Temporal correlations of orientations in natural scenes

Christoph Kayser, Wolfgang Einhäuser & Peter König
Institute of Neuroinformatics, ETH / UNI Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland.
Contact: kayser@ini.phys.ethz.ch
Keywords: natural scenes, orientations, Gestalt principles, image statistics, temporal coherence

Abstract

The visual system performs complicated operations such as visual grouping efficiently on its natural input. To study this adaptation to natural stimuli we measure spatio-temporal interactions of orientations in scenes with natural temporal structure recorded using a camera mounted to a cat's head. We find long range spatial and long lasting temporal correlations of orientations with collinear interactions being most prevalent and preserved over time. The spatial extent of correlations corresponds to the length of horizontal cortical connections and the temporal duration of the interactions allows co-activation of lateral and bottom up input by the same visual event.

Introduction

In recent years processing of natural stimuli by the visual system received increased attention (cf. [12]). Indeed it was found that early stages of visual processing are specifically adapted to the structure of natural scenes [1,4]. Furthermore, laws for object perception and visual grouping, the Gestalt rules [8,14], can be linked to the statistics of natural scenes. As an example the law of good continuation, favouring collinear arrangements of orientations over parallel, was shown to have a counterpart in the interaction of orientations in still images [7, 9,11]. Similar interactions of orientations are also found in contextual effects in psychophysical experiments [6,10], in surround interactions in V1 receptive fields [6] and in lateral connections in V1 [2,5]. Therefore it is of particular interest to link them to properties of natural scenes. To our knowledge, however, up to now correlations in natural scenes have only been investigated in still images. This neglects the temporal structure and it remains unclear whether these correlations persist on time scales relevant for lateral interactions in the cortex. Given the possibly long delays for tangential connections, correlations must extend over substantial temporal periods in order to fully cover the spatial extent of long-range connections. Furthermore, some of the previous studies did not report filter or correlation scales in units of degrees of visual angle leaving possible links to anatomical scales uncertain. Finally, some of the previous studies used still images captured by humans, therefore introducing a possible artistic or anthropocentric bias.

Here we address these issues and study spatio-temporal interactions of orientations in a large database of natural movies captured by a camera mounted to a cat's head.

Methods

We recorded movie sequences using a removable lightweight CCD-camera (Conrad electronics, Hirschau, Germany) mounted to the head of cats while taken for walks in different local environments like grassland, forest and the university campus. These videos contain a large variety of different speeds and accelerations as a result of the natural movements of the cat. Figure 1 shows four sample images of our database. For this study a total of three animals was used and all procedures are in agreement with national and institutional guidelines for animal care.

Videos were recorded via a cable connected to the leash onto a standard VHS-VCR (Pal) carried by the human experimenter and digitised offline at a temporal resolution of 25 Hz, 320 by 240 pixels (1 pixels $\approx$ 12 min of arc) and 16 bit color depth. For this study videos

were converted to 8-bit gray scale and 12 sequences (about 40000 frames total) were used. Before further processing the images were normalized to zero mean.

The image statistics was investigated using oriented wavelets. Single frames were convolved with pairs of circular Gabor wavelets of 90° relative phase shift. Filters had a envelope of 20 pixel width and a spatial frequency of 7 (1/pixels). The amplitude of the orientation was computed by summing the squared amplitudes of two phase shifted filters and subjecting the result to a square root, resembling a two subunit energy model. At each point the amplitudes of eight equally spaced orientations from 0° (horizontal) to 157.5° were computed. We define the 'prominent' orientation of each point by averaging the amplitude vectors (length = amplitude of filter response, orientation = orientation of the filter) of the eight filters. The resulting vector average has an orientation Θ, defining the prominent orientation of the point, and a length $A(\Theta,x,t)$, specifying the magnitude of the local orientation strength. For computational convenience these orientations Θ were binned into 16 bin between 0° and 180°. The second order statistics of these orientations was calculated assuming translation invariance of natural images. Thus correlations of two prominent orientations Θ_1 and Θ_2 were computed over all pairs of points with the same spatial separation Δx and temporal separation Δt (the mean < > runs over all points (x,t) with prominent orientation Θ_1).

$$C(\Theta_1,\Theta_2,\Delta x,\Delta t)=\frac{\left\langle \left(A(\Theta_1,x,t)-\left\langle A(\Theta_1,x,t)\right\rangle\right)*\left(A(\Theta_2,x+\Delta x,t+\Delta t)-\left\langle A(\Theta_2,x+\Delta x,t+\Delta t)\right\rangle\right)\right\rangle}{\sqrt{\left\langle \left(A(\Theta_1,x,t)-\left\langle A(\Theta_1,x,t)\right\rangle\right)^2\right\rangle}*\sqrt{\left\langle \left(A(\Theta_2,x+\Delta x,t+\Delta t)-\left\langle A(\Theta_2,x+\Delta x,t+\Delta t)\right\rangle\right)^2\right\rangle}}$$

Correlations were computed for temporal lags from $\Delta t=0$ to $\Delta t=30$ frames (1.2 seconds) and on a spatial grid of points spaced about 2 degrees apart. Therefore the kernels overlapped only for the smallest spatial distance used. As a control we also computed correlations using the maximally active orientation at each point instead of the 'prominent' orientation yielding similar results as reported below.

Results

First we investigate temporal correlations at the same point in space. Figure 2A demonstrates that if an orientation is present at one point in time then the amplitude of this orientation in the next frames at the same point is also likely to be high. Temporal correlations are strongest for the cardinal orientations, i.e. horizontal and vertical. For the other orientations correlations decay faster but are still significant over several hundreds of milliseconds (decay time constants for 0°: >1 s, 45°: 490 ms, 90°: 900 ms, 135°: 360 ms). Thus the presence of an oriented segment gives a strong prediction for the orientation at the same point later in time.

Next we look at the two dimensional spatial distribution of correlations as well as correlations of different orientations. Figure 2B shows the correlations between segments of 4 different orientations (0°, 45°, 90°, 135°) situated at different relative locations in the same frame. Iso-orientation correlations (panels on the diagonal) are stronger than cross-orientation correlations. Furthermore the contour lines of the iso-orientation correlations are elongated along the direction of the particular orientation. This shows that collinear structures are more prevalent than parallel shifted contours. Also parallel contours occur more likely than T-junctions since the iso-orientation correlations are at all points stronger than the correlations of this orientation with the orthogonal. An example of how the spatial correlations decay independently of the spatial direction is shown in Fig 2C for the horizontal (90°) orientation. Correlations decay fastest during the first 2 degrees of spatial distance but extend well up to 8 degrees.

Our dataset allows analysing how these spatial correlations evolve over time. Figure 2D shows the same data as in Fig 2B but for segments 400 ms apart in time. The spatial arrangement of correlations is the same as for zero time lag but the amplitudes decayed by a factor higher than 2. For the cardinal orientations again collinear interactions are prevalent.

This is in agreement with Fig 2A which shows that these orientations are very stable over time. Since the oblique orientations are less well correlated over time we would expect that collinearity will here be less prominent for larger time lags. Indeed the contour lines of the correlations for the oblique orientations are more circular symmetric. To quantify these changes over time we measure the aspect ratio (length / width) of the contour lines for the different time lags. Collinearity means a high aspect ratio and a loss of collinearity therefore is accompanied with a decrease in aspect ratio. Using this measure, Fig 3A shows that collinearity is preserved over long temporal lags and is strongest for the cardinal orientations.

To quantify the change in amplitude of the spatial correlations in a different way, we define areas of strong interactions by thresholding correlations. We chose a threshold of 0.4 to ensure that even for zero time-lag only iso-orientation correlations exceed this threshold (Fig. 3B). As expected from Fig 2 the decay times are slowest for the cardinal orientations but independent of the orientation there exist points with strong correlations for at least 280 ms (Fig. 3C).

We performed controls to see how these results depend on the amount of data used. The above data were averaged over our whole data-base. Since one feature of our video sequences is their variety in terms of landscapes etc. we look at the differences between different sequences. In Fig 4 we show the correlations for one oblique orientation (135°). The mean and standard deviation over 12 video sequences is shown in Figure 4A. The error is rather small compared to the correlation values. More importantly, the correlation surface plus minus the error (Fig 4B) shows the same spatial structure as the mean. Also, the distinct pattern of correlations is visible in averages over shorter sequences (data not shown). Thus the distinct patterns of spatial correlations are not introduced by averaging over a large data set.

As a further control, we use filters of a different spatial scale and frequency to measure the orientation content. The filters used for Fig 4C are twice as large as the ones used for the other experiments. The results are basically the same as with the lower frequency filters. Again collinearity is most prevalent. Therefore our results generalize over a wide range of filter parameters.

Discussion

We recorded natural image sequences from a camera mounted to a cat's head closely matching the animal's visual input. Thereby our database circumvents possible artistic or anthropocentric biases introduced in pictures and movies taken by humans. The database contains a large set of different environments, ranging from forest to grasslands and university campus. Furthermore the used sequences were recorded in different seasons and times of day providing a huge variety of lighting conditions. In respect to the temporal analysis it is worth noting that our video sequences contain natural movements of an animal, which might differ considerably from e.g. commercial movies filmed by humans.

In qualitative agreement with previous studies [7,9,11] we find spatial correlations corresponding to the Gestalt laws. For all orientations collinear contours are more prevalent then parallel contours and correlations between orthogonal orientations are weakest. However we find correlations over distances of up to 8 degrees of visual angle (Fig 2). This is considerably larger than distances reported in previous studies. For example Kaschube et al. [7] find that already for small distances correlations are relatively weak (<0.15 in a range from 1° to 4°). However, they do not indicate the size of their kernels in the same units. Sigman et al. [11] report similar correlations using filters of size smaller than 10 min of arc. Our higher correlation could be due to methodological differences to other studies besides the use of different and possibly larger kernels. We computed the 'prominent' orientation of a point by vector averaging the outputs of 8 oriented energy detectors. But correlations computed on these prominent orientations are very similar to correlations computed on the maximally active orientation (data not shown) a method used in [11].

The spatial distances of the correlations reported here fit well with anatomical data on long-range horizontal connections in primary visual cortex. In cat V1 8° of visual angle correspond roughly to 8 mm [13]. This is also the extent of long-range connections which

preferentially connect iso-orientation domains [5] and in some mammals preferentially mediate collinear interactions [2].

In the temporal domain we find long lasting correlations of orientations to extend several hundreds of milliseconds preserving their spatial structure i.e. collinearity. These persist sufficiently long to allow bottom up and long range lateral input to be coactive and driven by the same orientated structure even given the slow speeds of lateral connections reported in [3]. Therefore the spatio temporal interactions of orientations seem to fully cover the range of tangential connections and provide a substrate that could also guide the development of orientation maps and long-range connections in primary visual cortex.

Acknowledgements:

This work was financially supported by the Centre of Neuroscience Zurich, (ZNZ), Honda R&D Europe (Germany) and the Swiss national fund (SNF grant No. 31-65415.01).

References:

[1] J.J. Atick, A.N. Redlich. What does the retina know about natural scenes. Neural Comput. Vol. 4, pp196-210, 1992.
[2] W.H. Bosking, Y. Zhang, B. Schofield, D. Fitzpatrick. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. J. Neurosci. Vol. 17(6), pp.2112-2127, 1997.
[3] V. Bringuier, F. Chavane, L. Glaeser, Y. Fregnac. Horizontal propagation of visual activity in the synaptic integration field of area 17 neurons. Science, Vol. 283, pp.695-699, 1999.
[4] Y. Dan, J.J. Atick, R.C. Reid. Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. J. Neurosci. Vol. 16(10), pp.3351-3362, 1996.
[5] C.D. Gilbert, T.N. Wiesel. Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. J. Neurosci. Vol. 9(7), pp.2432-2442, 1989.
[6] M.K. Kapadia, M. Ito, C.D. Gilbert, G. Westheimer. Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. Neuron, Vol. 15, pp. 843-856, 1995.
[7] M. Kaschube, F. Wolf, T. Geisel, S. Löwel. The prevalence of collinear contours in the real world. Neurocomputing 38-40, pp. 1335-1339, 2001.
[8] K. Kofka. Principles of Gestalt psychology. Harcourt & Brace, New York. 1935.
[9] N. Krueger. Collinearity and parallelism are statistically significant second order relations of complex cell responses. Neural processing letters 8: 117-129. 1998.
[10] U. Polat, D. Sagi. Spatial interactions in human vision: From near to far via experience dependent cascades of connections. Proc. Natl. Acad. Sci. USA Vol. 91, pp. 1206-1209, 1994
[11] M. Sigman, G.A. Cecchi, C.D. Gilbert & M.O. Magnasco. On a common circle: Natural scenes andGestalt rules. Proc. Natl. Acad. Sci. Vol. 98 (4), pp 1935-1940, 2001.
[12] E.P. Simoncelli, B.A. Olshausen. Natural image statistics and neural representation. Ann. Rev. Neurosci. Vol. 24, pp. 1193-1215, 2001.
[13] R.J. Tusa, A.C. Rosenquist, L.A. Palmer. Retinotopic organization of areas 18 and 19 in the cat. J. Comp. Neur. Vol. 185, pp.657-678, 1979.
[14] M. Wertheimer. Laws of organization in the perceptual form. Harcourt & Brace. 1938

Figure 1: Four sample frames of our database are shown on the left. The amplitudes of the oriented energy detectors for the same frames is shown on the right. The bars indicate the orientation of the respective filter used.

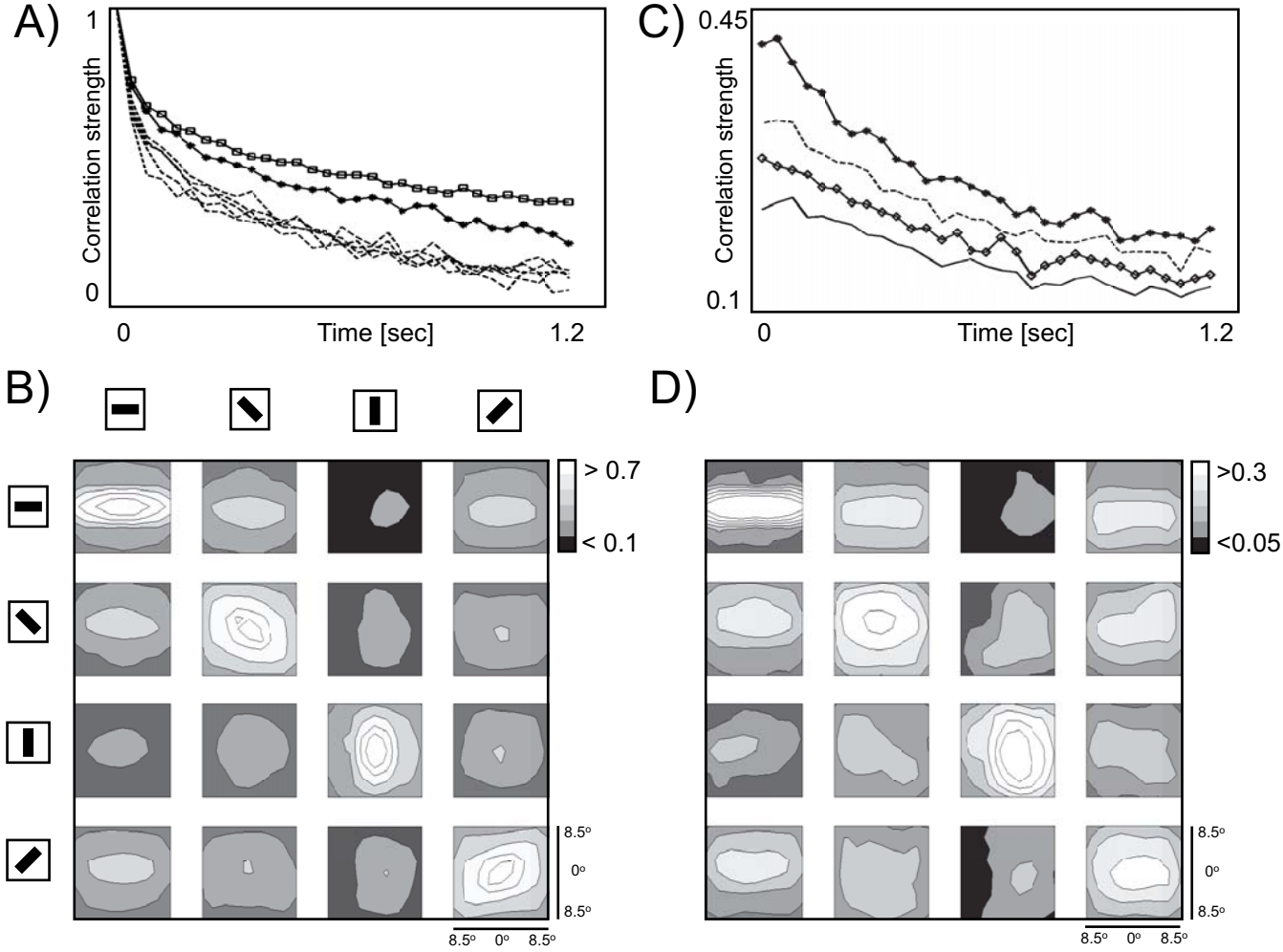

Figure 2: A) The correlation of orientation amplitude over time at the same pixel. Squares: 0^o, stars: 90^o, dashed: all other orientations (spaced 22.5^o). B) Correlations of different combinations of orientations and different spatial arrangements of the two points in the same frame. The orientations are (from top to bottom and left to right): 0^o, 135^o, 90^o, 45^o. C) Correlations over time of points with prominent horizontal orientation but which are spatially separated by different distances independent of relative orientation. Squares: 2.1 deg spatial distance, stars: 4.2 deg, dashed: 6.4 deg, diamonds: 8.4 deg.D) Same as in B but here the two points are also separated by 400 ms in time.

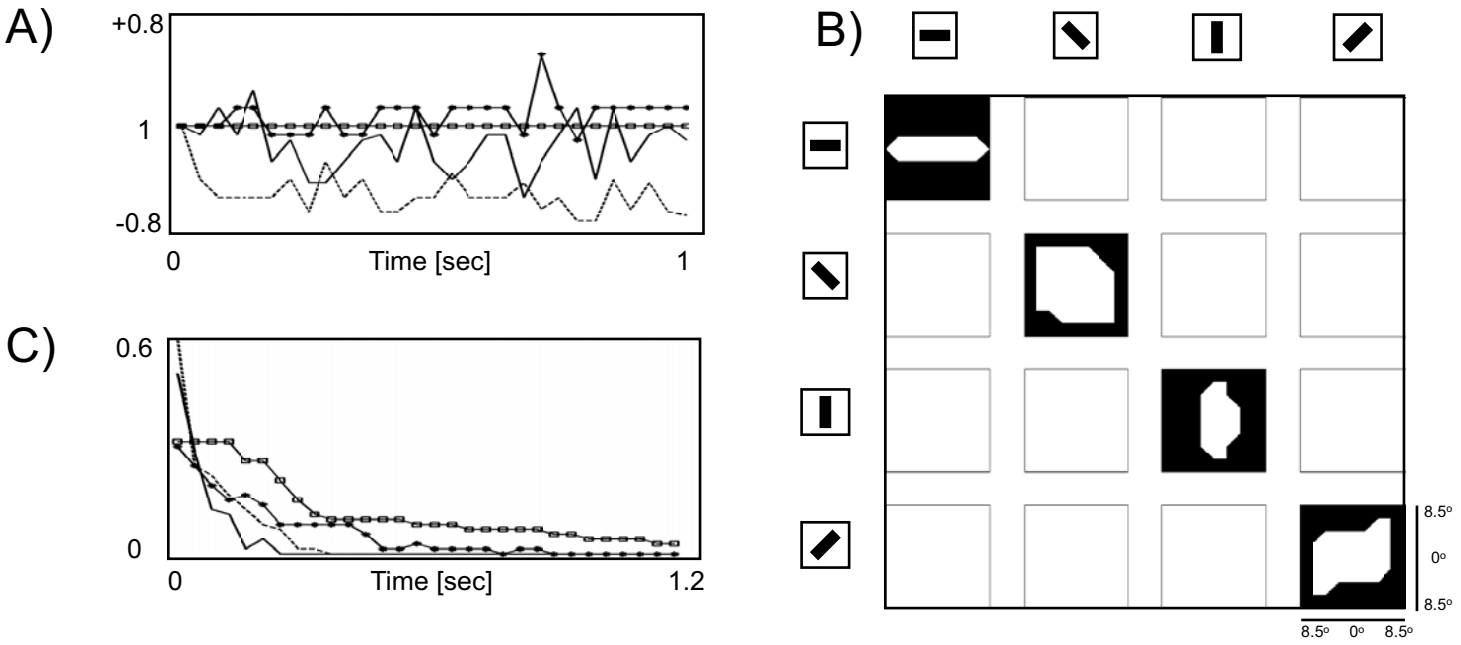

Figure 3: A) Relative change of the aspect ratio of the correlation contours in Fig. 2C as a function of time. Shown is the aspect ratio at each point in time divided by the aspect ratio at t =0. Squares: 0^o, stars: 90^o, solid: 45^o, dashed: 135^o. B) Areas of strong correlations. We defined spatio temporal separations with a correlation over 0.4 as strong. The figure shows these areas for the correlation diagram of Fig 2C. C) Shows the size of these areas relative to the total patch size over time. Lines are labeled as in A.

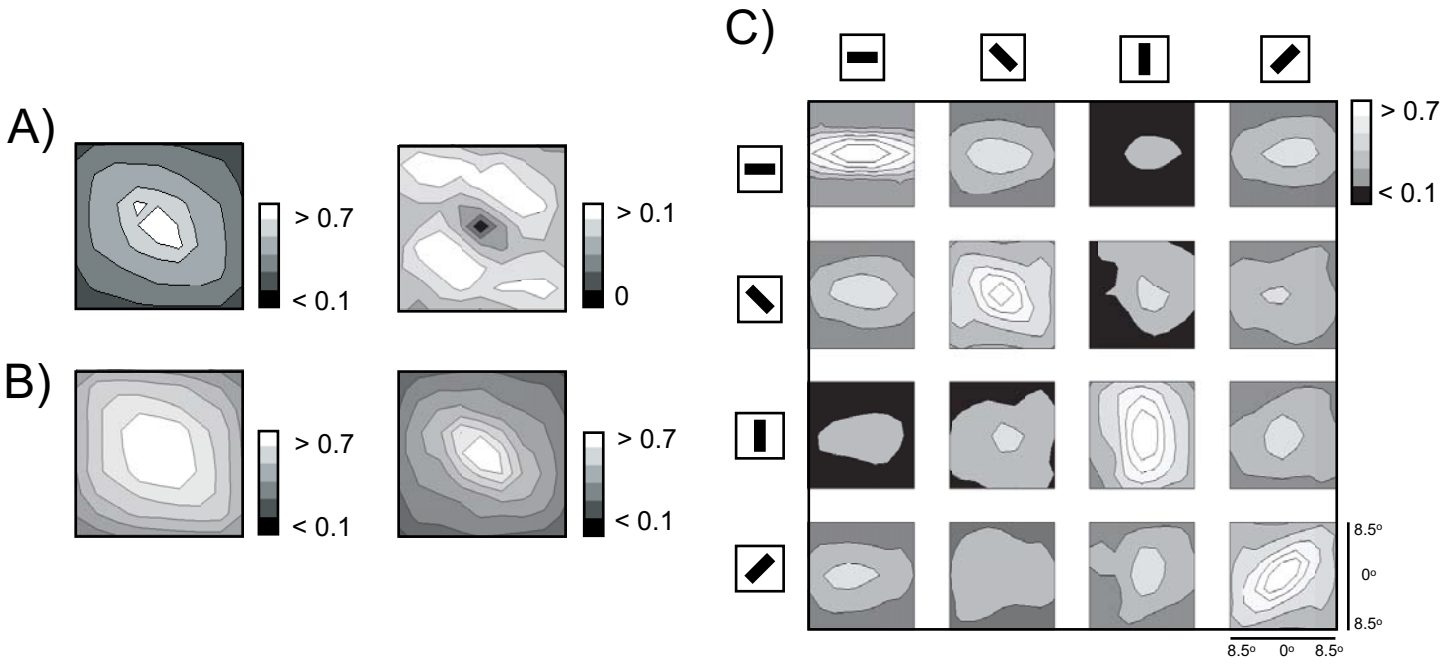

Figure 4: A) For an example orientation (45^o) we show the mean (left) over 12 video sequences together with the standard deviation (right). B) The mean plus / minus the standard deviation. C) Cross orientation correlations over space for filters of a higher spatial frequency.