Higher-Order Statistics of Natural Images and their Exploitation by Operators Selective to Intrinsic Dimensionality

Gerhard Krieger, Christoph Zetzsche, Erhardt Barth
Ludwig-Maximilians-Universität München
Institut für Medizinische Psychologie
Goethestr. 31, D-80336 München, Germany
krieger@imp.med.uni-muenchen.de

Abstract

Natural images contain considerable statistical redundancies beyond the level of second-order correlations. To identify the nature of these higher-order dependencies, we analyze the bispectra and trispectra of natural images. Our investigations reveal substantial statistical dependencies between those frequency components which are aligned to each other with respect to orientation. We argue that operators which are selective to local intrinsic dimensionality can optimally exploit such redundancies. We also show that the polyspectral structure we find for natural images helps to understand the hitherto unexplained superiority of orientation-selective filter decompositions over isotropic schemes like the Laplacian pyramid. However, any essentially linear scheme can only partially exploit this higher-order redundancy. We therefore propose nonlinear i2D-selective operators which exhibit close resemblance to hypercomplex and end-stopped cells in the visual cortex. The function of these operators can be interpreted as a higher-order whitening of the input signal.

1. Introduction

The extraordinarily large amount of information contained in images of natural scenes calls for efficient data compression techniques. In order to exploit redundancies, most of today's image encoding schemes rely primarily on second-order statistics in conjunction with linear systems theory. Examples are predictive, sub-band (wavelet), and transform coding (e.g. KLT). However, even such common image structures like locally oriented lines and edges cannot be represented by second-order statistics (for a detailed argumentation see [10]). Vector quantization could theoretically overcome these limitations but suffers from combinatorial explosion.

An alternative approach can be derived from the concept of *intrinsic dimensionality* which relates the degrees of freedom provided by a signal domain to the degress of freedom actually used by a given signal [8]. This concept provides a hierarchy of local image signals in terms of different degrees of redundancy:

i0D-signals are constant, i. e. $u(x, y) = const.$ within a local window.

i1D-signals can locally be approximated by a function of only one variable, i. e. $u(x, y) = u(ax + by)$. Examples are straight lines and edges. Sinusoidal gratings, the eigenfunctions of linear systems, are also a member of this class.

i2D-signals are neither *i0D* nor *i1D*. Examples are corners, junctions, curved lines and edges, etc.

The hierarchy of intrinsic dimensionality seems to be reflected in biological vision systems, which are well adapted to the statistical properties of the environment. Evolution caused not only the development of linear isotropic and orientation-selective neurons, but also of nonlinear *i2D*-selective cells known as 'hypercomplex' or 'end-stopped'. For example, in primates more than half of the neurons in area V2 of the visual cortex are selective to *i2D*-stimuli [5]. It has also been shown that the least redundant *i2D*-information is already sufficient for a reconstruction of the original image signal [1].

2. Limitations of the second-order approach

Our statistical investigations of natural images revealed that the hierarchy in terms of intrinsic dimensionality is reflected in different class probabilties [6] [7]. In this hierarchy *i0D*-signals are the most common, and hence most redundant ones. The corresponding type of redundancy can

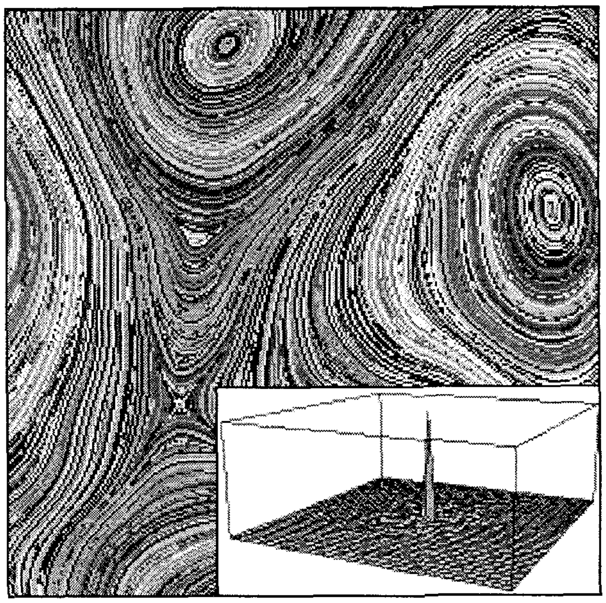

Figure 1. Sample image of an ergodic orientation-only process. This image completely lacks second-order correlations as can be seen from the autocorrelation function shown in the inlet (cf. [9]).

be related to second-order statistics, in form of the characteristic strong emphasis of low frequency components.

The next common class of signals are the *i1D*-signals, and any complete statistical description has to take them into account, too. Since the investigations of Oppenheim and Lim it is known, that the phase relations between the frequency components play an important role for the representation of the structural properties of images [4]. In this context the location of events such as lines or points has been identified as a property that cannot be represented by the Fourier amplitude, i. e. by second-order statistics, but requires the availability of Fourier phase. We have more recently extended this critique by suggesting that an further important deficit of the second-order approach has to be seen in its inability for the representation of *oriented structure* (*i1D*-signals) [10]. In order to illustrate this, we have constructed an orientation-only image that is full of oriented structure but lacks any second-order dependencies, i.e. it is 'white' (Fig. 1). Since second-order theory is insufficient to capture the oriented structure in this image, we have concluded that 'local orientedness' is a higher-order statistical property [10]. This point will be pursued further in the present paper.

3. Higher-order statistics of natural images

Surprisingly little is known about the higher-order statistics of natural images. In previous studies we noted substantial statistical dependencies remaining between the (second-order uncorrelated) outputs of orientation selective filter decompositions (wavelets) [6] [7] [10] [11]. Here we use higher-order spectra to investigate these dependencies in a more systematic way. The bispectrum $C_3^U(f_{x1}, f_{y1}, f_{x2}, f_{y2})$ is given by the Fourier transform of the third-order cumulant $c_3^u(x_1, y_1, x_2, y_2)$ of a stationary random process $\{u(x,y)\}$. Alternatively, the Fourier-Stieltjes representation of this process offers the possibility to express the bispectrum directly in terms of the components $dU(f_x, f_y)$ [3]:

$$E[dU(f_{x1}, f_{y1}) \cdot dU(f_{x2}, f_{y2}) \cdot dU^*(f_{x3}, f_{y3})] =$$

$$= \begin{cases} C_3^U(f_{x1}, f_{y1}, f_{x2}, f_{y2}) \cdot df_{x1} df_{y1} df_{x2} df_{y2} & \text{if } \binom{f_{x3}}{f_{y3}} = \binom{f_{x1}}{f_{y1}} + \binom{f_{x2}}{f_{y2}} \\ 0 & \text{else} \end{cases}$$

From this equation it is apparent that the bispectrum is a measure for the statistical dependencies between three frequency components, the sum of which equals zero. A direct computation in the frequency domain can also be derived from this notation [3].

We have extensively investigated the bispectra of natural images and compared them to those of noise images with almost identical first and second-order statistics. Only for the natural images we found a concentration of bispectral 'energy' exactly in those regions where the frequency components are aligned to each other with respect to their orientation (an example is shown in Fig. 2). Systematic statistical dependencies of a similar type have also shown up in our preliminary investigations using trispectra (see Fig. 3). The trispectral 'energy' is concentrated around the aligned frequency triples for natural images as it is for samples from the orientation-only process, whereas it shows a substantially different distribution for random images which are equivalent to the natural ones in their 1st-order pdf and 2nd-order acf. This prompts us to conclude that the predominance of oriented *i1D*-structures in natural images is reflected in the concentration of polyspectral 'energy' of co-oriented frequency-tuples.

4. Exploitation of the higher-order statistics of natural images

4.1. Linear filter decompositions

If local 'orientedness' is a higher-order property not reflected in the standard second-order statistics, shouldn't nonlinear operators be required for its exploitation? This is only partially correct. While it is certainly true that an exhaustive exploitation of higher-order dependencies will, in

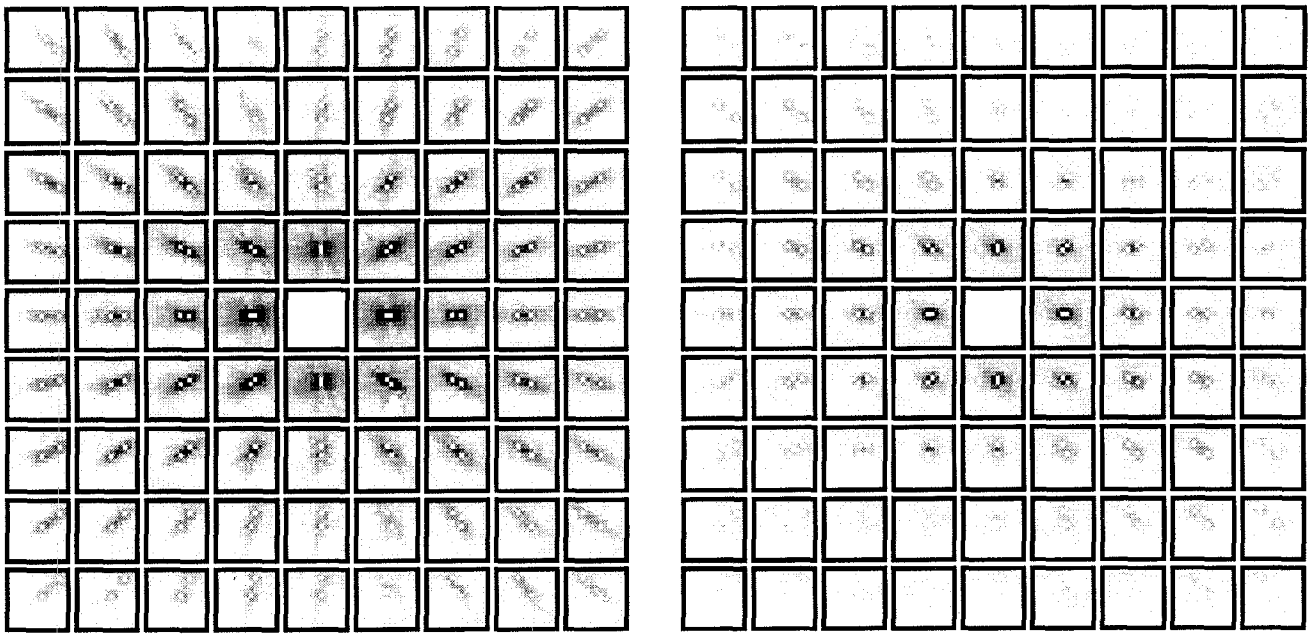

Figure 2. (a, left) Bispectral magnitude $|C_3(f_{x1}, f_{y1}, f_{x2}, f_{y2})|$ of a natural image. Shown are several sections with $f_{x1} = \text{const.}$ and $f_{y1} = \text{const.}$. Note the elongated concentration of bispectral 'energy' in regions where the frequency components are aligned to each other with respect to orientation. (b, right) Bispectrum of a noise image whose first-order pdf and second-order acf are approximately equivalent to that of the natural image. Here the 'energy'-concentration is more circular, around the void points $(f_{x1} + f_{x2} = 0, f_{y1} + f_{y2} = 0)$.

general, require nonlinear operations, a partial exploitation may well be achieved by suitable linear operators. In fact, the advantages of certain linear decompositions can only be properly conceived, once the higher-order dependencies are taken into account. We have suggested that this is the case with the encoding of images by orientation-selective decompositions like the Cosine transform or wavelet-like transforms [10].

In our view, the usual interpretation of the relationship between such transforms and the statistics of natural images requires a revision, although the current view appears to be consistent at a first glance: The KLT yields oriented basis functions, orientation obviously is an important structural property of natural images, hence usage of the optimally adapted, i.e. oriented basis functions for image compression will yield positive results. A closer look, however, reveals some problems with this interpretation.

In fact, strict application of the standard second-order reasoning would predict that the compression performance of an isotropic decomposition, like the Laplacian pyramid, should be close to the one achievable with a further orientation-selective splitting of the frequency bands [10]. This is due to the fact that no deviation from spectral flatness can be exploited along the orientation variable. However, practical experience and statistical investigations showed the contrary: substantial additional savings in data rate can be obtained if an image is not only decomposed by radial bandpass funtions but is also split up with respect to orientation , e.g. [12].

While this fact cannot be understood within the framework of second-order statistics, our results indicate that it might be explained by taking into account the structure of higher-order spectra since only orientation-selective filter decompositions can appropriately adapt to the concentration of higher-order spectral 'energy' along the '*i1D*-lines' as revealed by our measurements. An isotropic decomposition, like the Laplacian pyramid, will inevitably yield a mismatch to this specific type of structure.

4.2. Nonlinear operators

While an orientation-selective linear decomposition can partially exploit higher-order dependencies, nonlinear operators will be required in general. This raises the question of appropriate types of nonlinear operators. We will argue that *i2D*-selective operators are suitable candidates.

Starting from the Volterra-Wiener expansion of nonlinear functionals, a quite general class of *i2D*-selective operators can be derived. The Volterra series relates the input $u_1(x, y)$ of a nonlinear, shift invariant system to its output

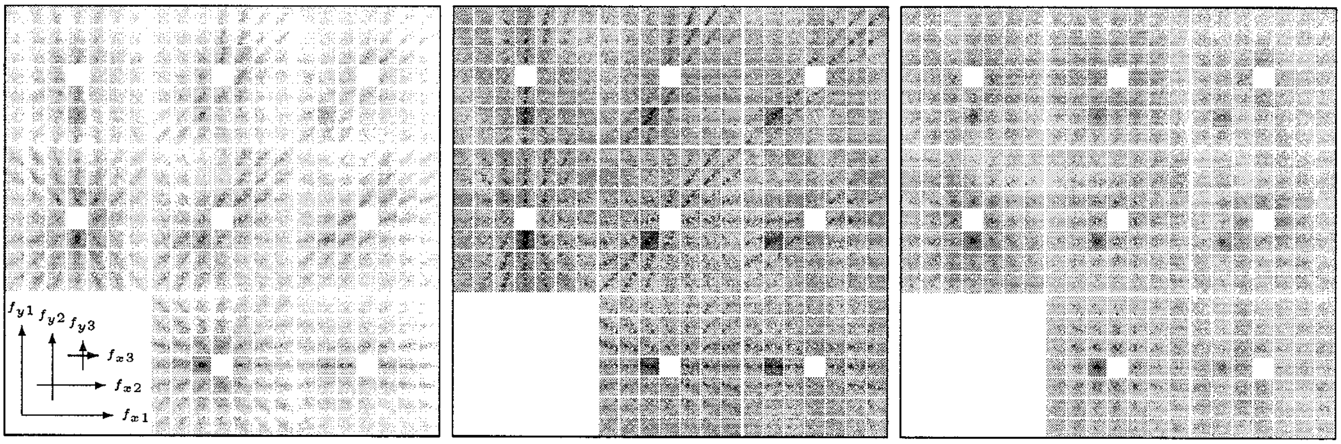

Figure 3. (a, left) Trispectral magnitude $|\mathrm{C}_4(\mathrm{f_{x1}, f_{y1}, f_{x2}, f_{y2}, f_{x3}, f_{y3}})|$ of a natural image. Note the elongated concentration of trispectral 'energy' in regions where the frequency components are aligned to each other. (b, middle) Trispectrum of the contour noise image given in Fig. 1. (c, right) Trispectrum of a noise image whose first-order pdf and second-order acf are approximately equivalent to that of the natural image. Here the 'energy'-concentration is more circular.

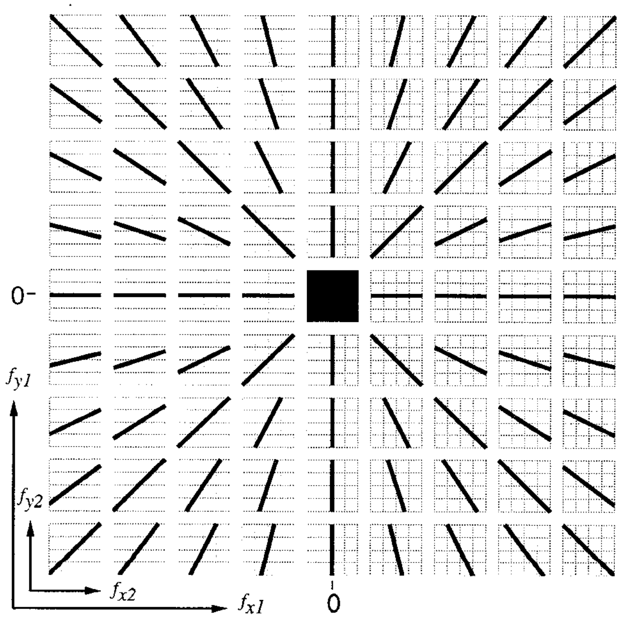

Figure 4. Illustration of the forbidden zones (indicated as black lines) for an *i2D*-selective quadratic Volterra kernel $\mathrm{H_2(f_{x1}, f_{y1}, f_{x2}, f_{y2})}$ in the frequency-domain. A system whose symmetric kernel vanishes at the forbidden zones may be regarded as blocking frequency components of equal orientation [2].

$u_2(x, y)$ in the following way:

$$\begin{aligned} u_2(x,y) &= h_0 + \qquad (1)\\ &+ \iint h_1(x_1,y_1)\cdot u_1(x-x_1,y-y_1)\cdot dx_1dy_1 + \\ &+ \iiiint h_2(x_1,y_1,x_2,y_2)\cdot u_1(x-x_1,y-y_1)\cdot \\ &\cdot u_1(x-x_2,y-y_2)\cdot dx_1dy_1dx_2dy_2 + \ldots \end{aligned}$$

The quadratic part of Eq. (2) may be expressed in the frequency domain as

$$U_2(f_x, f_y) = \iint \tilde{U}_2(f_{x1}, f_{y1}, f_x - f_{x1}, f_y - f_{y1})\; df_{x1}df_{y1}$$

where

$$\begin{aligned} &\tilde{U}_2(f_{x1}, f_{y1}, f_{x2}, f_{y2}) = \qquad (2)\\ &= H_2(f_{x1}, f_{y1}, f_{x2}, f_{y2})\cdot U_1(f_{x1}, f_{y1})\cdot U_1(f_{x2}, f_{y2}) \end{aligned}$$

is the expanded output spectrum and $H_2(f_{x1}, f_{y1}, f_{x2}, f_{y2})$ is the Fourier transform of the second-order Volterra kernel $h_2(x_1, y_1, x_2, y_2)$. Note that Eq. (2) may be regarded as the weighting of an AND-like conjunction between frequency components. A necessary and sufficient condition for a quadratic Volterra operator to be insensitive to *i0D*- and *i1D*- signals is given by [2]:

$$H_2(f_{x1}, f_{y1}, f_{x2}, f_{y2}) = 0 \quad \forall \quad f_{x1}\cdot f_{y2} = f_{y1}\cdot f_{x2}$$

The operation of systems adhering to this condition can be regarded as a blocking of aligned frequency components (see Fig. 4, an example is provided in Fig. 5). Since the redundancies of natural images are reflected in the concentration of polyspectral magnitude at aligned frequency components, *i2D*-selective systems have the potential to perform a kind of 'higher-order whitening', just as differentiating linear operators have the potential to whiten the $1/f^2$-decay of the power spectrum of natural images.

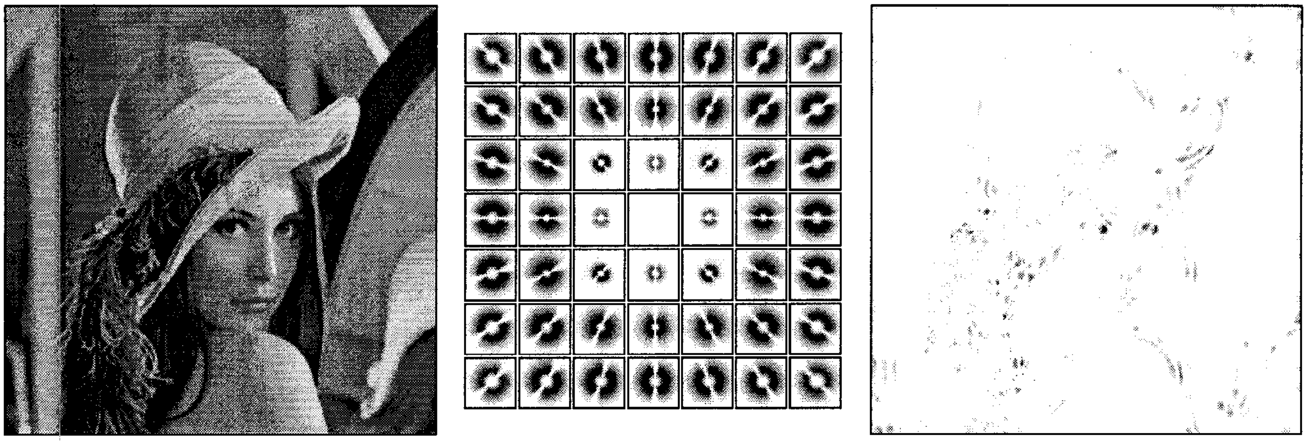

Figure 5. Application of *i2D*-selective operators to natural images: (a, left) Example of a natural image. (b, middle) Frequency-domain Volterra kernel of an *i2D*-selective operator. Note that the *i2D*-selectivity is guaranteed since the 'forbidden zones' of Fig. 4 are taken into account. (c, right) Image resulting from the application of the *i2D*-selective Volterra operator (only positive responses are shown). The original image can be reconstructed from such *i2D*-only representations if multiple scales are taken into account [1].

5. Conclusion

The frequent occurence of local oriented features is a basic structural property of natural images that cannot be captured by second-order statistics. Investigations of the polyspectra revealed that this property is reflected in the concentration of polyspectral magnitude in those regions where the frequency components are aligned in orientation. This fact can partially be exploited by orientation-selective linear filter decompositions. However, a full exploitation requires nonlinear schemes, for which *i2D*-selective operators are suitabe candidates.

Just like the distinction between a constant and a varying signal can be regarded as an elementary operation in linear signal processing, the distinction between *i1D* and *i2D*-signals can be regarded as a basic operation in the nonlinear processing of images. In this sense, *i2D*-operators can be seen as performing a 'nonlinear differentiation' beeing capable of exploiting the elementary redundancies of natural images.

References

[1] E. Barth, T. Caelli, and C. Zetzsche. Image encoding, labelling and reconstruction from differential geometry. *CVGIP:GRAPHICAL MODELS AND IMAGE PROCESSING*, 55(6):428–446, 1993.

[2] G. Krieger and C. Zetzsche. Nonlinear image operators for the evaluation of local intrinsic dimensionality. *[Special Issue on Nonlinear Image Processing] - IEEE Trans. Image Processing*, 5(6):1026–1042, 1996.

[3] C. L. Nikias and A. P. Petropulu. *Higher-Order Spectral Analysis: A Nonlinear Signal Processing Framework.* Prentice-Hall, 1993.

[4] A. V. Oppenheim and J. S. Lim. The importance of phase in signals. *Proc. IEEE*, 69:529–541, 1981.

[5] G. A. Orban. *Neuronal operations in the visual cortex.* Springer, Heidelberg, 1984.

[6] B. Wegmann and C. Zetzsche. Statistical dependence between orientation filter outputs used in a human vision based image code. In M. Kunt, editor, *Visual communication and image processing*, volume 1360 of *Proc. SPIE*, pages 909–923, Bellingham, WA, 1990.

[7] B. Wegmann and C. Zetzsche. Feature-specific vector quantization of images. *[Special Issue: Vector Quantization] - IEEE Trans. Image Processing*, 5:274–288, 1996.

[8] C. Zetzsche and E. Barth. Fundamental limits of linear filters in the visual processing of two-dimensional signals. *Vision Research*, 30:1111–1117, 1990.

[9] C. Zetzsche, E. Barth, G. Krieger, and B. Wegmann. Neural network models and the missing link between cortical orientation selectivity and the natural environment. *submitted*, 1997.

[10] C. Zetzsche, E. Barth, and B. Wegmann. The importance of intrinsically two-dimensional image features in biological vision and picture coding. In A. Watson, editor, *Digital images and human vision*, pages 109–138. MIT Press, Cambridge, MA, 1993.

[11] C. Zetzsche, E. Barth, and B. Wegmann. Nonlinear aspects of primary vision: entropy reduction beyond decorrelation (Invited Address). In J. Morreale, editor, *SID International Symposium - Digest of Technical Papers*, volume XXIV, pages 933–936. Soc. Inform. Display, Playa del Ray, CA, 1993.

[12] C. Zetzsche and W. Schönecker. Orientation selective filters lead to entropy reduction in the processing of natural images. *Perception*, 16:229, 1987.