43.1: *Invited Address:* Nonlinear Aspects of Primary Vision: Entropy Reduction Beyond Decorrelation

C. Zetzsche, B. Wegmann
Technische Universiät München, München, Germany
E. Barth
Ludwig-Maximilians-Universität, München, München, Germany

ABSTRACT

Recent developments like sub-band and wavelet coding show the relevance of biological vision for image data compression. These decomposition schemes rest on linear filter operations and remove second-order statistical dependencies by decorrelation. We discuss how nonlinear detectors in biological vision (complex and hypercomplex cells) may contribute to the reduction of higher-order dependencies.

INTRODUCTION

Modern techniques for image data compression and recent concepts in biological vision research seem to converge to a common framework. This becomes apparent in the trend from block transforms towards biology-like filter decomposition schemes (subband, wavelet, etc.), and in recent approaches to an analysis of the primary vision stages in terms of signal statistics and entropy reduction [6][18][22][26].

Up to now, this convergence of concepts seems to be confined to the level of linear filter theory and, consequently, to the exploitation of second-order properties. In terms of neural machinery this corresponds to an emphasis on retinal lateral inhibition and on cortical simple cells. However, the biological system contains various other cell types with substantial nonlinear behavior. Our goal, therefore, is to investigate how these nonlinear mechanisms may be engaged in the exploitation of higher-order properties of the signal statistics.

The first relevant nonlinearity, besides retinal adaptivity, occurs on the level of cortical complex cells. In this context we have investigated the properties of the joint two-dimensional probability density function (pdf) resulting from the output of even- and odd-symmetric simple cell pairs. We could show that the outputs of these cells are, though decorrelated, not statistically independent [19]. Interestingly, complex cells can be seen to make use of these dependencies by a nonlinear transform of the even/odd signal space into a polar amplitude/phase representation. Furthermore, the complex cells outputs are suited to make statistical dependencies between mechanisms with different tuning properties accessible in a more efficient way.

An even more essential nonlinearity arises on the level of hypercomplex, or end-stopped cells. Their operation is clearly beyond the scope of filter theory and requires the incorporation of novel signal processing concepts. We show how the necessary "and"-like operations can be derived from differential geometry. A key feature of our approach is the notion of "intrinsic dimensionality". This results in a hierarchy of 0D-, 1D-, and 2D-signals which has a close relation to the degree of redundancy of the signals. We demonstrate, that the least redundant and sparsely distributed 2D-information extracted by nonlinear operators with end-stopped behavior is sufficient for a reconstruction of the original image signal.

STATIONARITY, SECOND-ORDER STATISTICS, AND LINEAR SYSTEMS

The classical methods for the design of data compression schemes follow one common strategy: the optimum parameters of a linear device (coefficients of a linear predictor, basis (eigen) images of a linear transform) are determined from the second-order statistics (covariance) of the input signal (e.g. [10]). For signals from a stationary source powerful optimization theorems can be derived. The basic technique for the exploitation of redundancies in such signals is the Karhunen-Loeve transform (KLT). In particular, it can be shown mathematically that (i) it can transform a set of correlated input data into a set of uncorrelated output coefficients, and that (ii) it provides among the unitary transforms the optimal "compaction of energy" if only a subset of the coefficients is retained for the reconstruction of the signal (e.g. [9]). Similar arguments can be provided for the optimization of predictive and subband coding schemes.

Due to its dominating role in the history of the subject the linear/second-order approach is often seen as the ideal theoretical framework for data compression. It is known, however, that its practical suitability depends on several assumptions which may not be valid in the situations of interest. The main problem is that natural image signals often can not be adequately described by a statistical model. This failure of statistical models has given rise to the following, widely accepted paradigm of statistical approaches to image processing :

Natural image signals are highly nonstationary, hence they have to be processed by adaptive procedures.

One goal of this paper is to remind the reader that this nonstationarity hypothesis is, in spite of its wide acceptance, disputable and potentially misleading for the development of alternative image processing concepts. We will demonstrate this by means of a counter example.

The basic reasoning behind the statistical approach to image compression is that the system should be designed for an entire class and not just an individual image (e.g. [9]). In our example this class may be thought of as satellite images of Jupiter cyclones (see Fig. 1)..

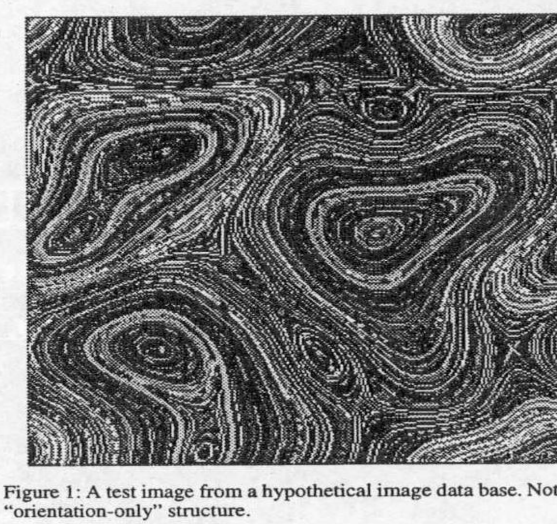

Figure 1: A test image from a hypothetical image data base. Note the "orientation-only" structure.

Clearly, the image contains considerable structural redundancies, i.e., there are strong statistical dependencies in the associated random

ISSN0097-0966X/93/0000-0933-$1.00 + .00 © 1993 SID

field. Surprisingly, however, the autocorrelation function tells us that the image is nearly perfectly decorrelated (Figure 2).

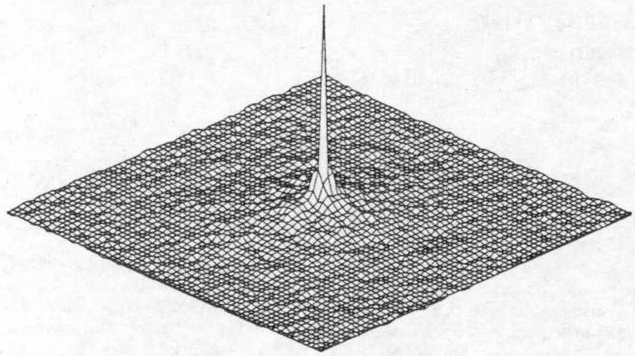

Figure 2: Autocorrelation function of the test image shown in Figure 1. A 64 x 64 section is shown. The sharp decrease in the immediate vicinity of the origin implies that second-order statistics "see" the image as almost completely decorrelated.

Unlike as in realistic situations, however, here we have full access to the process generating the images. The source of our example is a computer simulation - and can be shown to be perfectly stationary and ergodic. Therefore, the classical nonstationary interpretation is definitely wrong in this case

The reason for this failure lies in the presence of strong higher-order dependencies. Second-order estimates depend on fourth-order properties (e.g. [15]). Since natural images possess far reaching, higher-order dependencies second-order estimates based on local windows must be expected to show considerable variance across spatial position. Observation of such a variation can thus not be used as a reliable indicator for nonstationarity.

If higher-order dependencies, rather than nonstationarity, are seen as the key feature of image statistics, the current approaches to image coding should be further extended towards the exploitation of such dependencies. In an implicit fashion, adaptive, frequency specific decomposition schemes do already make partial use of higher-order dependencies. Adaptive schemes are designed to track the presumed nonstationarities of natural images but they can equally well be regarded as static but nonlinear devices which exploit stationary higher-order dependencies [26]. Seen in this way, they are a restricted case of the most general nonlinear approach, namely vector quantization (VQ). Since straight-forward VQ of the image itself (e.g. by means of the LBG algorithm) has not lead to substantial advantages over conventional coding schemes, additional structural knowledge has to be incorporated into the general VQ framework, and various, mostly heuristically motivated, modifications of the basic VQ approach are under investigation (e.g. [7]).

It has to be expected that the dominance of linear/second-order concepts will further decline in favour of nonlinear/higher-order approaches. VQ is the framework with the highest potential but, due to its generality, needs to be structured with respect to the nature of the signals being processed. Therefore, more specific nonlinear concepts for the exploitation of higher-order dependencies are of interest. We think that biological vision systems may provide us with suitable candidates in the search for such principles.

NONLINEAR OPERATIONS IN BIOLOGICAL VISION

So this might not be evident from the vast amount of psychophysical and neurophysiological literature on the spatial filter properties of visual mechanisms, only a small part of the visual machinery can be adequately described in terms of linear system theory. Rather, biological image processing is notoriously nonlinear right from the first stage.

Retinal adaptivity

This is evident from the common usage of *contrast* as the determining variable in most models. It is only in terms of contrast and for threshold stimuli that retinal processing can be characterized as a linear filtering stage. However, the two-dimensional luminance distribution of natural images cannot be described in terms of contrast. The retinal processing of a natural image has to be modeled as adaptive, i.e., as dependent on the local context, and must, therefore, be described within a nonlinear framework. The deeper reason for this nonlinear processing is the multiplicative coupling of reflectance and illumination component, and the full understanding in terms of statistical signal processing would require elaborated two-source concepts.

Recent models of retinal processing use multi-resolution schemes with logarithmic and divise nonlinearities [16][21][23]. A substantial coding gain, however, is already obtained with the common nonlinear transducer properties of cameras and monitors.

Complex cells

On the cortical level, the incoming signals are split up by different types of cells with a variety of tuning preferences (e.g. orientation, velocity, color). However, only one cell type, namely that of cortical simple cells, dominates the theories on cortical visual processing. Maybe this is because simple cells are the only cortical cells for which a modeling in terms of linear spatial filters is easily feasible. Currently, the discussion is focussed on the appropriate mathematical functions (Gabor functions, derivatives of Gaussian, wavelets, etc.; see e.g. [11]).

The linear simple cells represent only a subpopulation in the primary cortical areas. Depending on the area, estimates for the percentage of end-free simple cells range between 30% and 10% [14]. A considerable population of similar size is established by complex cells. These cells seem to play an important role in feature detection and in texture and motion processing (e.g. [1][13][22]). Since they show a considerable nonlinear behavior they may contribute to the exploitation of higher-order dependencies.

It is important to point out, that second-order theory predicts no substantial advantage of an orientation specific signal decomposition by linear filters (QMF, wavelet, etc.). Since the second-order properties of natural image signals show no general anisotropies (besides the minor "oblique effect" [2]) the classical exploitation of deviations from spectral flatness by appropriate subband filters [10] should yield no advantage for an angular decomposition but only for a radial decomposition of the spectral domain (cf. [26]). It should also be clear from our "Jupiter cyclone" example that the presence of locally oriented structures cannot be adequately described as a second-order property.

Filter outputs are decorrelated by definition if the filters are orthogonal. Since this is, at least approximately, the case for the spatial filters in biological vision, the outputs of visualsimple cells are also more or less decorrelated. However, this does not imply that there are no higher-order statistical dependencies left. Due to the structure of natural images substantial inter-position, inter-orientation, and inter-resolution dependencies have still to be expected. The exploitation of such dependencies between decorrelated signals requires nonlinear operations.

A statistical dependency between the (decorrelated) outputs of jointly located even- and odd-symmetric simple cells can be easily demonstrated by the difference between the shape of the joint two-dimensional pdf and the shape of the pdf which results from multiplication of the marginal pdf's [19]. According to the concept of the analytic signal the operation of complex cells can be seen as a nonlinear coordinate transformation from a cartesian representation (even/odd) into a polar amplitude/phase representation which is better matched to the circular symmetric shape of this joint pdf [19].

Even more interesting than the dependencies between the outputs of detectors with differing symmetry are the dependencies between detectors tuned to different orientations or different sizes (resolution). We have recently demonstrated that there are substantial inter-orientation effects [20]. Here we show that there are also substantial inter-resolution dependencies and that the nonlinear operation of complex cells can partially transform this higher-order inter-resolution dependencies into second-order dependencies.

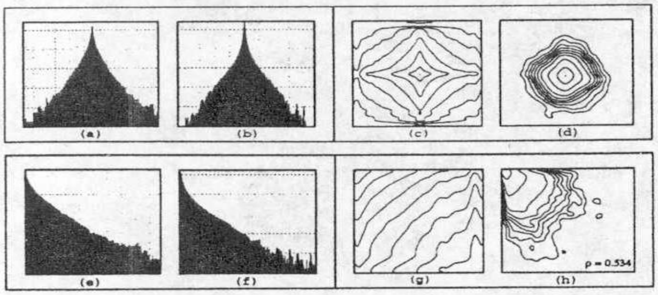

Figure 3: Statistical analysis of inter-resolution dependencies. The output statistics of even-symmetric orientation specific filters with a center frequency of 0.0625 cycle/pel and 0.25 cycle/pel are shown in (a) and (b). The joint statistics under the assumption of independency are shown in (c) and the actual joint statistics are shown in (d). Note that (d) is different from (c), i.e. there are higher-order statistical dependencies, but (d) shows no correlation. The lower row shows the same statistics but for the amplitudes of the filter outputs (complex cell outputs). Now some statistical dependencies do show up as correlations in (h).

The upper row of Figure 3 demonstrates that the joint output statistics (3d) of two linear orientation filters with sufficiently different center frequencies show no correlation (as has to be expected) but reveal a statistical dependency. If the same analysis is performed for the (nonlinear) amplitudes (lower row of Figure 3), it is evident that part of this higher-order dependency is transformed into second-order (i.e., correlation). Furthermore, significant information is now more easily accessible by coincidence detectors.

2D-cells (hypercomplex, end-stopped)

A natural hierarchy for local image signals is established by their *intrinsic dimensionality* [24]. 0D-signals (flat areas) are most frequent in natural images, 1D-signals (such as locally straight edges) are the next frequent group, while 2D-signals (such as corners) are the most rare events.

0D- and 1D-signals being redundant in a certain sense, the 2D-signals convey the essential information in images. This importance of 2D-signals is also taken into account by biological vision systems. The "Attneave cat" [3] and the insensitivity of hypercomplex [8] or dot-responsive cells [17] to any kind of 0D- or 1D-signal make this point evident.

However, though 2D-specific cells constitute a substantial part of the cells in the primary visual areas [14] they have received far less attention than simple and complex cells. This is probably due to the impossibility of a description of their function in terms of linear spatial filter operations. We have recently developed a theoretical framework [24][25] for the modelling of 2D-cells which combines the nonlinear classification properties of differential geometry [12] with the flexible tuning properties of linear filters.

The essential point of this framework is that all 2D-operators have necessarily to be based on an "and"-like (e.g., multiplicative) combination of the outputs of elementary units which differ in their orientation specificity. Furthermore, this is a necessary but in general not sufficient requirement. In many cases an additional "inhibitory" compensation is necessary to compensate for erroneous responses to certain 1D-signals.

RECONSTRUCTION FROM 2D-SCALE-SPACE

2D-detectors are of relevance for image coding since their operation can be seen as a nonlinear transformation of the image signal into a representation where the nonredundant information is represented by a few active coefficients. This operation can hence be regarded as the nonlinear analog to the decorrelation of signals by linear coding schemes. However, this nonlinear transformation exploits the higher than second-order dependencies induced by the frequent occurrence of locally oriented signals in natural images.

Reconstruction of the image from such a nonlinear representation poses a severe problem, since their exists no adequate theory for the derivation of the necessarily nonlinear inverse transformation (if there is any). For these reasons, we have studied how relaxation schemes may be used for the reconstruction [5]. Figure 4 shows, how an image can be reconstructed from its nonlinear 2D-scale-space representation. Although the relaxation algorithm may not yet be fully optimized for the solution of the reconstruction problem the results are a clear demonstration of our assumption that the essential information of natural images is captured in their local 2D-features.

CONCLUSION

Biological vision can serve as a model for the efficient exploitation of statistical dependencies in natural images. It is well known that retinal lateral inhibition and cortical orientation specificity can be seen as decorrelating operations performed by linear spatial filters. We have shown, however, that decorrelation is insufficient to fully exploit the statistical dependencies induced by the frequent presence of oriented structures in natural images, since these dependencies are of higher than second order.

The exploitation of higher-order dependencies requires nonlinear operators, and we have considered two types of cortical cells as potential candidates for such an operation. Complex cells have been shown to transform higher-order inter-scale dependencies into second-order correlations. Furthermore, we have established a signal hierarchy based on the notion of intrinsic dimensionality. We have shown that the nonredundant information in natural images can be extracted by nonlinear, 2D-specific cells (hypercomplex/end-stopped and dot-responsive cells). By means of a relaxation based reconstruction of an image from its 2D-scale-space representation we could finally demonstrate that the 2D-information is in fact the important information in natural images. The investigation of nonlinear mechanisms in biological vision may hence provide us with valuable methods for the efficient exploitation of higher-order dependencies in image data compression.

ACKNOWLEDGEMENTS

This work has been supported by DFG grant Re337/7-1 to I. Rentschler and C.Z. We thank M. Ferraro for comments on the manuscript.

REFERENCES

1. Adelson,. E.H. & J.R. Bergen (1985) Spatiotemporal energy models for the perception of motion. Journal of the Optical Society of America A2, 284-299
2. Appelle, S. (1972) Perception and discrimination as a function of stimulus orientation: the oblique effect in man and animals. Psychological Bulletin 78, 266-278
3. Attneave, F. (1954) Some informational aspects of visual perception. Psychological Review 61, 183-193
4. Barlow, H. B. (1960) The coding of sensory messages. In W.H Thorpe and O.L Zangwill (Eds.), Current problems in animal behavior Cambridge: Cambridge University Press, 331-360
5. Barth, E., T.M. Caelli, & C. Zetzsche (1993) Image encoding, labelling and reconstruction from Differential Geometry. Computer Vision, Graphics, and Image Processing, submitted
6. Field, D. J. (1987) Relations between the statistics of natural images and the response properties of cortical cells. Journal of the Optical Society of America A4, 2379-2394
7. Gersho, A. & R.M. Gray (1992) Vector quantization and signal compression. Boston: Kluwer Academic Publishers
8. Hubel, D.H. & T.N. Wiesel (1965) Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. Journal of Neurophysiology 28, 229-289
9. Jain, A.K. (1989) Fundamentals of digital image processing. Englewood Cliffs, NJ: Prentice Hall
10. Jayant, N. S. & P. Noll (1984) Digital coding of waveforms. Englewood

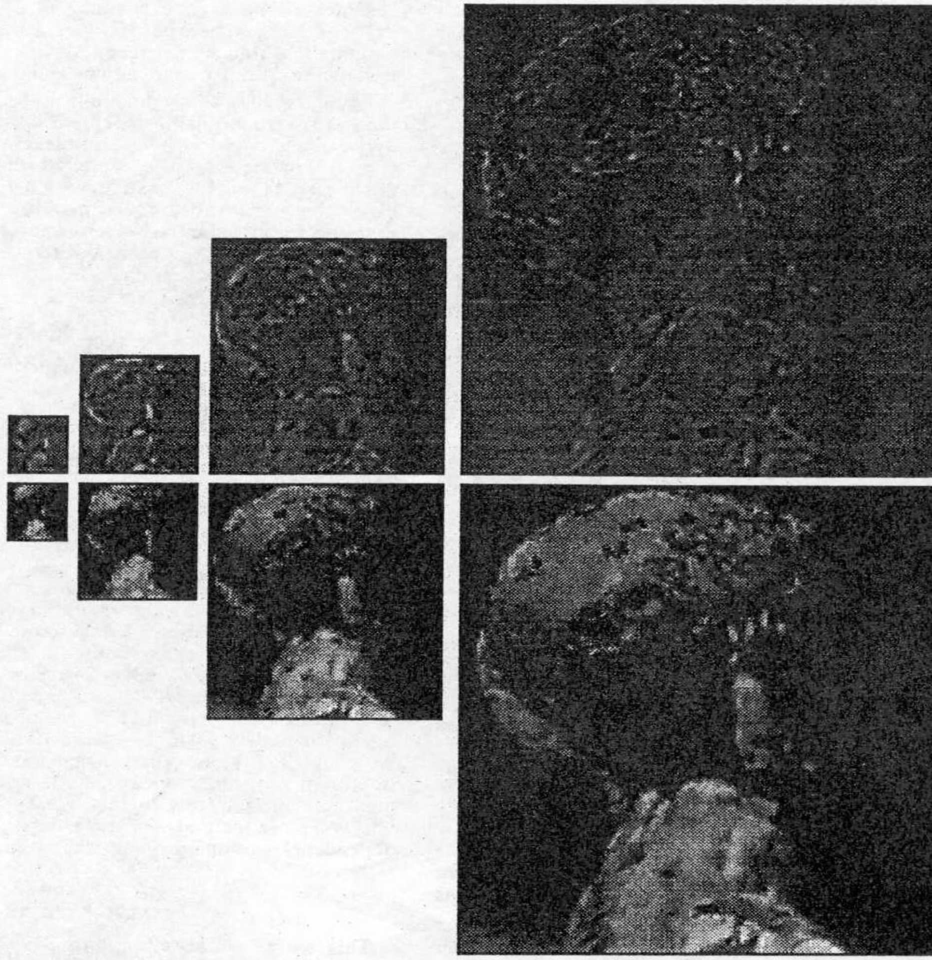

Figure 4. Reconstruction of an image from its nonlinear 2D-scale-space representation. The upper pyramid is a result of the multi-scale application of a bio-inspired 2D-operator to the original input image. The reconstruction is performed by a corresponding multi-grid relaxation algorithm which proceeds from the lower left picture to the final result shown on the lower right.

Cliffs, NJ: Prentice Hall.

11. Klein, S.A. & B. Beutter (1992) Minimizing and maximizing the joint space-spatial frequency uncertainty of Gabor-like functions: comment. Journal of the Optical Society of America A9, 337-340
12. Koenderink, J.J. & W. Richards (1988) Two-dimensional curvature operators. Journal of the Optical Society of America A 5, 1136-1141
13. Morrone, M.C. & D.C. Burr (1988) Feature detection in human vision: a phase dependent energy model. Proceedings of the Royal Society London B235, 221-245
14. Orban, G.A. (1984) Neuronal operations in the visual cortex. New York, Heidelberg: Springer.
15. Papoulis, A. (1984) Probability, Random Variables, and Stochastic Processes. New York: McGraw-Hill
16. Peli, E. (1990) Contrast in complex images. Journal of the Optical Society of America A7, 2032-2040
17. Saito, H., K. Tanaka, Y. Fukada & H. Oyamada (1988) Analysis of discontinuity in visual contours in area 19 of the cat. Journal of Neuroscience 8, 1131-1143
18. Watson,A.B. (1987) Efficiency of a model human image code. Journal of the Optical Society of America A4, 2401-2417
19. Wegmann, B. & C. Zetzsche (1990a) Visual system based polar quantization of local amplitude and local phase of orientation filter outputs, In B. E. Rogowitz and J. P. Allebach (Eds.) Human Vision and Electronic Imaging: Models, Methods, and Applications: Proc. SPIE 1249. Bellingham, WA: SPIE, 306 - 317.
20. Wegmann, B. & C. Zetzsche (1990b) Statistical dependence between orientation filter outputs used in an human vision based image code. In M. Kunt (Ed.) Visual Communications and Image Processing, Proc. SPIE 1360. Bellingham, WA: SPIE, 909 - 923
21. Xie, Z. & T.G. Stockham, Jr. (1989) Toward the unification of three visual laws and two visual models in brightness perception. IEEE Transactions on Systems, Man, and Cybernetics 19(2), 379-387
22. Zetzsche, C. & W. Schönecker (1987) Orientation selective filters lead to entropy reduction in the processing of natural images. Perception 16, 229
23. Zetzsche, C. & G. Hauske (1989) Multiple channel model for the prediction of subjective image quality. In B.E. Rogowitz (Ed.), Human Vision, Visual Processing, and Digital Display, Proc. SPIE 1077, 209-216
24. Zetzsche, C. & E. Barth (1990a) Fundamental limits of linear filters in the visual processing of two-dimensional signals. Vision Research 30, 1111-1117
25. Zetzsche, C. & E. Barth (1990b) Image surface predicates and the neural encoding of two-dimensional signal variations, In B. E. Rogowitz and J. P. Allebach (Eds.) Human Vision and Electronic Imaging: Models, Methods, and Applications, Proc. SPIE 1249, Bellingham, WA: SPIE, 209-216.
26. Zetzsche, C, E. Barth & B. Wegmann (1993) The importance of intrinsically two-dimensional image features in biological vision and picture coding, In A.B. Watson (Ed.) Visual Factors in Electronic Image Communication, Cambridge, MA: MIT Press, in print