Classification of Natural Scenes: Critical Features Revisited

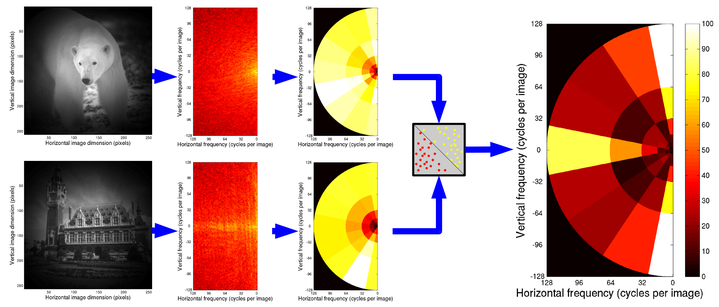
Jan Drewes, University of Giessen, Giessen, Germany (Jan.Drewes@psychol.uni-giessen.de)
Felix Wichmann, Max Planck Institute for Biological Cybernetics, Tübingen, Germany (felix@tuebingen.mpg.de)
Karl R. Gegenfurtner, University of Giessen, Giessen, Germany (Karl.R.Gegenfurtner@psychol.uni-giessen.de)

Motivation

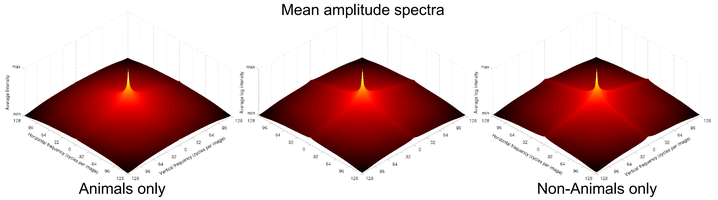
The algorithmic classification of complex, natural scenes is generally considered a difficult task due to the large amount of information conveyed by natural images. Work by Simon Thorpe and colleagues showed that human observers are capable of detecting animals within novel natural scenes with remarkable speed and accuracy. Despite the seeming complexity of such decisions it has been hypothesized that a simple global image feature, the relative abundance of high spatial frequencies at certain orientations, could underly such fast image classification [1][2].

Methods

We classified a set of ~11.000 handpicked grayscale images into animal and non-animal images, using a Fourier fingerprint with 8 orientations and 6 frequency bands, totaling 48 dimensions [3]. With classifier based on linear discriminant analysis, we were able to successfully classify up to 74% of the images correctly. We inferred a rating of bin importance based on multiple classifications, showing that the bins most important for classification are the highest frequencies at the vertical and horizontal orientations [2].



Mean amplitude spectra



Animals only

Non-Animals only

Conclusion

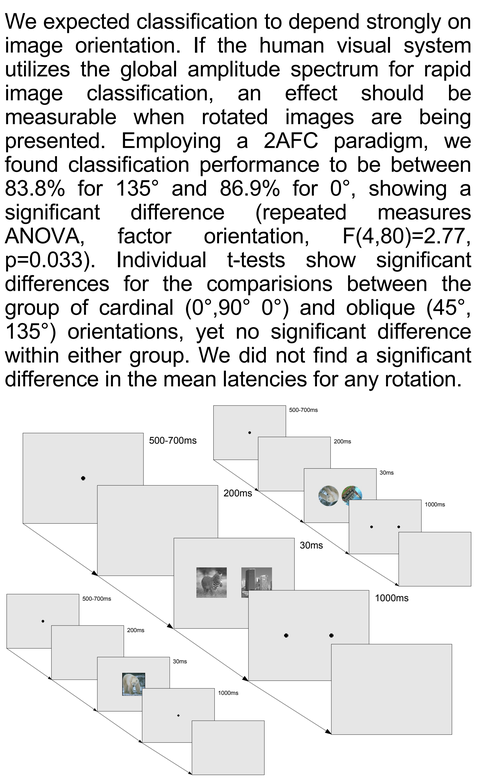
Our results show that a classifier based solely on the global amplitude spectrum indeed behaves remarkable similar to our human subjects relative classification performance. However, we have shown human classification performance to be largely independent of the global amplitude spectrum; instead, after equation of the amplitude spectrum, human performance is reduced by only a few percent, while the effect of classification difficulty persists. The relative differences between individual images found in the global amplitude spectrum must therefore result from other image properties that are not eliminated by spectral equation.

[1] Johnson & Olshausen, Journal of Vision, 2003
[2] Torralba & Oliva, Network: Comput. Neural Syst., 2003
[3] Drewes, Wichmann, Gegenfurtner, VSS 2005

Supported by DFG Ge 879/6 Wahrnehmung natürlicher Szenen

Results

We expected classification to depend strongly on image orientation. If the human visual system utilizes the global amplitude spectrum for rapid image classification, an effect should be measurable when rotated images are being presented. Employing a 2AFC paradigm, we found classification performance to be between 83.8% for 135° and 86.9% for 0°, showing a significant difference (repeated measures ANOVA, factor orientation, $F(4,80)=2.77$, $p=0.033$). Individual t-tests show significant differences for the comparisions between the group of cardinal (0°,90° 0°) and oblique (45°, 135°) orientations, yet no significant difference within either group. We did not find a significant difference in the mean latencies for any rotation.



We then used a Go/No-go paradigm to evaluate human performance on a set of images that our Fourier fingerprint classifier found to be “easy” or “difficult”. Those images that our computer algorithm found to be "easily" classified were apparently also found to be easier by our human subjects; classification performance showed a significant difference on the mean hit ratios ($F(1,10)=30.621$, $p<0.001$) as well as the mean latencies of the valid go-responses ($F(1,10)=8.569$, $p=0.014$).
Next, we equated the amplitude spectra of the 400 images previously tested, eliminating any clue that might have been contained in the global amplitude spectrum. Retesting human performance, we found performance was only slightly reduced, and the relative difference between “easy” and “difficult” images remained ($F(1,10)=119.301$, $p<0.0001$) for the hit ratios, ($F(1,10)=27.584$, $p<0.001$ for the latencies).