Binocular rivalry favors naturalistic stimuli in space and time

Daniel Baker & Erich Graf

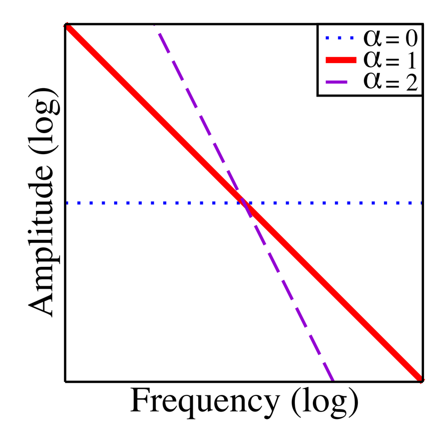
School of Psychology, University of Southampton

contact: d.h.baker@soton.ac.uk

May 2009

1. Background

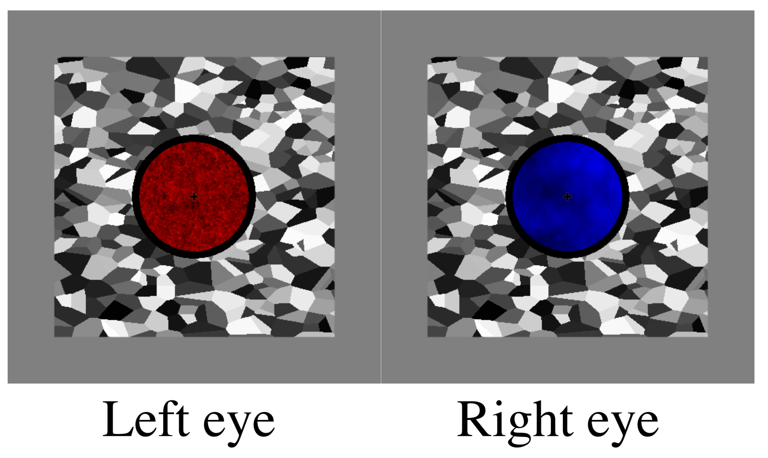
The visual system evolved to process images in the natural environment (see Geisler, 2008). It follows that statistical properties associated with natural stimuli might be processed more effectively than more artificial stimuli. A well known statistical regularity (e.g. Dong & Attick, 1995) is the lawful reduction in amplitude as a function of frequency, which approximates 1/f both spatially and temporally, (red line in right panel). Natural images also contain correlations across scale space in the phase spectrum which correspond to edges.



A central premise of ecological accounts of vision is that stimuli with natural properties should be preferrentially processed. Here, we test this using binocular rivalry to ask if naturalistic stimuli are favoured.

2. Methods

For the first two experiments, stimuli were noise images generated in Matlab by applying filters in the Fourier domain. We varied the amplitude spectral slope (α) in either the spatial (for a static image) or temporal (for a dynamic image sequence) domain over five values (α = 0, 0.5, 1, 1.5, 2). Stimuli were tinted red and blue to aid identification, and had a fixed RMS contrast of 0.15. Counterbalancing was across eye of presentation and image tint, for 15 factorial combinations of α. Four observers each completed 300 one-minute trials per experiment.

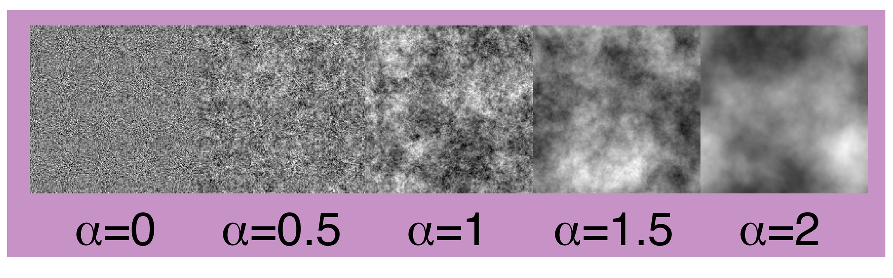


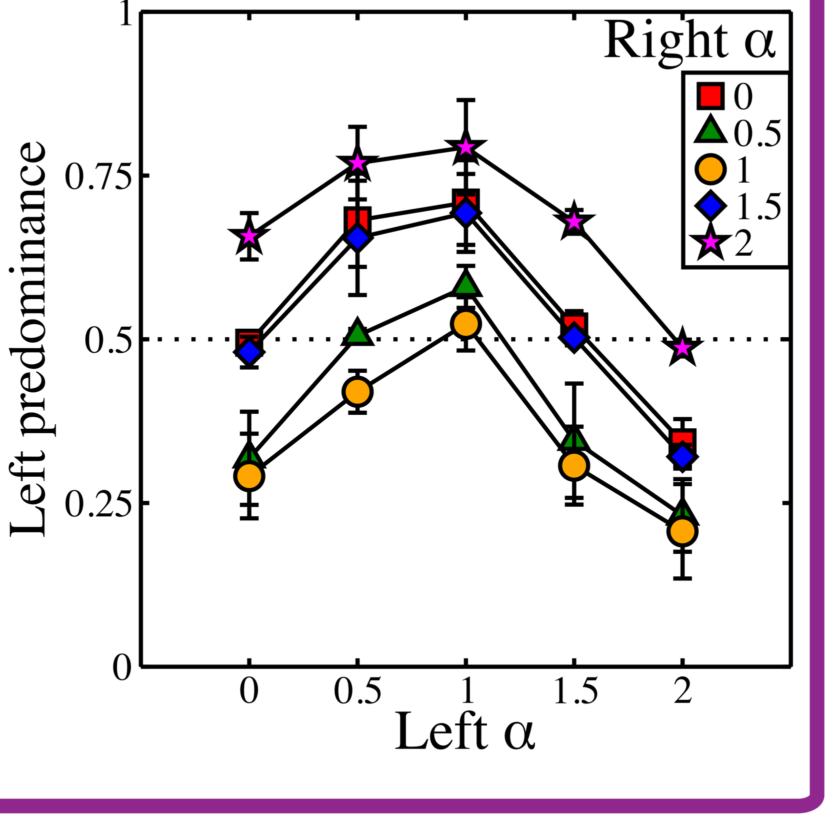
The third experiment used colour photographs from the McGill Calibrated Color Image Database (Olmos & Kingdom, 2004) which rivalled with their phase-scrambled versions. Stimuli had equal RMS contrast, and were counterbalanced across eye. A rivalry simulation condition (see below) was also included to measure observer biases. Six observers each completed 192 one-minute trials, half rivalry and half simulated rivalry.

In all experiments, observers indicated their percept continuously using a keyboard. A mirror stereoscope enabled dichoptic viewing. Predominance scores were calculated as the proportion of all responses allocated to a given stimulus, and averaged across trials and observers. To aid stable fusion, the stimuli were surrounded by a binocular Voronoi texture (see image).

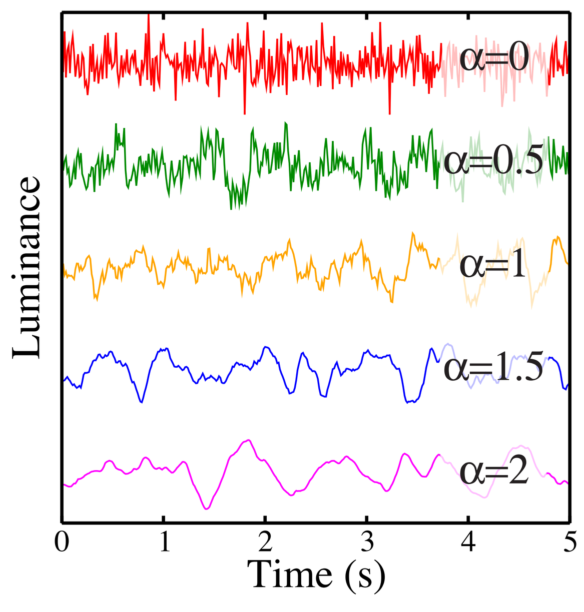
3. Results: spatial amplitude spectrum

The average results across four observers are shown in the panel. Symbols indicate the α value shown to the right eye, and position on the ordinate gives the left eye's α value, though these are interchangeable as the experiments were counterbalanced. For all functions, there is a clear peak at α=1. This indicates that stimuli most similar to those in the natural environment tended to dominate, relative to all other α values. It is consistent with previous results, such as the finding that noise images for which α=1 produce stronger surround suppression than those with other spectral slopes (McDonald & Tadmor, 2006).



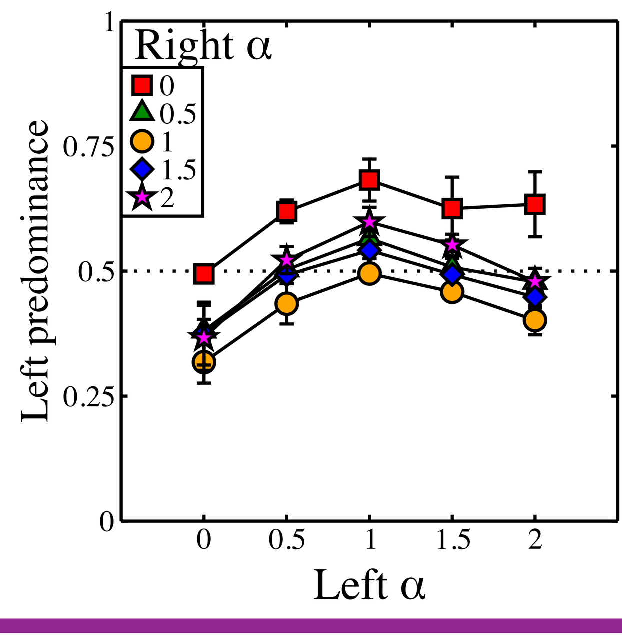


4. Results: temporal amplitude spectrum



The second experiment used temporally filtered dynamic noise sequences, and retained the design and methodology of the spatial version of the experiment. Illustrative luminance values from single pixels of example stimuli are shown in the left panel for five α values. The stimuli always had a spatial α of unity, and resembled moving cloud patterns at different speeds (see Billock et al., 2001, for a principled approach to describing temporally filtered noise stimuli).

Please ask to view a movie of the stimuli when the poster is attended



Results are shown in the right panel, again averaged over four observers. As for the previous experiment, a peak is apparent for all functions at α=1. Again, this was evident in the individual data for each observer. This extends our findings to the temporal domain, and is evidence that the visual system is preferrentially selective for image sequences with natural properties.

5. Ecological validity of contrast metrics

In the previous two experiments, all stimuli had the same RMS contrast (and hence the same total energy). But is the *effective* contrast also the same in terms of:

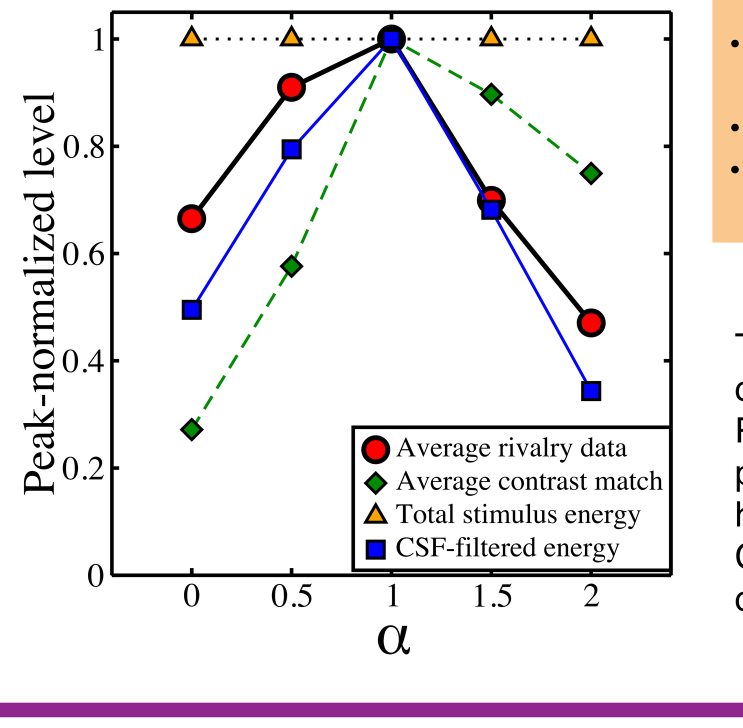
i) perceived contrast?

ii) suprathreshold contrast energy?

We assessed this using (i) a contrast matching task and (ii) by attenuating the stimulus energy using a model CSF.

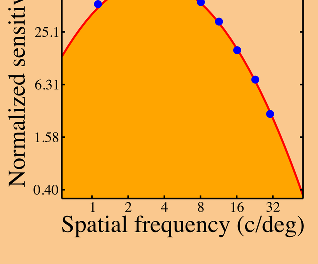
Contrast matching procedure

- Match noise stimulus to 3c/deg grating
- 1-up-1-down staircase controls grating contrast
- Measure PSE for 3 observers and average



Calculating CSF-filtered energy

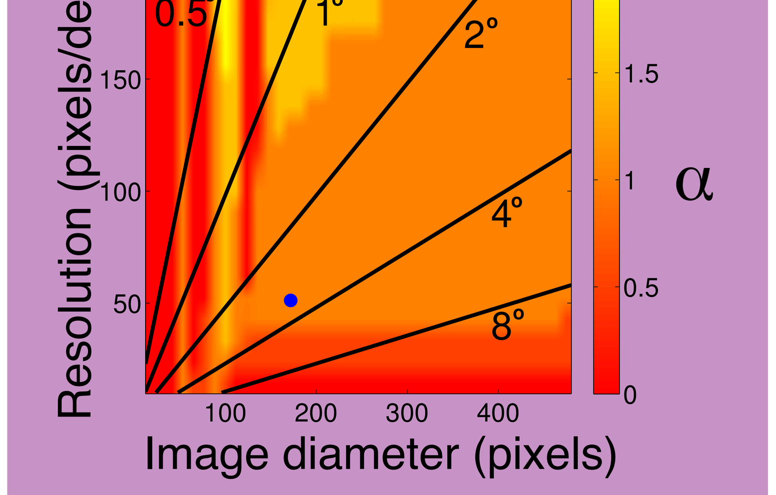
- Fit average Modelfest grating detection data using 2nd order polynomial
- Attenuate amplitude spectra of noise images
- Sum energy across spatial scales



Results

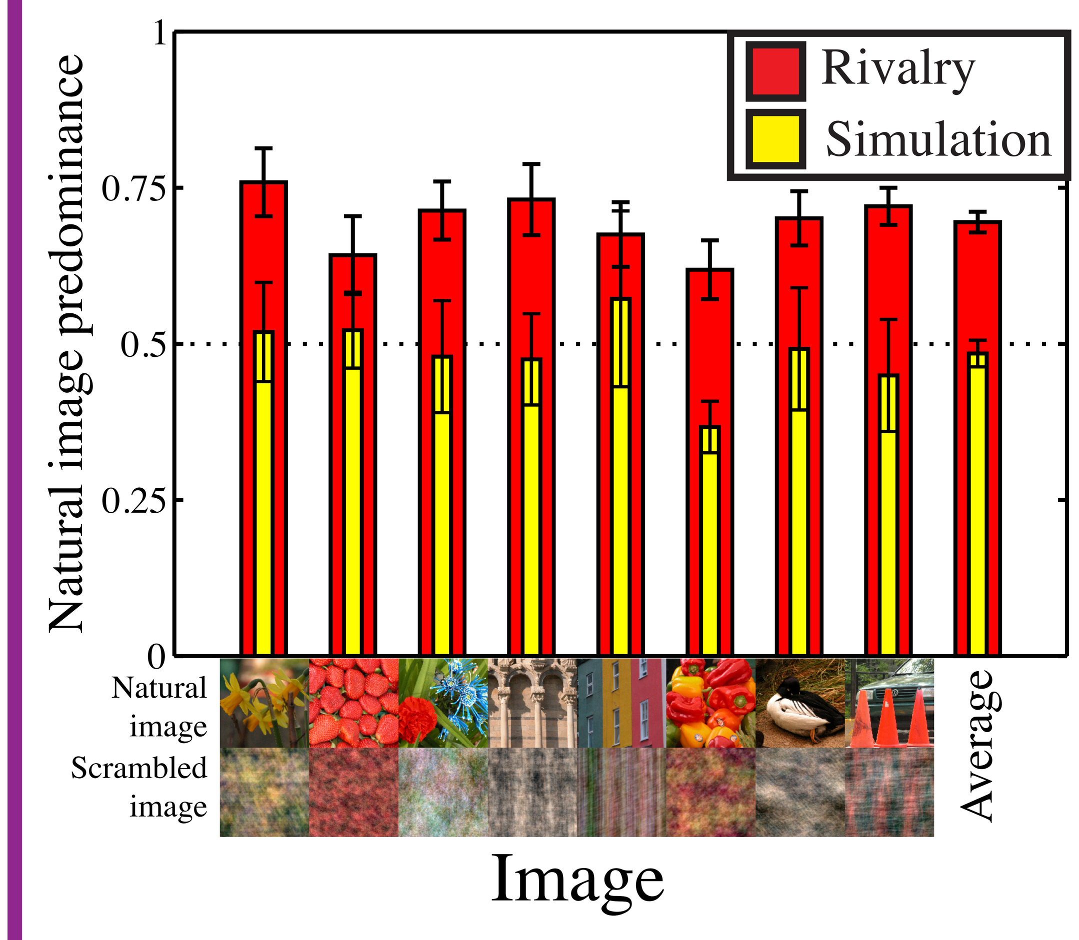
Total energy (orange triangles) is constant, and is a poor predictor of the average rivalry data for static noise images (red circles). Perceived contrast (green diamonds) measured by matching does peak at α=1, but is too steep at low α values, and too shallow at high α values.

CSF-filtered energy (blue squares) is a surprisingly good predictor of dominance.



To check that these findings are not peculiar to our stimulus parameters, we calculated CSF-filtered energy over wide range of stimulus sizes and resolutions. Energy was greatest at α=1 for 61% of values, and only differed when resolution was very poor or stimuli were small.

6. Results: phase randomized images

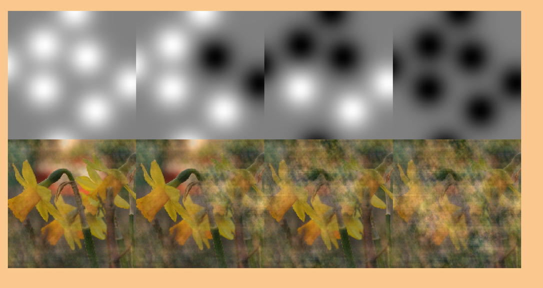


Rivalry between natural images and their phase-scrambled counterparts produced clear and striking results. For all stimuli, images with intact phase spectra were dominant for more than 50% of presentation duration (red bars), with a mean predominance of 69.5%. This was highly statistically significant (F=329.57, p<<0.001), as were all ANOVAs for the six individual observers.

A binocular simulation of rivalry was also run as a control condition (see box below for method). Dominance of natural images in this condition (yellow bars) was not significantly greater than 50% (F=0.72, p>0.41), with a mean duration of 51.4%. This indicates that observers were not biased to report dominance of the phase-intact images.

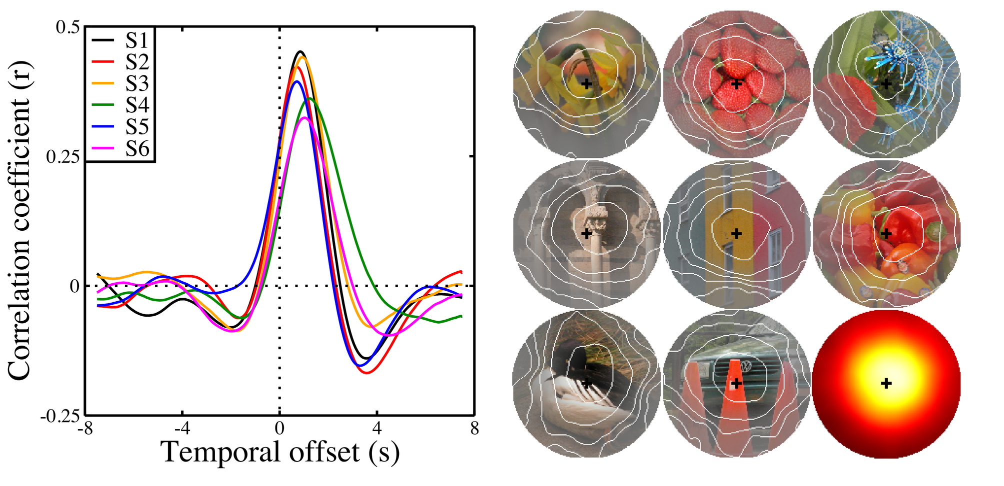
Rivalry simulation - method

Using a similar method to Lee & Blake (2004), we created binocular movie sequences, in which different amounts of phase-intact and phase-scrambled images were visible. This was achieved by inserting Gaussian windows into the alpha (transparency) layer of an RGBA image. The state of each Gaussian varied over time, according to a gamma distribution, producing smooth local transitions between the two images, very similar to piecemeal rivalry. Observers were given the same instructions - to report which image was most dominant.



Rivalry simulation - detailed analysis

The results of the simulation condition were further analysed by cross-correlating observer responses with the average state of the stimulus (natural or phase-scrambled). There was a peak in the cross-correlation function with a mean latency of 895ms (see below). We then generated correlation maps by calculating how well the state of each pixel predicts observer responses at the appropriate latency. Any bias towards focusing on salient image features should show up as a spatial inhomogeneity in the resulting heat maps. These are shown below for each image (averaged over 6 observers). All maps peak close to fixation, with no bias around features.



7. Summary and conclusions

- When given a choice, the visual system preferrentially selects stimuli with naturalistic properties, in both space and time
- An important factor for rivalry dominance is the distribution of energy across spatiotemporal scale available following attenuation by the CSF
- Correlations in the phase spectrum corresponding to contours may also contribute, both to dominance and percept coherence
- Human vision is tuned to the statistics of natural images, and this has measurable consequences for bistable stimuli

8. References

Billock, VA, Cunningham, DW, Havig, PR & Tsou, BH (2001). Perception of spatiotemporal random fractals: an extension of colorimetric methods to the study of dynamic texture. *J Opt Soc Am A*, **18**: 2404-2413.

Dong, DW & Atick, JJ (1995). Statistics of natural time-varying images. *Network*, **6**: 345-358.

Geisler, WS (2008). Visual perception and the statistical properties of natural scenes. *Annu Rev Psychol*, **59**: 167-192.

Lee, S-H & Blake, R (2004). A fresh look at interocular grouping during binocular rivalry. *Vision Res*, **44**: 983-991.

McDonald, JS & Tadmor, Y (2006). The perceived contrast of texture patches embedded in natural images. *Vision Res*, **46**: 3098-3104.

Olmos, A & Kingdom, FAA (2004). McGill Calibrated Colour Image Database. http://tabby.vision.mcgill.ca

Please help yourself to a card containing the citation for this work. We are also happy to provide reprints by email.