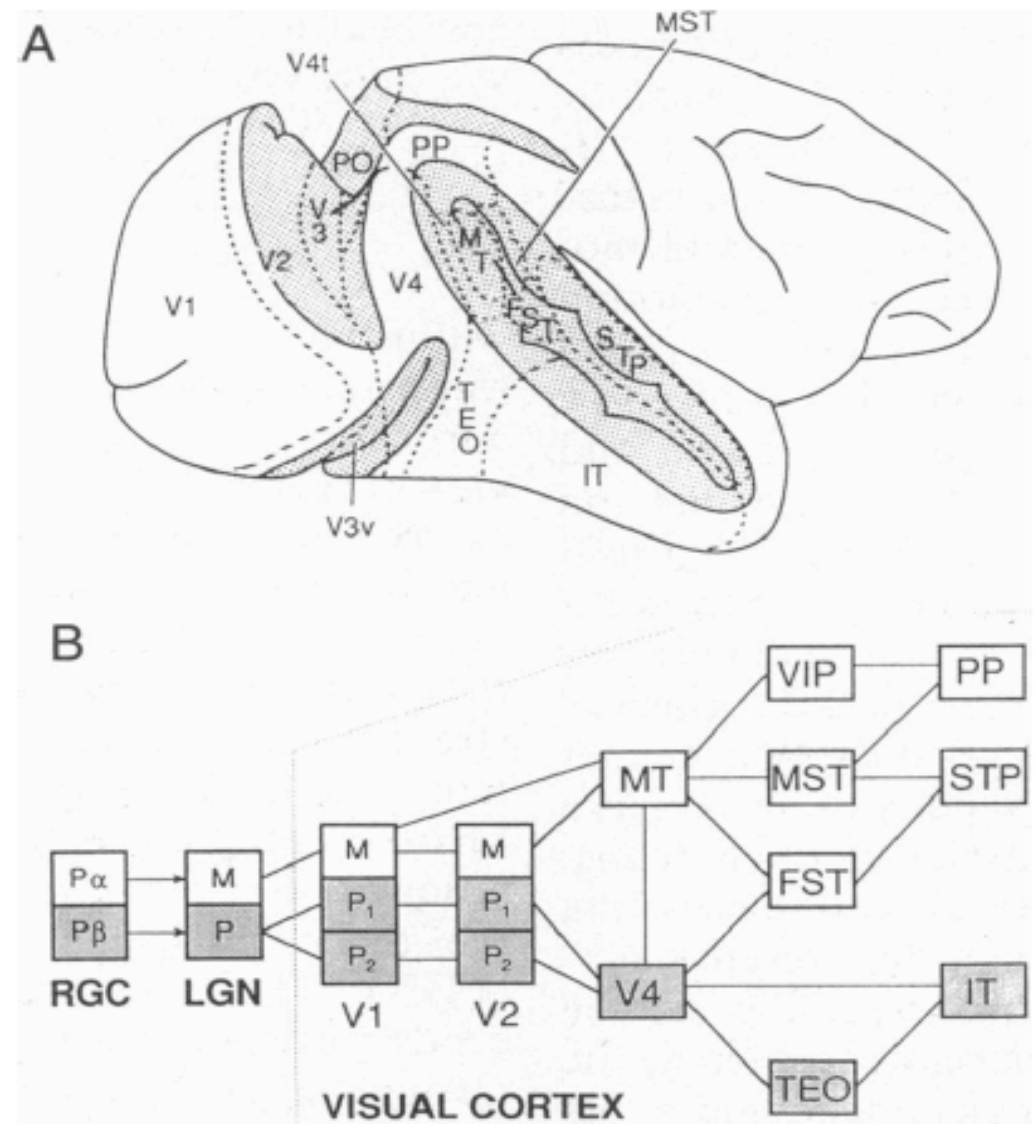
Movimento

O movimento é, provavelmente, a mais importante e poderosa dimensão da visão. Praticamente tudo que é de interesse no mundo visual se move.

- Detecção / segmentação de objetos
- Extração da forma a partir do movimento
- Extração da 3D (ex: paralaxe)
- Determinação do tempo colisões
- Controle da postura, da locomoção e dos movimentos oculares
- Percepção de causalidade (processos cognitivos)

O sistema visual



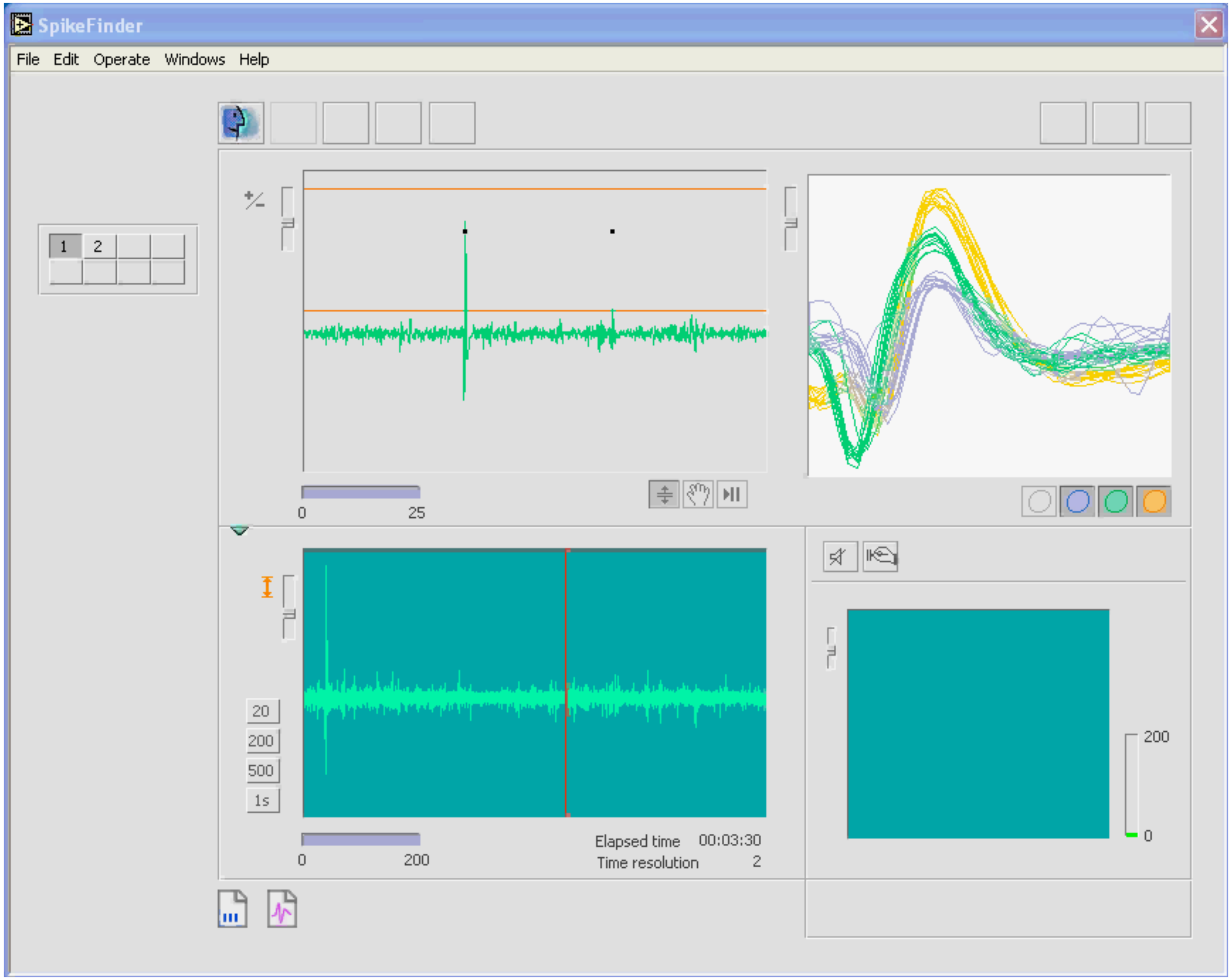


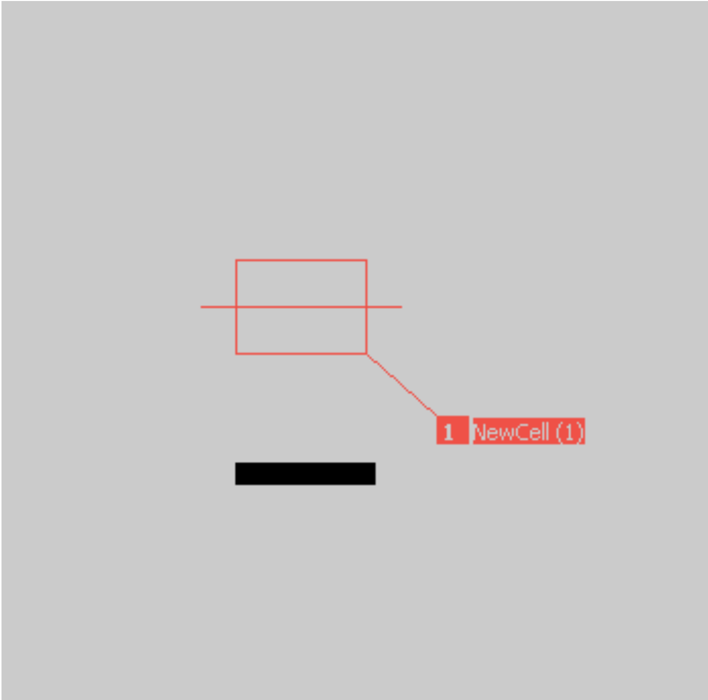
Albright & Stoner (1995) – *PNAS*

Retina e LGN

Hubel & Wiesel (1965) – *Harvard Medical School, Cambridge, MA*

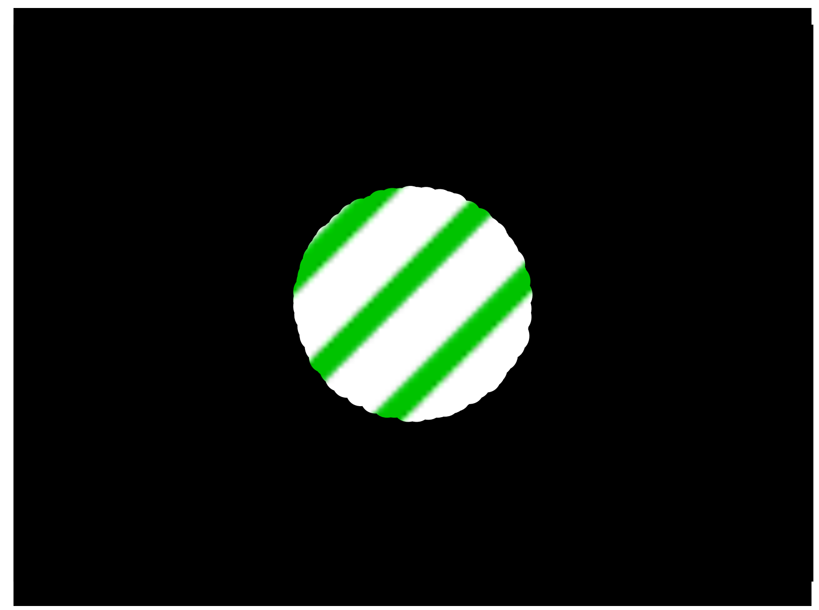
V1 (ou wulst)

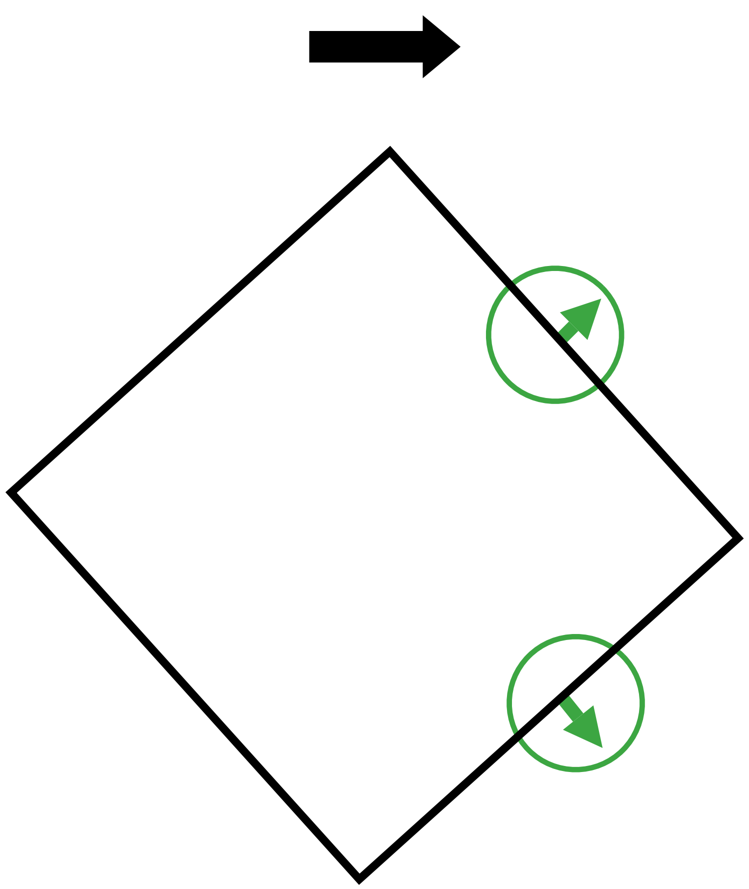


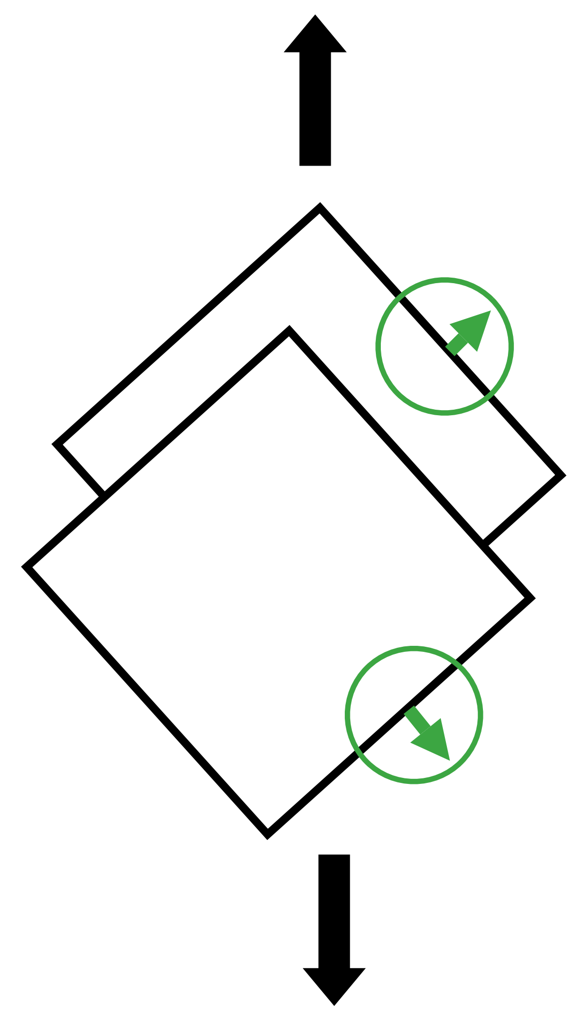


Pinto, Gabriele & Baron - *Laboratório de Neurodinâmica - UFMG,* 2007

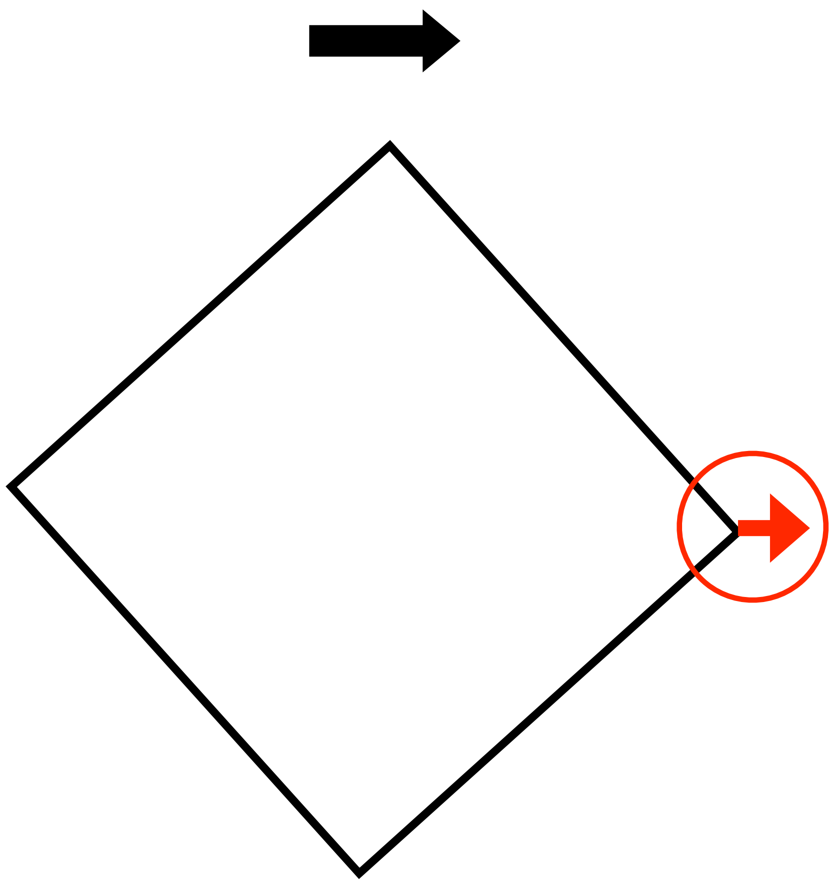
O problema da abertura...

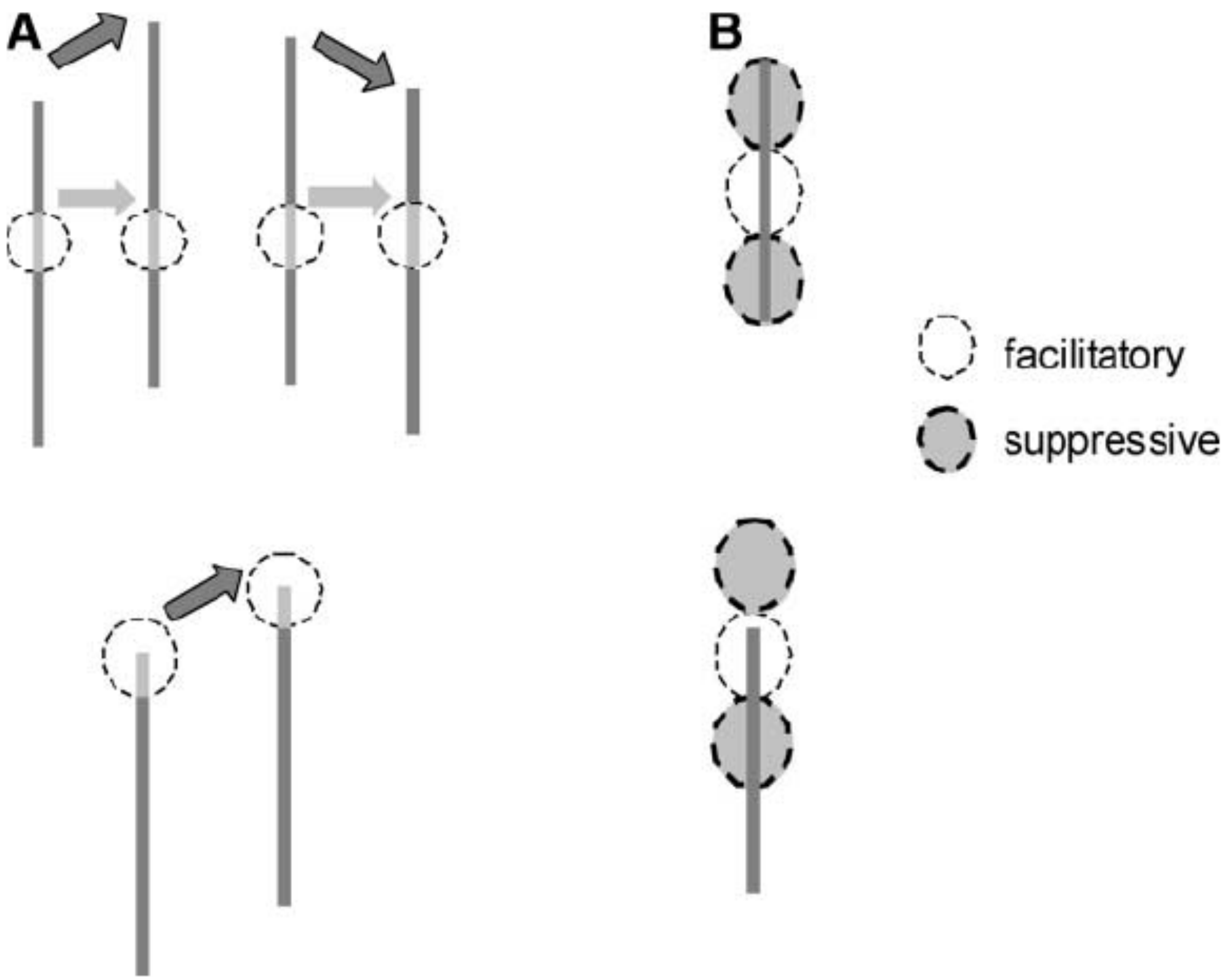






... e uma possível solução

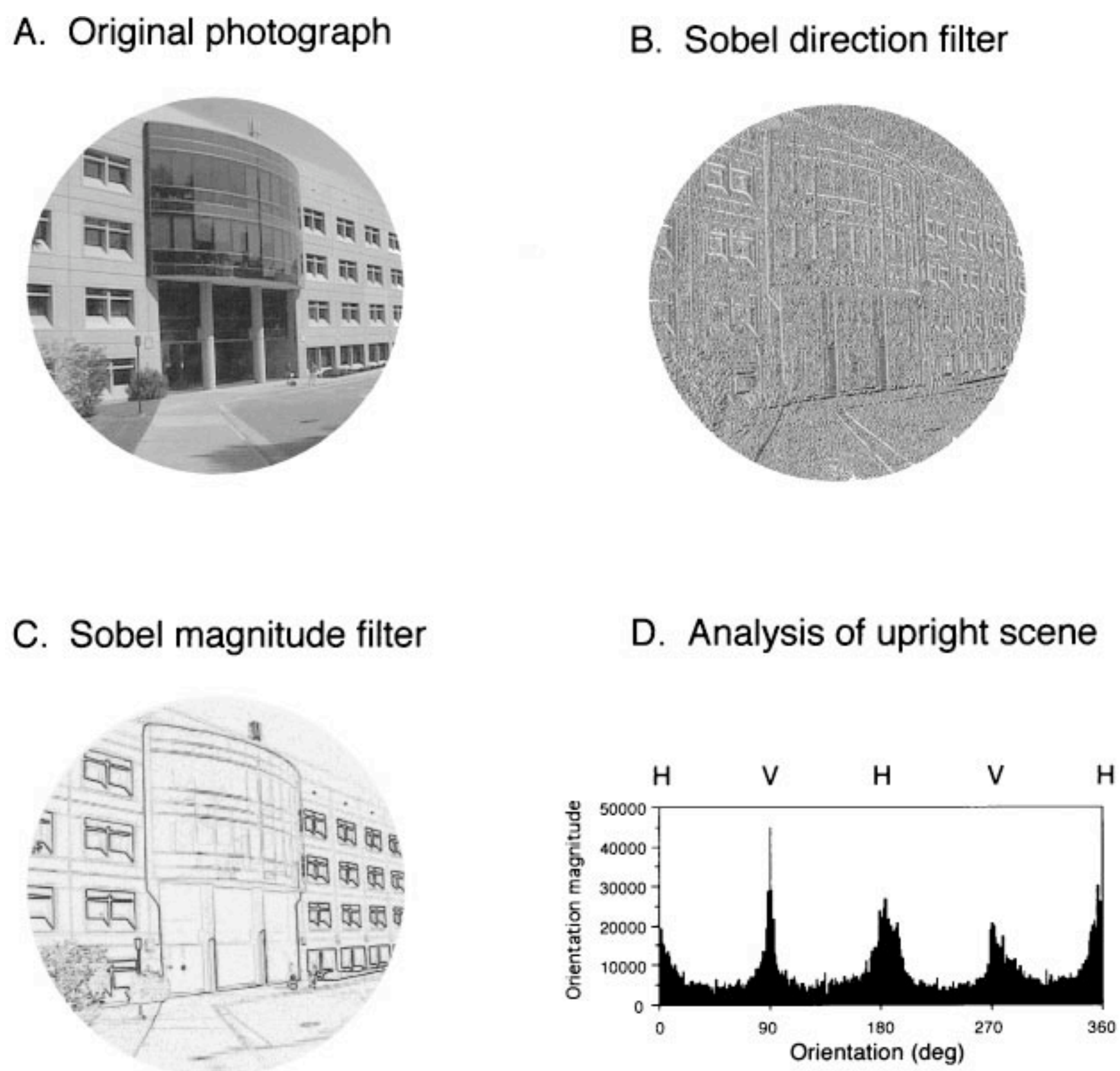




Pack et al. (2003) *J Neurosci*

Cenas naturais

Foto: Gustav Sprecht



Coppola et al. (1998) – *PNAS*

Metodologia

Seqüência natural

Seqüência urbana

Detector de Harris

1) Calculation of partial image derivatives along the x and y dimensions:

$$X = I \otimes [-1 \;\; 0 \;\; 1] = \partial I / \partial x$$

$$Y = I \otimes [-1 \;\; 0 \;\; 1]^T = \partial I / \partial y$$

2) For a given Gaussian kernel *w* given by:

$$w_{u,v} = \exp - (u^2 + v^2) / 2\sigma^2$$

With $\sigma = m/5$, where *m* is the size of *w* in pixels [23], a covariance 2x2 matrix of the form

$$\begin{pmatrix} A & C \\ C & B \end{pmatrix}$$

Is calculated, where

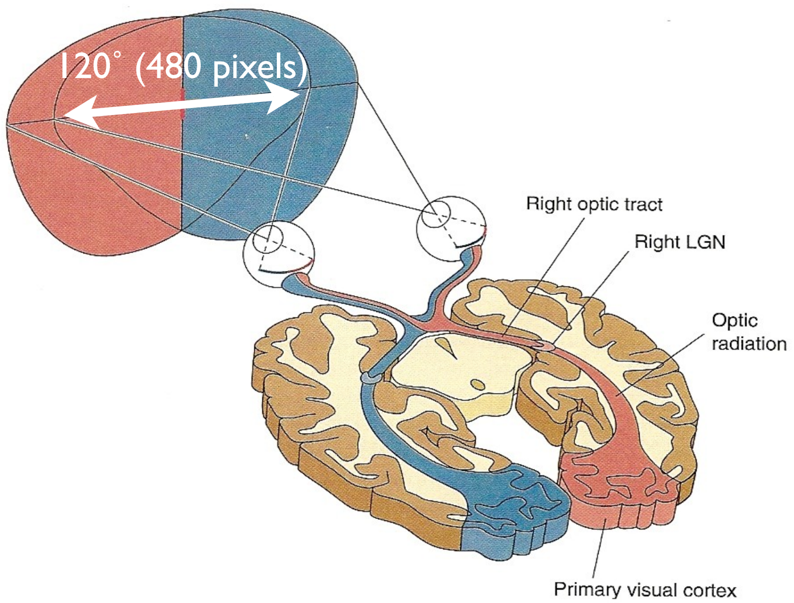
$$\begin{aligned} A &= X^2 \otimes w \\ B &= Y^2 \otimes w \\ C &= (XY) \otimes w \end{aligned}$$

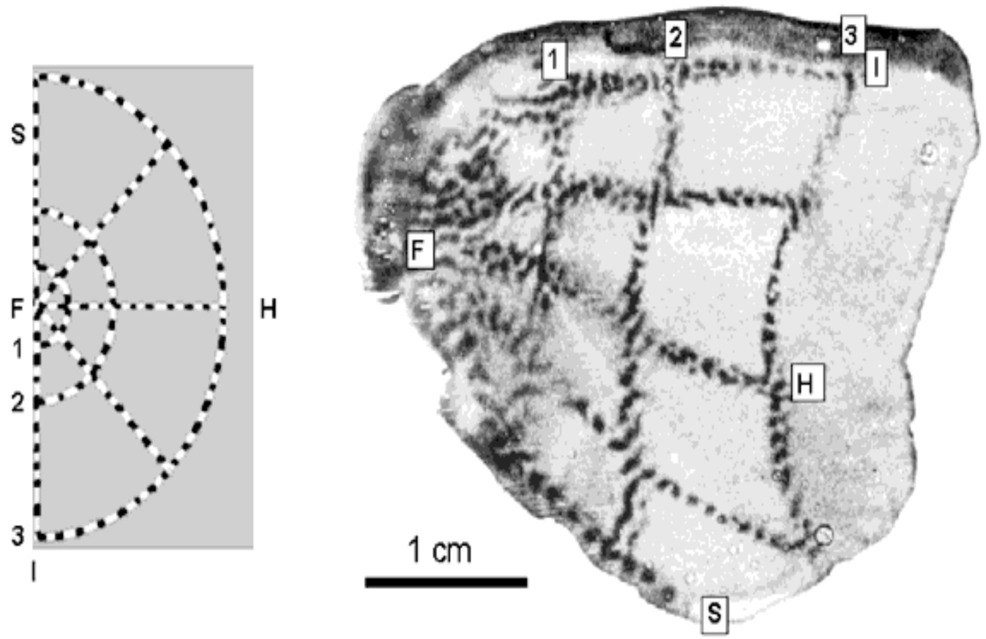
The matrix eigenvalues, λ_1 e λ_2, are then calculated from the matrix trace and determinant

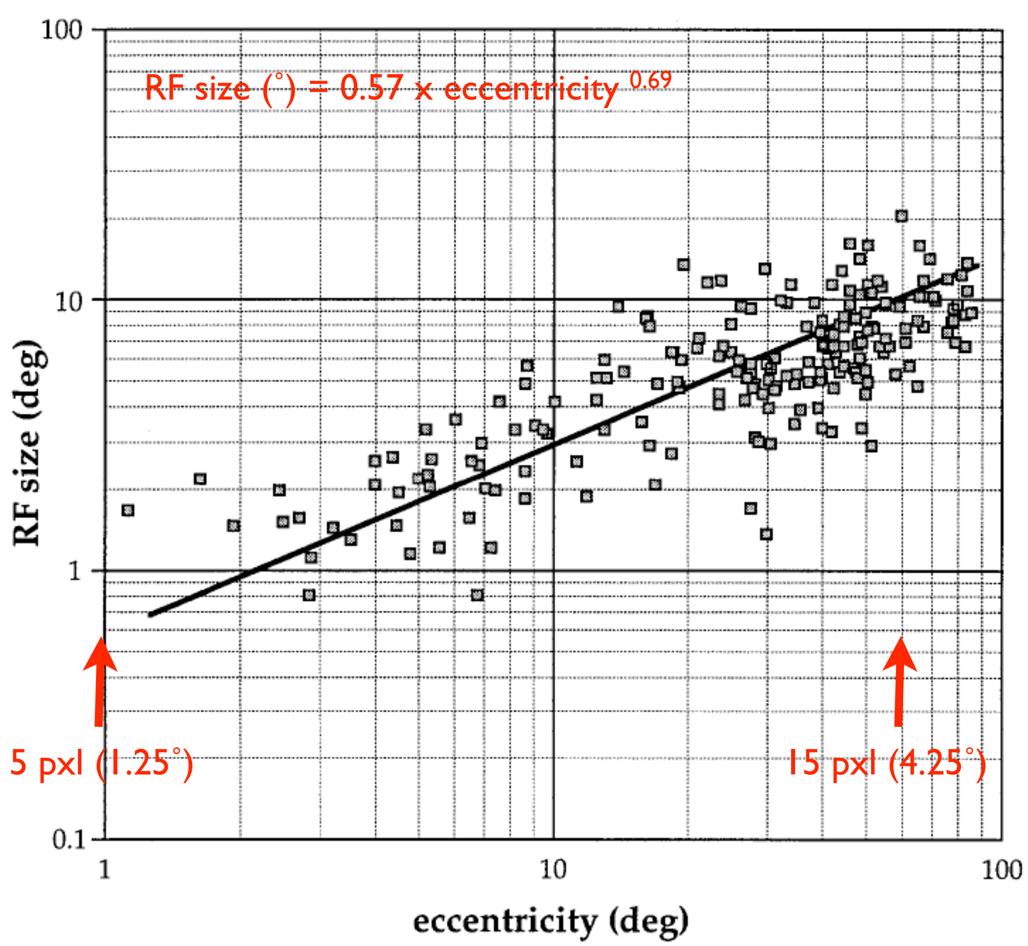
$\lambda_1 \approx \lambda_2 \approx 0$ indicates absence of covariance, which in turn indicates a uniform region;
$\lambda_1 >> \lambda_2 \approx 0$ indicates the presence of an edge, as there is significant covariance only along one dimension;
$\lambda_1 > \lambda_2 > 0$ indicates the presence of a corner, by virtue of significant covariance along both dimensions.

Harris & Stephens (1988) *Proceedings of the 4th Alvey Vision Conference*

Kernel de tamanho variável







Rosa et al. (1997) *J. Neurophysiol.*

Classificação dos campos (RFs)

Receptive fields (regions centered around a pixel, with the size of the kernel) were classified into three categories that relate to the aperture problem:

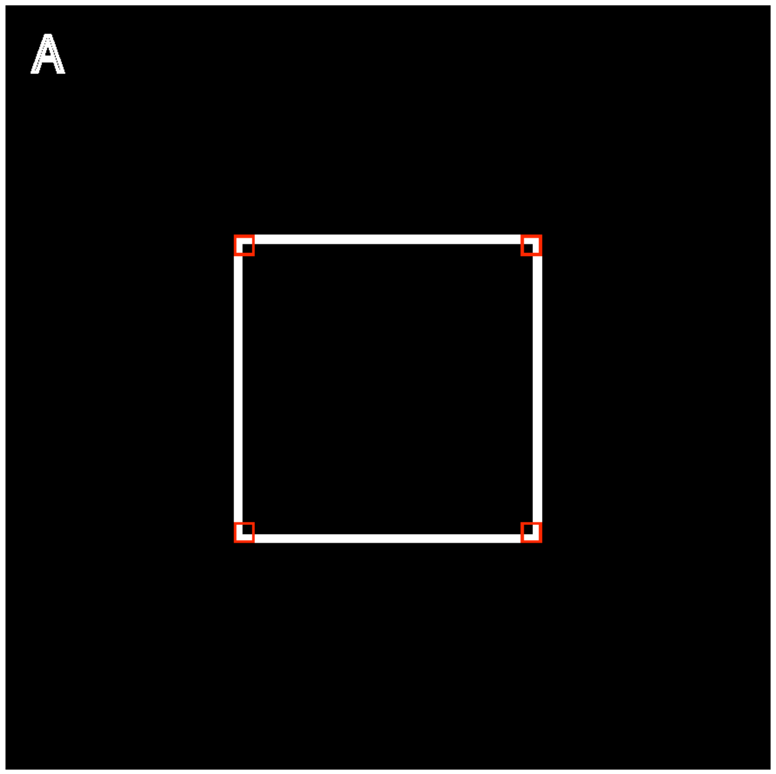
1) **uniform field**, if the output of the algorithm was equal for a given receptive field in all frames;

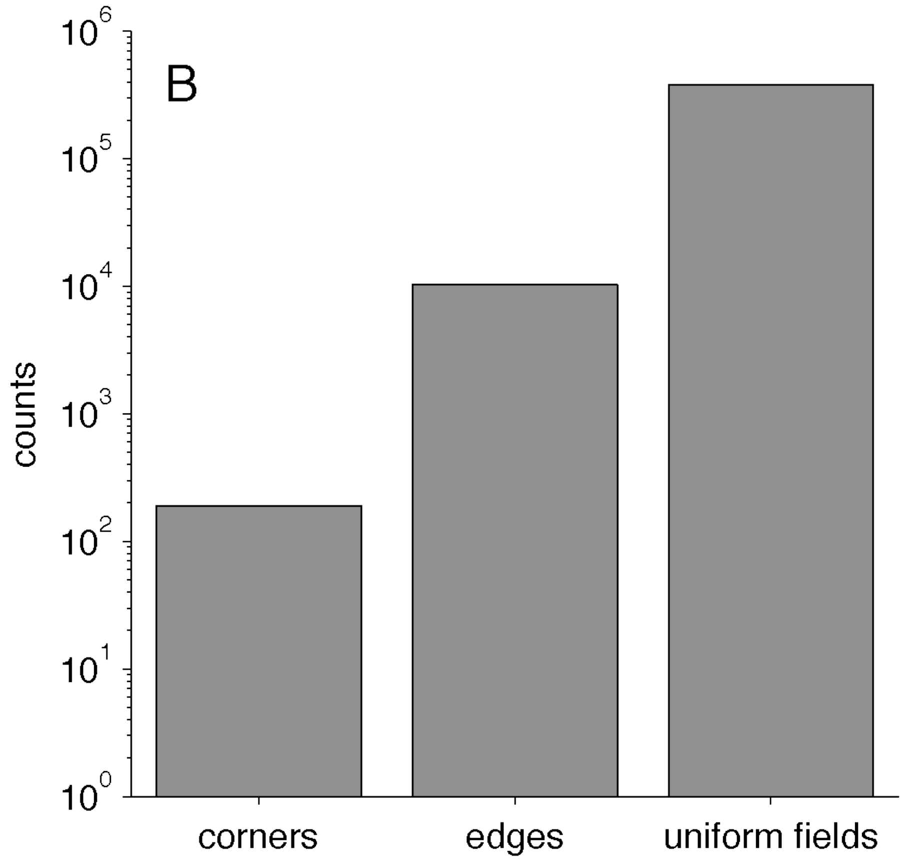
2) **edge**, if an edge was present in at least one frame and at most the total number of frames minus one (meaning that it moved), and no corners were present in none of the frames;

3) **corner**, if a corner was present in at least two frames and at most the total number of frames minus one.

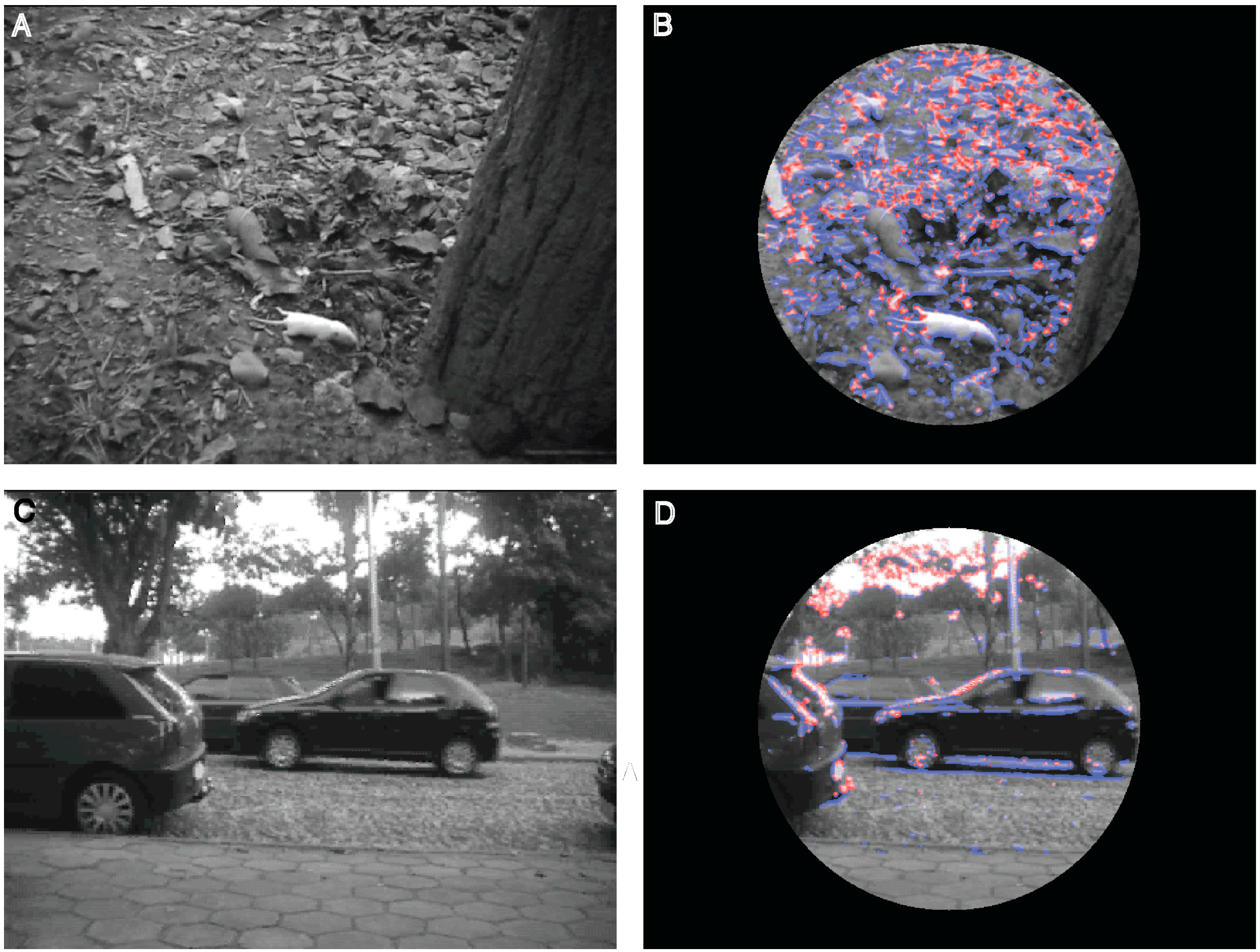
Resultados

Testes c/ imagens sintéticas



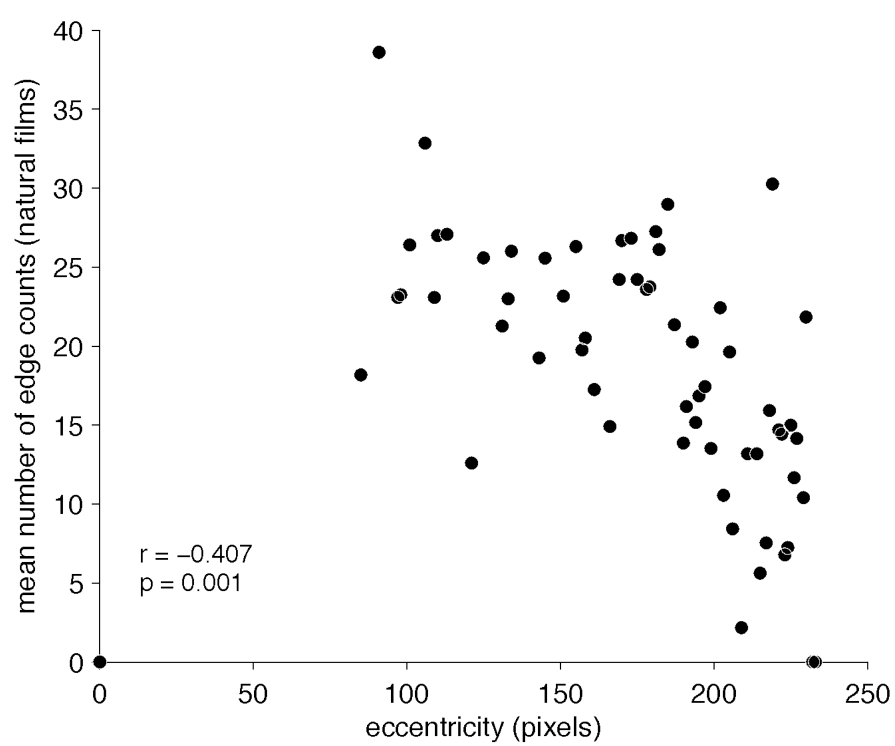


Exemplo de output

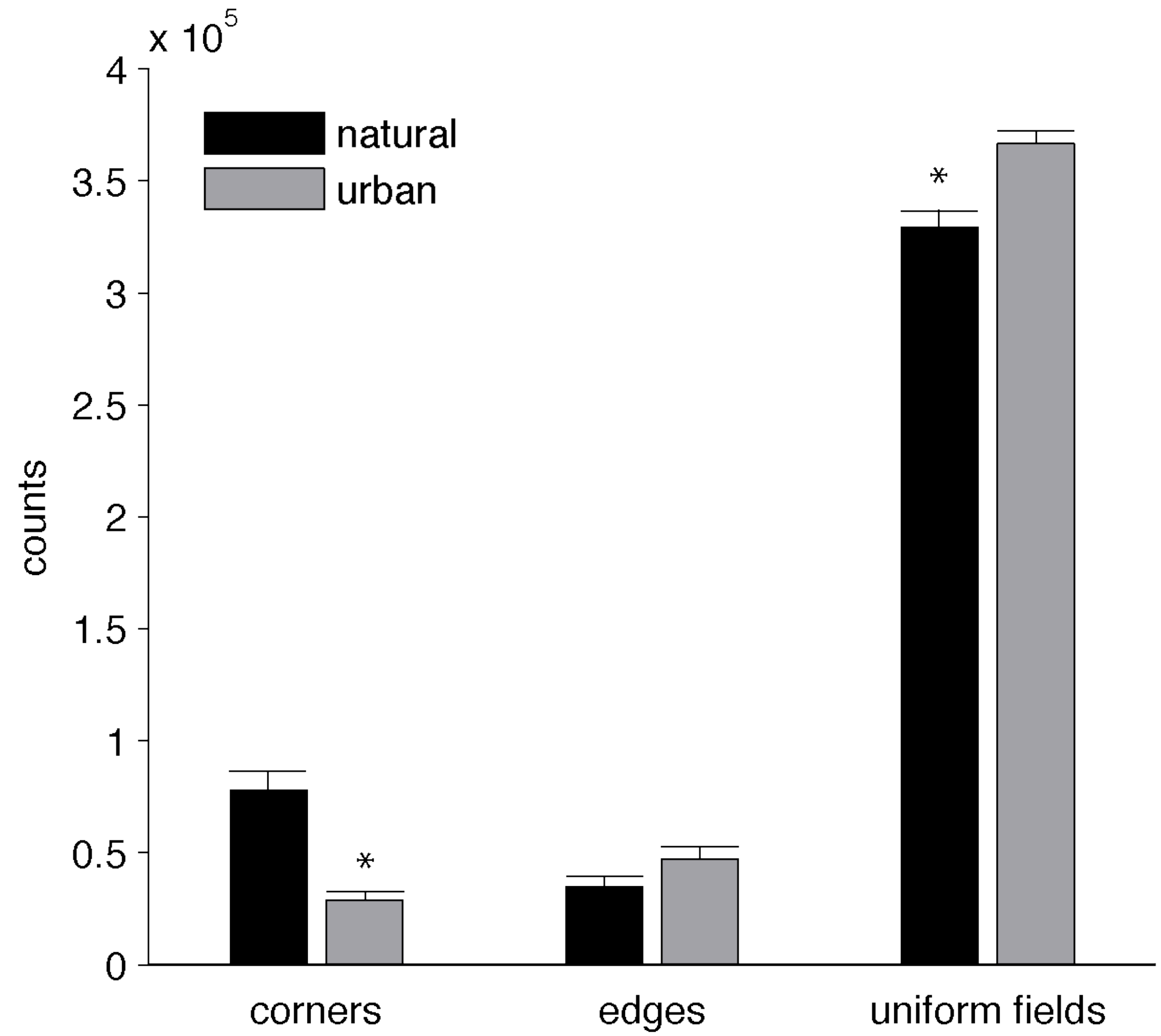


N° de bordas e quinas

	Corners		Edges		Uniform fields	
Film type	**Absolute count (mean ± S.E.M.)**	**Percentage (mean ± S.E.M.)**	**Absolute count (mean ± S.E.M.)**	**Percentage (mean ± S.E.M.)**	**Absolute count (mean ± S.E.M.)**	**Percentage (mean ± S.E.M.)**
Natural	18,907 ± 2,477	11.2 ± 1.6*	49,637 ± 6,997	29.4 ± 4.1*‡	100,490 ± 9,001	59.4 ± 5.3†
Urban	5,174.1 ± 771.5	3.1 ± 0.4	24,924 ± 2,297	14.7 ± 1.4‡	138,940 ± 2,715	82.2 ± 1.6*†



Classificação dos RFs



Conclusões

- To the best of my knowledge, this is the first time that the structure of corners is studied in natural scenes, and that edges and corners are studied in moving scenes.
- Moving corners are more abundant than moving edges in natural scenes, under which our visual system has evolved. Thus, there is a large set of 2D cues to make use of, suggesting that the very existence of end-stopped neurons is yet another piece of evidence of the coupling between visual cortical physiology and the statistics of natural images.
- It is important to note that this conclusion is contrary to what the analysis of single frames, or static images, could lead to believe (see table 1), which highlights the importance of studying image sequences, besides static scenes.
- This study is an important first step towards the understanding of which environmental pressures have shaped the evolution of edge and corner processing by biological motion detection systems.

Futuros trabalhos

- Adquirir mais filmes (fora do campus).
- Utilizar um sistema de classificação alternativo que leve em conta também terminações de bordas.
- Utilizar análise de fluxo óptico para determinar se uma borda e suas quinas correspondentes se movem em direções diferentes ou iguais (incidência real do problema da abertura).
- Utilizar também a distribuição de tamanhos de campos receptores do macaco rhesus.

Referências

[1] L. Spillmann, "From elements to perception: local and global processing in visual neurons". *Perception*, 1999, vol. 28, pp. 1461-1492.

[2] S. Grossberg, E. Mingolla, and C. Pack, "A neural model of motion processing and visual navigation by cortical area MST". *Cereb.Cortex*, 1999, vol. 9, pp. 878-895.

[3] C. Fennema, and W.B. Thompson, "Velocity determination in scenes containing several moving images". *Comput.Graphics Image Process*, 1979, vol. 9, pp. 301-315.

[4] D. Marr, and S. Ullman, "Directional selectivity and its use in early visual processing". *Proc.R.Soc.Lond B Biol.Sci.*, 1982, vol. 211, pp. 151-180.

[5] D.H. Hubel, and T.N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex". *J.Physiol*, 1962, vol. 160, pp. 106-154.

[6] D.H. Hubel, and T.N. Wiesel, "Receptive fields and functional architecture of monkey striate cortex". *J.Physiol*, 1968, vol. 195, pp. 215-243.

[7] J.D. Pettigrew, "Binocular visual processing in the owl's telencephalon". *Proceedings of the Royal Society of London*, 1979, vol. 204, pp. 435-454.

[8] J. Baron, L. Pinto, M.O. Dias, B. Lima, and S. Neuenschwander, "Directional responses of visual wulst neurons to grating and plaid patterns in the awake owl". *Eur. J. Neurosci.,* 2007, vol. 26, pp. 1950-1968.

[9] G.H. Henry, P.O. Bishop, and B. Dreher, "Orientation, axis and direction as stimulus parameters for striate cells". *Vision Res.*, 1974, vol. 14, pp. 767-777.

[10] P.H. Schiller, B.L. Finlay, and S.F. Volman, "Quantitative studies of single-cell properties in monkey striate cortex. II. Orientation specificity and ocular dominance". *J. Neurophysiol.*, 1976, vol. 39, pp. 1320-1333.

[11] J.A. Movshon, E.H. Adelson, M.S. Gizzi, and W.T. Newsome, "The analysis of moving visual patterns". In: C. Chagas, R. Gattass, and C. Gross (eds), "Pattern recognition mechanisms". *Pontificiae Academiae Scientiarvm Scripta Varta*, 1985, vol. 54, pp. 117-151.

[12] T.D. Albright, and G.R. Stoner, "Visual motion perception". *Proc. Natl. Acad. Sci. U.S.A*, 1995, vol. 92, pp 2433-2440.

[13] H.R. Rodman, and T.D. Albright, T.D. "Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT)". *Exp.Brain Res.*, 1989, vol. 75, pp. 53-64.

[14] J.W. Scannell, F. Sengpiel, M.J. Tovee, P.J. Benson, C. Blakemore, and M.P. Young, "Visual motion processing in the anterior ectosylvian sulcus of the cat". *J.Neurophysiol.*, 1996, vol. 76, pp. 895-907.

[15] R.J. van Wezel, and M.J. van der Smagt, "Motion processing: how low can you go?" *Curr.Biol.*, 2003, vol. 13, pp. R840-R842.

[16] W. Singer, "Synchrony, oscillations and relational codes". In L.M. Chalupa, and J.S. Werner (eds), *The Visual Neurosciences*. A Bradford Book, The MIT Press, Cambrige, MA, 2004, pp. 1665-1681.

[17] N.J. Majaj, M. Carandini, and J.A. Movshon, "Motion integration by neurons in macaque MT is local, not global". *J.Neurosci.*, 2007, vol. 27, pp. 366-370.

[18] C.C. Pack, "The aperture problem for visual motion and its solution in primate cortex". *Sci Prog*, 2001, vol. 84, pp.255-266.

[19] C.C. Pack, M. Livingstone, K.R. Duffy, and R.T. Born, "End-stopping and the aperture problem: two-dimensional motion signals in macaque V1". *Neuron*, 2003, vol. 39, pp. 671-680.

[20] E.P. Simoncelli, and B.A. Olshausen, "Natural image statistics and neural representation". *Ann. Rev. Neurosci.*, 2001, vol. 24, pp. 1193-1216.

[21] W.S. Geisler, "Visual perception and the statistical properties of natural scenes". *Ann. Rev. Psychol.,* 2008, vol. 59, pp. 167-192.

[22] C. Harris, and M. Stephens, "A combined corner and edge detector". *Proceedings of the 4th Alvey Vision Conference*, 1988, pp. 147-151.

[23] E. Trucco, and A. Verri, *Introductory Techniques for 3D Computer Vision*, Prentice-Hall, New Jersey, NJ, 1998.

[24] R. van der Willigen, "On the perceptual identity of depth vision in the owl". Tese de Doutorado. Universidade de Aachen, Alemanha, 2000, 155 pp.

[25] P.M. Daniel, and D. Witteridge, "The representation of the visual field on the cerebral cortex in monkeys." *J Physiol (Lond)*, 1961, vol. 159, pp. 203-221.

[26] R. Gattass, A.P. Sousa, and M.G. Rosa, "Visual topography of V1 in the Cebus monkey". *J Comp Neurol,* 1987, vol. 259, pp. 529-548.

[27] M.G. Rosa, V.A. Casagrande, T. Preuss, and J.H. Kaas, "Visual field representation in striate and prestriate cortices of a prosimian primate (*Galago garnetti*)". *J Neurophysiol,* 1997, vol. 77, pp. 3193-3217.

[28] D.M. Coppola, H.R. Purves, A.N. McCoy, and D. Purves, "The distribution of oriented contours in the real world". *Proc Natl Acad Sci USA*, 1998, vol. 95, pp. 4002-4006.

[29] W.S. Geisler, J.S. Perry, B.J. Super, and D.P. Gallogly, "Edge co-occurrence in natural images predicts contour grouping performance". *Vision Res*, 2001, vol. 41, pp. 711-724.

[30] D.W. Dong, and J.J. Atick, "Statistics of natural time-varying images". *Network - Comp Neural*, 1995, vol. 6, pp. 345-358.