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

Lucas Pinto
Lab. of Neurodynamics - Dept. Physiology and Biophysics - ICB/UFMG
lucasmpinto@yahoo.com.br

Abstract

The correct signaling of motion velocity depends on the solution of a problem that affects biological and artificial motion detection systems alike: the aperture problem, whereby it is only possible to perceive the orthogonal component of the velocity of an edge seen through a small aperture. One of the ways to overcome this problem is by using two-dimensional cues, such as those provided by corners. In fact, a subset of primary visual cortical cells seems to use precisely this strategy. Thus, the incidence of edges and corners in natural scenes, under which biological visual systems evolved, can be a good indicator of the magnitude of the aperture problem in primary visual areas. In the present study, I sought to determine this incidence using the Harris edge and corner detector with kernels that varied in size mimicking the variation of cortical receptive field sizes with eccentricity. The results suggest that the magnitude of the aperture problem may be much smaller in natural scenes when compared to urban scenarios, as they contain a higher number of moving corners.

1. Introduction

Motion is probably the most important and powerful dimension of vision. Almost anything of interest in the visual world moves. Stimuli that share approximately the same velocity are seen as a unit, a figure which is easily parsed from background, according to the Gestalt factor of common fate [1]. For instance, an animal that is perfectly camouflaged while still will appear as soon as it moves. Thus, motion can connect stimuli that are widely distributed throughout the visual field, overriding proximity and similarity as factors of perceptual grouping.

The neural implementation of an unequivocal association between an object and its motion, however, is far from being a trivial task, as it depends on the solution of two complementary problems: integration of moving contours belonging to the same object and segregation of the motion of distinct objects [2]. These problems, in turn, are corollary to the so-called aperture problem, which arises due to the nature of receptive fields in the primary visual cortex [3,4]. In both mammals and owls, a large fraction of neurons in this structure presents at least some degree of motion direction selectivity [5,6,7,8]. Nevertheless, their receptive fields are typically small. Hence, individual local motion signals coded by these neurons are often inherently ambiguous, because the direction of motion to which a given cell responds is always orthogonal to the orientation of the contour [9,10], and may differ from the global motion direction of the object. These ambiguities somehow need to be overcome for the operations of integration and segmentation to be carried out properly.

A vast body of psychophysical and electrophysiological evidence suggests that the processing of visual motion occurs in at least two stages [11,12]. In mammals, the first of these stages would take place in the primary visual cortex, the ambiguous local motion signals of which would be transmitted to another area, where a subsequent process of integration and/or segmentation in space and time would occur, thus forming a globally coherent motion perception. This second level of analysis is believed to take place in hierarchically superior cortical visual areas, such as the medial temporal area (MT) in primates [11,13] and the anterior ectosylvian cortex (AEV) in cats [14]. However, the exact neural mechanism of integration is still unknown. In fact, it is not even consensual whether such two or more processing stages are indeed necessary for an integrated perception of visual motion [15,16,17].

Indeed, the aperture problem is not universal in primary visual areas, since non-oriented stimulus features, such as corners and contour terminations, may be present within the receptive field of a neuron. In theory, such two-dimensional features of a stimulus can contain unambiguous information on its motion [18,19]. In accordance with this notion, there is a subset of primary visual cortex cells that are suited for the use of such two-dimensional cues: the end-stopped cells (Hubel and Wiesel's hypercomplex cells), which respond better when line terminations fall inside their receptive fields [6]. Pack *et al.* (2003) [19]

demonstrated that this sub-population of neurons is indeed capable of signaling motion regardless of stimulus orientation -- in other words, they solve the aperture problem.

Although there are experimental demonstrations both of the occurrence of the aperture problem and its resolution by a subset of cells in the primary visual cortex, the proportion and relative contribution of these two phenomena to motion signaling within this area is presently unknown. Actually, not even the relative incidence of contours and corners/line terminators in moving natural scenes seen through apertures the size of receptive fields is known, as most studies use synthetic images. It is likely that these two things are intimately related: sensorial systems are believed to be adapted to the processing of information in their environments, by means of both phylogenetic and ontogenetic mechanisms [20,21]. In the specific instance of vision, mechanisms of neuronal computation would be determined by the statistical structure of natural scenes [20,21]. Therefore, a greater knowledge on the structure of visual scenes might be an important aid in elucidating the neuronal mechanisms underpinning a given perceptual process. The goal of the present study was, thus, to analyze the incidence of edges and corners in sequences of natural scenes seen through apertures that simulated the size and distribution of the receptive fields of primary visual cortical neurons.

2. Methods

2.1. Film acquisition

Images were acquired at a frame rate of 25/s, with a resolution of 576 x 768 pixels in grayscale, with autofocus and manual white balance, using a Sony Handycam DCR-TRV-8E®, with the following specifications:
Type: Camcorder
Sensor: CCD 800 Kpixels
Effective resolution: 400 Kpixels
Media type: Mini-DV
Lens: Carl Zeiss, optical zoom 10x, focal distance: 3.3 - 33 mm.

The films were captured using the software Final Cut Express® on an Apple MacBook® with a 2.16 GHz Intel Core 2 Duo Processor and 2 GB RAM, and converted to '.avi' format using QuickTime Pro®.

Several sequences of scenes considered by the author as being representative of natural scenes, with non-controlled ambient light, and divided into two groups: urban environments (with man-made structures, e.g. buildings and cars) and natural environments (trees, foliage, animals etc). All sequences were acquired from different locations at the University campus. Motion in the image sequences could be in three configurations: motion of objects in the scene with static camera; self-motion (of the camera against a static or semi-static scene); or both. Whenever possible, the same scene was shot from at least three different distances (as measured from a chosen reference point), typically of 1, 2 and 4 m, to control for the possible effects of distance on the incidence of edges and corners. One or more short segments (length of 3 s, or 75 frames) were extracted from each sequence for the analysis.

2.2. Edge and corner detection

All frames from each sequence were submitted to a Harris corner and edge detector [22], programmed and executed in Matlab 7.5 (The Mathworks, Cambridge, MA, EUA) running on Microsoft Windows XP®. Briefly, the algorithm consists of:

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

$$A = X^2 \otimes w$$
$$B = Y^2 \otimes w$$
$$C = (XY) \otimes w$$

The matrix eigenvalues, λ_1 e λ_2, are then calculated from the matrix trace and determinant [22]. From these two values, an edge, a corner or an uniform field is detected according to the following relations:
$\lambda_1 \approx \lambda_2 \approx 0$ 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.

In the algorithm, the decision between edge and corner is made by means of a user-defined threshold τ, typically set at 20 intensity levels in the present study. Edge detection was based on the fact that the higher

eigenvalue represents edge intensity, while its associated eigenvector encodes edge direction [22,23].

2.2.1. Image and kernel size. Only a circular window of the image with 480 pixels in diameter was analyzed, so as to simulate the shape of human visual field. A ratio between number of pixels and degrees of visual angle was assumed so that the whole circular window would be equivalent to 120˚, which is the approximate size of the binocular field (Arditi, 1986 apud van der Willigen, 2000 [24]). This, each degree of visual angle corresponded to four pixels..

The size of the convolution kernel varied according to pixel position, in a way that simulated the change in receptive filed size observed in the primary visual cortex as one goes from the center to the periphery of the visual field. In primates, the representation of different regions of the visual field is not uniform: the central 10˚ occupy roughly half of the primary visual cortex surface [25,26,27]. Furthermore, receptive field size increase from center to periphery following a logarithmic relation [25,26,27].

In the present study, the specific relation between kernel size and eccentricity in pixels was established following the mathematical relation between receptive field (RF) size and visual field eccentricity found experimentally for the prosimian primate *Galago* by Rosa *et al.* [27]:

$$\text{RF size } (^\circ) = 0.57 \times \text{eccentricity}^{0.69}$$

Minimal kernel size was set to 5 x 5 pixels, in order to avoid aliasing in the sampling of the Gaussian function [23], and size at maximal eccentricity (240 pixels, or 60˚ of visual angle) was, according to the relation above, of 13 x 13 pixels. These values correspond to 1.25˚ and 4.25˚ of visual angle, respectively, which is in accordance with the values found in the *Galago* [27]. Edge and corner suppression was not carried out, so as to allow the detection of the same feature by more than one convolution window, similarly to what happens in the primate visual system.

The algorithms used in this study are available on the web at http://www.verlab.dcc.ufmg.br/doku.php?id=cursos:visao:2008-1:grupo06:index

2.3. Data and statistical analysis

For each image sequence, receptive fields (regions centered around a pixel, with the size of the kernel, as previously explained) were classified into three categories that relate to the aperture problem 1) **uniform field**, if this was the output of the algorithm f or 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; and 3) **corner**, if a corner was present in at least two frames and at most the total number of frames minus one. In addition to this, the incidence (in number of frames) of edges, corner and uniform fields was quantified for each region (receptive field).

All data were tested for normality using the Lilliefors' normality test. Normally distributed variables were analyzed with Student's t-test and/or Analysis of Variance (ANOVA), while their non-parametric correspondents, namely Wilcoxon's rank sum test and Kruskal-Wallis, were used in the absence of normal distribution. Correlations were evaluated with the Pearson's correlation coefficient. Significance level considered in al tests was 0.05. All the analyses were carried out in Matlab 7.5 (The Mathworks, Cambridge, MA, EUA).

3. Results

3.1. Algorithm validation

Before being used to analyze the acquired image sequences, the implemented algorithms first underwent a series of tests with simple synthetic image sequences. In all the tests, the algorithms' outputs fitted exactly to the expected values based on the known image features and number of frames. Fig. 1 shows an example of this. The left-hand side image is a close-up of a frame of one of the synthetic sequences used, in which the white square moved against a black background. In this case, the edge and corner detector was run with edge and corner suppression to facilitate visualization. As the bar plot in Fig. 1B shows, the number of receptive fields classified as "corner" equals the number of corners in a frame (centers of red squares in 1A) multiplied by the number of frames (48). Running the algorithm without suppression yields similar results, with a difference in scale, as expected (data not shown).

3.2. Image sequences

A total of 12 sequences of natural scenes and 12 of urban scenes were analyzed. Fig.2 shows representative examples of frames from a natural (Fig. 2A) and an urban (Fig. 2C) sequence, and their respective algorithm outputs (without corner and edge suppression), where blue pixels indicate edges and red pixels indicate corners. It is clear from the pictures that the natural setting contains both more corners (red pixels) and edges (blue pixels) than the more urban scenario. This was confirmed when all image frames were analyzed: as shown in Table 1, average corner and edge counts per movie frame were larger in natural scenes. Two-way ANOVA revealed differences among edge, corner and uniform field counts (degrees of freedom (df)=2, p=0) and a significant interaction between natural and urban scenes (df=2, $p<0.0001$). In

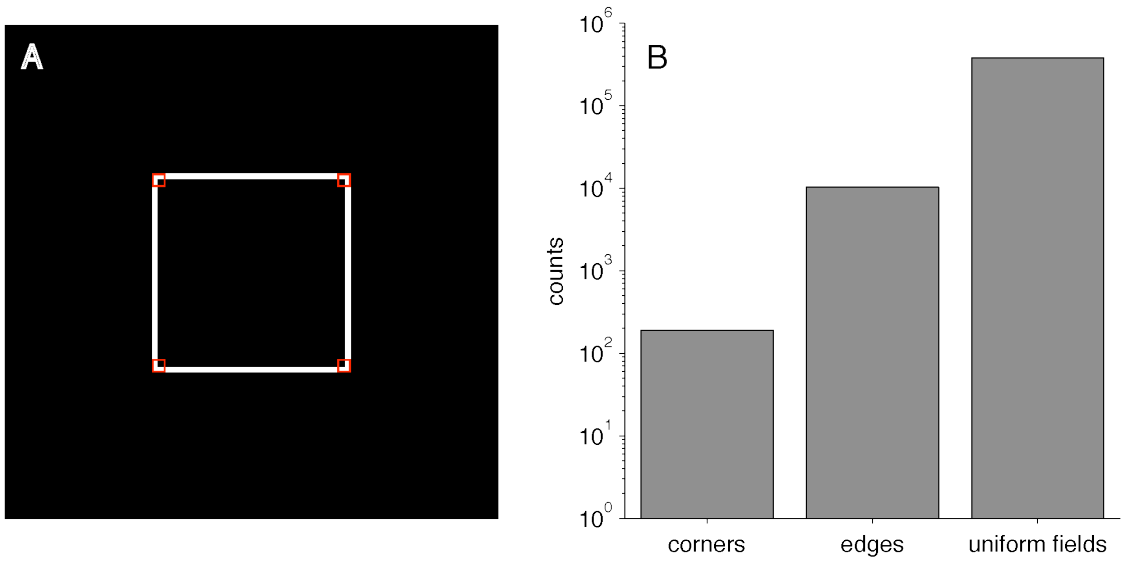

Figure 1. A) Close-up of the first frame of a test synthetic sequence. Red squares indicate corners. B) Counts (y axis, in logarithmic scale) of receptive fields classified as corners, edges and uniform fields for the entire sequence.

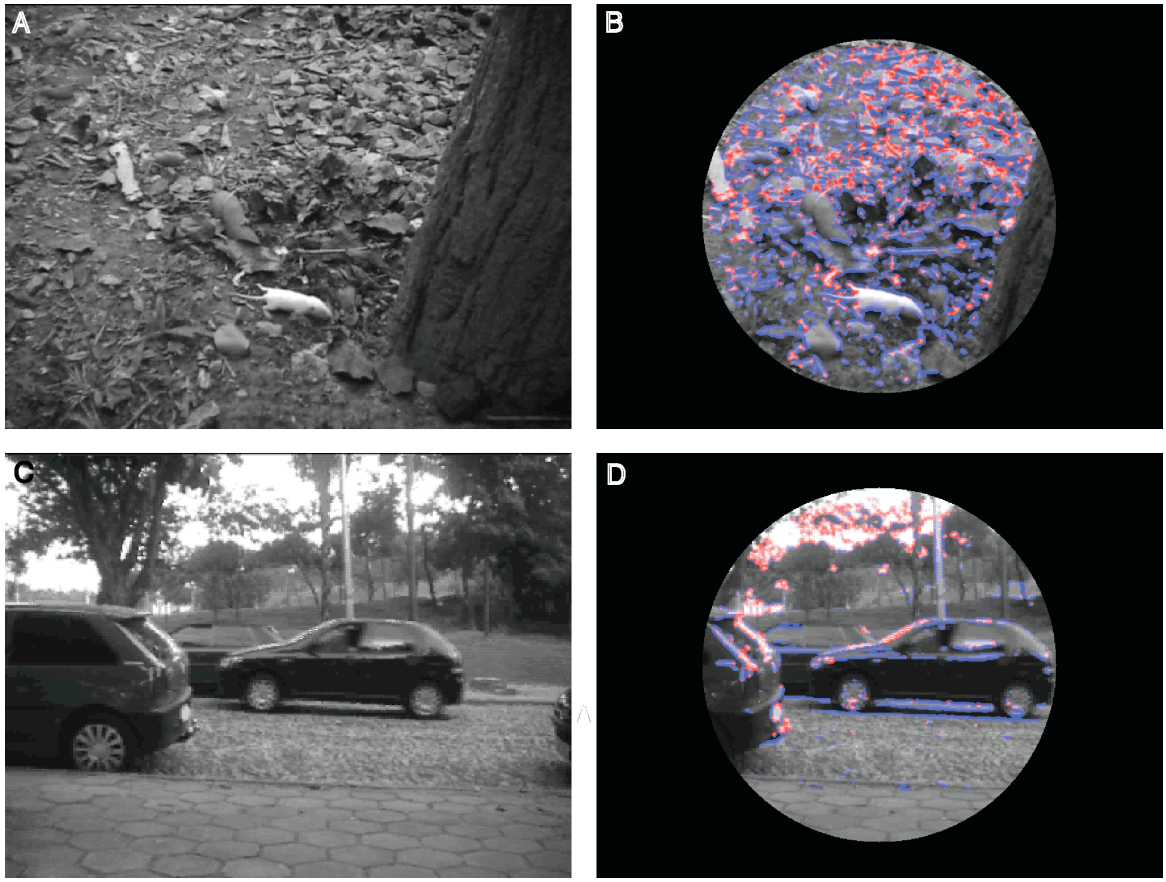

Figure 2. Examples of frames from a natural (A) and an urban (C) sequence and their corresponding edge and corner detector outputs (B and D), whithout suppression (see text for details). Blue pixels indicate edges and red pixels indicate corners. Circular window diameter is 480 pixels.

Table 1. Average corner, edge and uniform field counts per frame

	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*†

Table 1. In both natural and urban films, average edge counts per frame are higher than corner counts (†, $p<0.01$), and uniform fields are more common the the latter two (‡, $p<0.001$). Corner and edge counts are higher in natural than in urban scenes (*, $p<0.05$), while the opposite is true for uniform fields (*).

both natural and urban sequences, average uniform field counts per frame were significantly larger than those of edges (Wilcoxon's rank sum, p=0.0006 and 0.0004, respectively for natural and urban) and corners (two-tailed unpaired Student's t-test, $p<0.0001$ for both); and that edges were significantly more common than corners (Wilcoxon's rank sum, p=0.0035 and 0.0004, respectively). Natural and urban scenes differed significantly in the average count per frame of all categories, confirming that indeed individual frames of natural scenes are more abundant in both corners (Student's t-test, p=0.00009) and edges (Wilcoxon's rank sum, p=0.019), while urban scenes had, on average, more uniform fields (Student's t-test, $p<0.00001$).

3.2.1. Receptive field classification. Solely the incidence of corners and edges in individual frames, however, does not tell much about the incidence of the aperture problem *per se*, as the problem is eminently one that arises with stimulus motion. The classification of "receptive fields" into "corner," "edge" and "uniform field" with the criteria explained in the methods session, on the other hand, is more directly related to the problem, as RFs with the same feature over all frames (indicating lack of motion) were classified as uniform fields. Fig. 3 is a bar plot showing the average distribution among the receptive field categories for natural and urban image sequences. Two-way ANOVA revealed significant differences among categories for both groups (p=0, df=2) and a significant interaction between the groups (p=0.0004, df=2), indicating that counts in each of the categories were not equal for urban and natural films. Individual comparisons with unpaired two-tailed Student's t-test revealed that, for natural sequences, the average number of RFs classified as "corner" (77,840 ± 9,595, 17.6%) is significantly higher that that of "edge" RFs (35,077 ± 4,798, 7.9%)($p<0.0001$), in contrast with the findings for individual frames. The number of "uniform" RFs (329,450 ± 10,057, 74.5%) is higher than both ($p<0.0001$ versus "corner" and "edge"). Among urban sequences, the opposite was observed: "edge" RFs (47,208 ± 6,073, 10.7%) were significantly more incident than "corner" RFs (28,638 ± 4,231, 6.5%)(p=0.0199). Also in this film type, uniform fields predominated over the other two (366,520 ± 9,073, 82.8%)($p<0.0001$ for both). Individual comparisons between the two film types showed that the number of "corner" RFs in natural sequences is significantly larger than in urban scene sequences ($p<0.0001$), while no statistical difference was observed in "edge" RF counts (p=0.13). The average number of uniform fields, in turn, was higher in the urban sequences

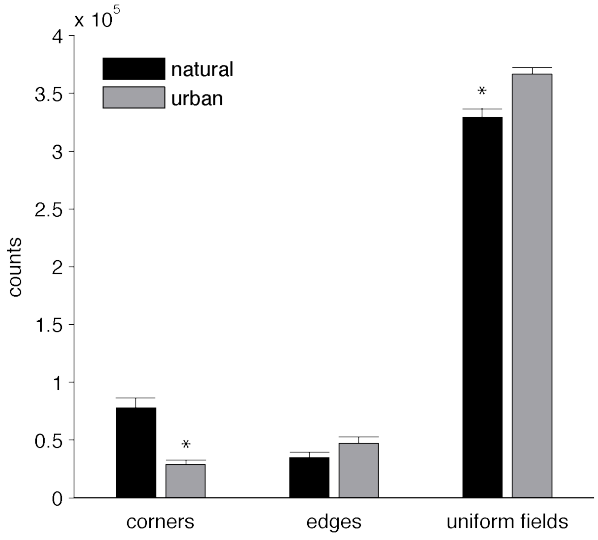

Figure 3. Number of RFs classified as "corner", "edge" or "uniform" for natural (black bars) and urban (gray bars) sequences. Asterisks indicate significant differences between film types. See text for further details.

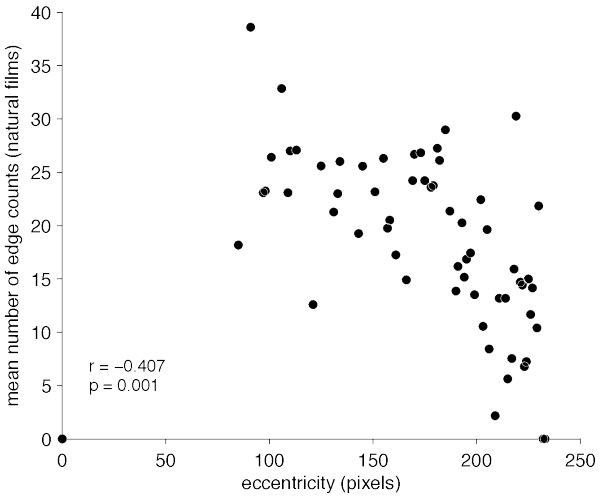

Figure 4. Negative correlation between the mean number of edge counts (y axis) and eccentricity (x axis) in natural scenes.

(p=0.012), showing that the observed differences were indeed probably due to a higher incidence of moving corners in natural sequences.

3.2.2. Influence of eccentricity. I next sought to investigate whether edges, corners and uniform fields are differently distributed with eccentricity. I first analyzed the correlations between counts averaged for all frames and pixel eccentricity. For both urban and natural settings, no correlations were observed for either corners or uniform field counts, nor for edge counts in urban sequences. In natural films, on the other hand, edge counts displayed a significant negative correlation with eccentricity (Pearson's correlation coefficient (r)=-0.407, p=0.001), meaning that edge counts were higher in the more central portion of the frames (see Fig. 4). It is possible that this result may have been a consequence of the sampling technique, whereby objects of interest would have been preferentially sampled in the central part of the image. This seems unlikely, however, as the films were captured under a variety of scene/camera settings; as this distribution was not observed for any feature other than edges; and as this was observed only for natural sequences, even though the same sampling procedures were adopted for both film types.

To verify whether this observation was also valid for receptive field classification, the visible image circle (see Fig. 2) was divided into two concentric portions: in one instance, the central circle comprised half of the field, while in the other it had a radius equivalent to 10˚ of visual angle. However, in none of these two configurations were the incidences of the three types of RF (corrected by the circle areas) different between the central and the peripheral circles, for neither urban nor natural sequences, suggesting that RF types are evenly distributed across different eccentricities (data not shown).

4. Discussion

In the present work, I provided evidence favoring the notion that the aperture problem may be less severe in sequences of natural scenes as compared to their urban counterparts. The main findings may be summarized as follows: 1) images of static natural scenes contain both more corners and edges than those of static urban scenes; in both types of images, edges are more abundant; 2) in images of static natural scenes (individual frames), edges are negatively correlated with eccentricity, i.e., they are more abundant around the image center; corners are evenly distributed across the image; 3) the RF classification scheme proposed in this study categorized more RFs as "corner" than as "edge" in natural scene sequences, while the opposite was observed in urban sequences; the number of "corner" RFs was significantly larger in the former than in the latter.

4.1. Methodological considerations

4.1.1. Film acquisition. I see two potential weaknesses regarding the film acquisition procedures adopted in this study. First of all, the fact that all sequences were acquired at the University Campus made it more difficult to film "exclusively urban" scenes, as there were trees in all the outdoor scenes. This problem is somewhat relieved by the fact that, in sequences shot with a static camera, features coming from these trees would be classified as "uniform" receptive fields. Secondly, an n of 12 for each film type could potentially be somewhat insufficient. The robustness of the results, however, prompts me to think otherwise: for instance, in all the individual natural sequences, the number of "corner" RFs was quite larger than "edge" receptive fields. All the same, it would be interesting to increase the number of sequences of both types, and acquire some of them outside campus.

4.1.2. Corner and edge detection. An alternative approach to the one used in this study could be the use of a classical Harris corner and edge detector (with fix kernel size) and a subsequent stage of receptive field classification whereby the output of the detector would be viewed through several apertures of varying size, which would classify the receptive field as "corner" also if an edge terminated within it. This would make the classification even closer to what happens in cortical physiology -- end-stopped cells respond best to edges terminating inside their receptive fields [5,6,19]. I intend to use this approach in a follow-up study.

4.1.3. Receptive field classification. The adopted receptive field classification procedure may also suffer from a few drawbacks. Most importantly, it does not take into account whether a feature in given pixel of a given movie frame is the same as in another frame. In other words, it does not tackle the issue of feature correspondence. Thus being, it is possible that some fields were incorrectly classified as "uniform", because the presence of a same *type* of feature (i.e. edge or corner) in a same pixel does not necessarily mean that the *same* feature is present, which means that the number of "corner" and "edge" fields may have been underestimated. Of course, there is no reason to believe that urban or natural sequences, and corners or edges, have been differently affected by this underestimation, which does not affect the main conclusion. Also noteworthy is the fact the presence of an edge alone does not imply in the occurrence of the aperture problem, because the global motion vector may coincide with the component orthogonal to edge orientation. It would thus be very interesting to approach this particular issue, perhaps using optical flow analysis to compare the motion direction of an edge and its associated corners. This is also a follow-up I propose for this study.

4.1.4. Receptive field sizes. As a final methodological consideration, it is prudent to remember the fact that the specific relation between receptive field size and eccentricity used in the present study was taken from data from the *Galago* [27]. Even though they are qualitatively very similar to data from other primates, such as the macaque monkey, they do have a few quantitative differences. It would be interesting to test whether the results of the present study would hold for other receptive field size distributions.

4.2. Comparison with the literature

Several studies on the statistical structure of natural scenes have been published over the past few years (for reviews, see references 20 and 21). This notwithstanding, the specific issue of the structure of corners and edges has received considerably less attention. In particular, two studies approached the structure of edges in static scenes. One of them [28] studied the distribution of contour orientation in an image databank, as determined by the Sobel edge detector, and discovered that horizontal and vertical edges are more common than those in oblique angles, suggesting that our knowingly better perceptual ability for vertical and horizontal edges may have an ontogenetic and/or phylogenetic origin. Another study [29] was a characterization of the probability of edge co-occurrence in natural static images. The authors showed that edges have a higher probability of being found together if they are co-circular, which is likely to reflect the smooth curves found in natural settings. These results also agreed with a number of psychophysical findings, and effectively showed the physical and statistical counterpart of the "good continuation" principle of *Gestalt*. I stress that these two latter studies were performed with static images. The statistics of time-varying images had already been the subject of a previous study [30], which, differently from the present one, approached the issue exclusively with a frequency-domain analysis.

To the best of my knowledge, however, 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. As in the above-mentioned studies, the present findings may have important implications for our understanding of the perceptual physiology of motion.

4.3. Implications for the aperture problem

Since the initial works of Hubel and Wiesel [5,6], the so-called end-stopped cells (first termed hypercomplex cells) are known to exist in the mammalian primary visual cortex. The name of this subset of neurons comes from the fact they are under strong surround inhibition, i.e., stimuli extending outside their receptive fields strongly inhibits them, to the point that they respond best to edge discontinuities within their receptive fields. In other words, they respond to two-dimensional image features, such that, at least theoretically, they are not subject to the aperture problem. Naturally, this property is very desirable in systems that have to perform feature matching, or correspondence [4] -- and this is precisely what our visual system does, both to match disparities between successive moments in time and between the two eyes. Earlier investigations, however, failed to identify the ability of any primary visual cortical neuron to signal the motion of two-dimensional cues (see, for instance, ref. 11), which was probably due to the nature of the stimuli used [19]. Only more recently has evidence favoring this function of end-stopped neurons emerged: when tested with the proper stimuli, a large fraction of these neurons responds to 2D stimulus motion, regardless of edge orientation [19]. This is where the present study fits in: 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 therefore 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, a knowledge that could certainly benefit the area of computer and machine vision as well.

5. References

[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.

6. Acknowledgements

The author would like to thank Jerome Baron for excellent ideas and fruitful discussions and Marcelo G. O. Dias for help with some of the footage. Lucas Pinto receives a M.Sc. Scholarship from the Brazilian National Council for Scientific and Technological Development (CNPq).