Collinearity and Parallelism are Statistically Significant Second Order Relations of Complex Cell Responses*

Norbert Krüger

Abstract

By investigating the second order statistics of Gabor wavelet responses derived from natural images, we show that collinearity and parallelism are conspicuous relations. We give a precise mathematical characterization of these Gestalt principles by the conditional probability of two responses. Essential for our investigations is a non–linear transformation, initially utilized within the object recognition system [5], which transforms continuous Gabor wavelet responses into a binary code indicating the presence or absence of local oriented line segments.

Keywords: Natural images, Gabor Wavelets, Gestalt Principles, Learning, Contextual Information, Banana Wavelets,

Norbert Krüger
Institut für Neuroinformatik
Ruhr–Universität Bochum
44801 Bochum
Universitätsstraße 150
Germany

*Supported by grants from the German Ministry for Science and Technology 01IN504E9 (NEUROS) and 01M3021A4 (Electronic Eye).

Collinearity and Parallelism are Statistically Significant Second Order Relations of Complex Cell Responses

Abstract

By investigating the second order statistics of Gabor wavelet responses derived from natural images, we show that collinearity and parallelism are conspicuous relations. We give a precise mathematical characterization of these Gestalt principles by the conditional probability of two responses. Essential for our investigations is a non–linear transformation, initially utilized within the object recognition system [5], which transforms continuous Gabor wavelet responses into a binary code indicating the presence or absence of local oriented line segments.

1 Introduction

A lot of evidence exists for the assumption that Gabor wavelets (see, e.g., Daugman [1]) play an important role in the first stages of visual processing of humans. Jones et al. [4] recorded neurons with Gabor wavelet like sensitivity in V1 and Gabor wavelet like filters can be learned from the statistics of natural images by applying the restrictions "information preservation" and "sparse coding" (Ohlshausen et al. [7]). There is also strong evidence that the human object recognition system processes the visual input through a couple of stages in which features of increasing complexity are extracted and in which Gabor wavelets represent only an early stage of processing (see, e.g., Hummel at al. [3] and Oram et al. [8]). Evaluating the second order statistics of Gabor wavelet responses we give statistical evidence for important second order relations for the class of natural images and therefore we support the understanding of the stages of the visual system beyond the extraction of Gabor wavelets responses.

On the one hand there exists little knowledge about the statistical properties of the class of natural images. As the most precise quality Field [2] showed that the amplitude of the power spectrum of natural images decreases with frequency ($1/f$–law) and that this statistical quality holds true for at least three octaves of scaling. On the other hand it is known for a long while that humans utilize grouping mechanisms in their perception called "Gestalt principles" (see e.g., Ellis [9]). Hummel and Biederman [3] showed that there is a mutual dependence of the occurence of local oriented line segments for line drawings of randomly generated triangles. Here we go beyond this result by characterizing the Gestalt principles collinearity and parallelism in a mathematical framework from the second order statistics of Gabor wavelet responses. By showing that pairs of Gabor wavelet responses derived from natural images have a characteristic conditional probability distribution corresponding to the Gestalt principles mentioned above we describe an additional statistical quality of natural images.

Essential for the extraction of Gestalt principles from natural images is a non–linear transformation which transforms continuous Gabor wavelet responses (see figure 1b) to the binary code indicating the presence or absence of local oriented line segments (see figure 1d). This corresponds to a transformation of a quantitative description to a more qualitative description of an image. Without this transformation, looking only at the second order statistics of the magnitude of Gabor wavelet responses, the Gestalt principles collinearity and parallelism are barely detectable. Initially

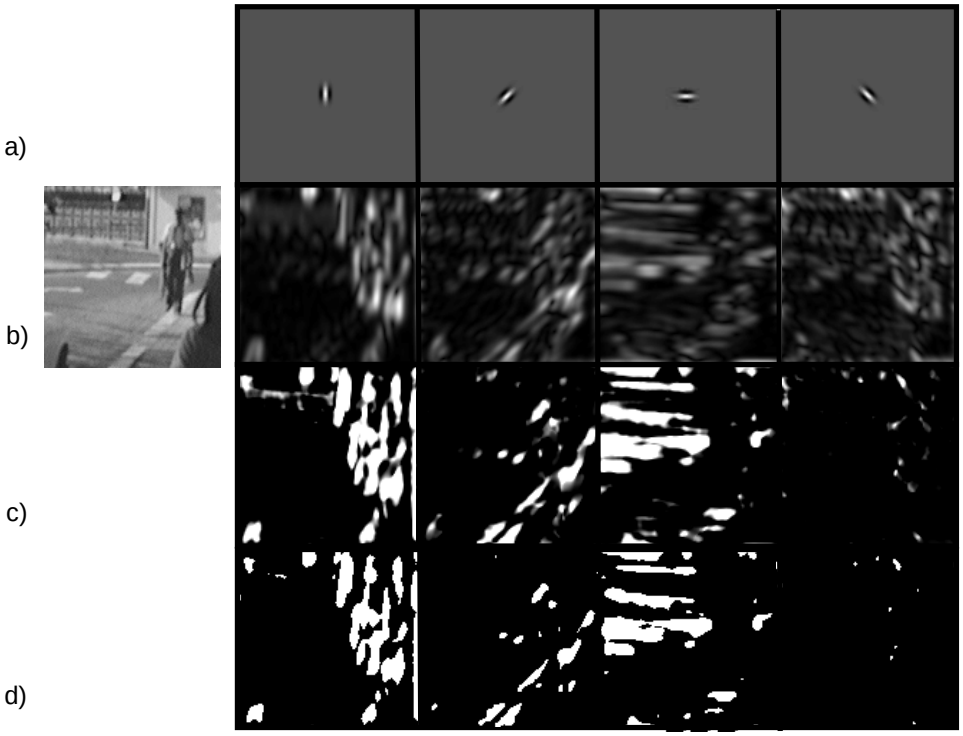

Figure 1: Gabor Transformation a) Gabor wavelets of four different orientations. b) An image and the Gabor responses corresponding to the kernels in a). c) Normalized response after applying the normalization function N. d) Binarized response after thresholding.

the transformation was derived to match a sparse and binary object representation with an image (see figure 2 and [5]). We see the ability to detect these important Gestalt relations with this non-linear transformation as a positive example of task–driven approaches yielding answers to more fundamental questions.

2 Second Order Statistics of Gabor Wavelet Responses

Gabor wavelets (see figure 1a) can be parameterized by a vector $\vec{c} = (x, y, f, \alpha)$. The tupel (x, y) gives the position of the Gabor wavelet's center in the image, f the frequency of the wave function and α its orientation. In all of our simulations we use only one frequency level because we are only interested in the relation of line segments corresponding to similar scale. We apply a slight modification of Gabor wavelets. Instead of a Gaussian with $\sigma_x = \sigma_y$ as envelope function we choose stretched the envelope function perpendicular to the orientation of the wave function by a factor 1.5. In this way we achieve a more specific tuning to local oriented lines. Here a Gabor wavelet response $r(\vec{c}, I) = \left| G^{\vec{c}} * I \right|$ is defined as the magnitude of the convolution of the Gabor wavelet $G^{\vec{c}}$ with image I. Figure 1b) shows a Gabor transformation with four orientations of an 128×128 image. The space of second order relations consists of $(128^2 \cdot 4)^2 \cong 4 \cdot 10^9$ elements. In order to reduce this huge number we restrict ourselves to the relative distance of two Gabor wavelets assuming that the second order statistics of natural images are translation invariant. This implies that we compute our statistics over all tupel of Gabor wavelets which have same distance vector. In other words, we introduce the equivalence relation "$\stackrel{T}{\equiv}$":

$$(\vec{c}_1, \vec{c}_2) \stackrel{T}{\equiv} (\vec{c}_1', \vec{c}_2') \Longleftrightarrow (x_1 - x_2, y_1 - y_2, f_1, f_2, \alpha_1, \alpha_2) = (x_1' - x_2', y_1' - y_2', f_1', f_2', \alpha_1', \alpha_2').$$

Let now $\mathcal{I}$ be a set of images. In all our simulations we used an image data base from British

Figure 2: A learned representation of a face. In [5] a large number of objects (such as, hand postures, faces, and cans) are coded by localized curved line segments allowing for a fast and efficient detection of these objects in complicated scenes.

Telecom[1] consisting of 98 pictures of natural images showing outdoor scenes with and without human made objects like houses and cars (see figure 3a). The images have the pixel format 512×512. By extracting sub–images of size 128×128 we can produce a large set of training images used for our statistics (see figure 3b). In our simulation we used 900 randomly selected sub–images.

We compute the correlation of two Gabor wavelet responses

$$\begin{aligned} Cor^{(\overset{T}{\equiv},\mathcal{I})}(\vec{c}_1,\vec{c}_2) &= Cor^{(\overset{T}{\equiv},\mathcal{I})}((0,0,f,\alpha_1),(x_2,y_2,f,\alpha_2)) \\ &= \frac{\sum_{(\mathcal{I},\overset{T}{\equiv})}(r(\vec{c}_1,I)- <r(\vec{c}_1)>)\cdot(r(\vec{c}_2,I)- <r(\vec{c}_2)>)}{\sqrt{\sum_{(\mathcal{I},\overset{T}{\equiv})}(r(\vec{c}_1,I)- <r(\vec{c}_1)>)^2}\cdot\sqrt{\sum_{(\mathcal{I},\overset{T}{\equiv})}(r(\vec{c}_2,I)- <r(\vec{c}_2)>)^2}} \end{aligned} \quad (1)$$

where $r(\vec{c},I)$ represents the response if the Gabor wavelet $G^{\vec{c}}$ for image $I \in \mathcal{I}$ and $<r(\vec{c})>$ represents the mean response of $\vec{c}$ on all images. Because of the translation invariance of the statistics of natural images $<r(\vec{c})>$ does not depend on the position (x,y) but only on the angle α (see also table 1). The expression $\sum_{(\mathcal{I},\overset{T}{\equiv})}$ indicates that we sum over all instances $(\vec{c}_1,\vec{c}_2)$ in $\mathcal{I}$ for which the eqivalence relation "$\overset{T}{\equiv}$" holds true. Because we are only interested in the relative distances of two Gabor wavelets, i.e., because of our equivalence relation "$\overset{T}{\equiv}$", we can set $x_1 = y_1 = 0$. We investigate the correlation for Gabor wavelets with a maximal distance of 40 pixels. For 4 orientations and one frequency level we have $4^2 \cdot 81^2 = 104976$ equivalence classes described by the parameters $\alpha_1^i, \alpha_2^j : 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$ and $x_2, y_2 : -40, \ldots, 40$. For each tupel $(\vec{c}_1,\vec{c}_2)$ we extract in each image 729 values falling in the same equivalence class, i.e., our estimate of $Cor^{(\overset{T}{\equiv},\mathcal{I})}(\vec{c}_1,\vec{c}_2)$ is based on more than 300.000 data values. Figure 4) shows the result. For each pair of orientations, i.e., one contour plot in figure 4, the computation of $Cor^{(\overset{T}{\equiv},\mathcal{I})}(\vec{c}_1,\vec{c}_2)$ takes approximately 3 hours on a Sun UltraSparc (167Mhz).

Most interesting are the results for which $\vec{c}_1,\vec{c}_2$ have the same orientation. We can see a distinct expansion along and perpendicular to the orientation of the kernel corresponding to the fact that collinear edges and parallel edges occur with statistically significant frequency. In case of horizontal and vertical orientation this expansion is most distinct. Furthermore there is always positive correlation, expressing that the occurrence of structure at one position makes the occurrence of structure nearby more likely.

[1] Access to these images may be obtained by anonymous ftp to site ftp.teleos.com. The images may be found in the subdirectory VISION-LIST-ARCHIVE/IMAGERY/BT_scenes.

Figure 3: a) Original Images of the BT data base. b) Extracted subimages. c) Rotated subimages.

To ensure that these effects are not caused by the correlation of the kernels themselves we compute the correlation on noise with $\frac{1}{f}$–power spectrum (see figure 5). Here the correlation is very different compared to natural images. Firstly correlation drops down to zero where the kernels do not overlap, i.e., there is no *global dependency* of the correlation for $\frac{1}{f}$–noise images. Secondly the expansion along the orientation of the kernel is much weaker compared to natural images. Thirdly the second order statistics of $\frac{1}{f}$–noise are rotation invariant.

Now we apply a non-linear transformation to the Gabor responses to translate the continuous response to the pure binary information indicating the presence or absence of local line segments with certain orientation (see figure 1d). We say

AAC a line segment corresponding to the Gabor wavelet $G^{\vec{c}}$ is present if the corresponding Gabor wavelet response is distinctly above the average response.

We call this criterion the "Above Average Criterion" which can be formalized by first applying a normalization function $N(t)$ to the Gabor wavelet response transforming $r(\vec{c}, I)$ to the interval $[0, 1]$. The normalized response $N(r(\vec{c}, I))$ represents a confidence in the interval $[0, 1]$ indicating the presence or absence of a line segment corresponding to a certain Gabor wavelet parametrized by $\vec{c}$ (see figure 1c). We worked with different normalization functions and different parameters. The results do not depend crucial on our specific choices as long as they formalize our criterion AAC. The normalization function used for our simulations is described in the appendix. A similar normalization function is applied in the object recognition system[2] described in Krüger et al. [5].

[2]In this system an object is represented as a sparse and spatially organized arrangement of curved local line

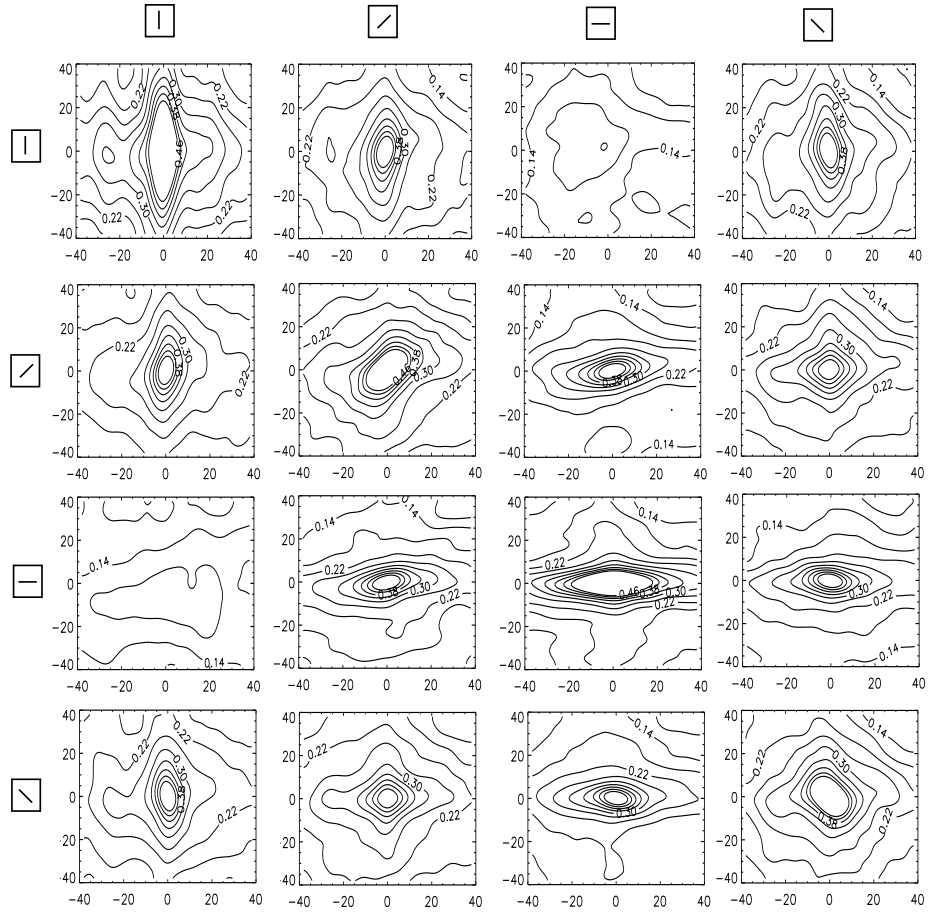

Figure 4: Correlation of Gabor wavelet responses. 1^{st} line: $Cor((0,0,f,0),(x,y,f,\alpha_2^j))$ for α_2^j : $0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$. $2^{nd}, 3^{rd}, 4^{th}$ line: $Cor((0,0,f,\alpha_1^i),(x,y,f,\alpha_2^j))$ for α_1^i : $\frac{\pi}{2}, \pi, \frac{3\pi}{2}$ and α_2^j : $0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$. The value with coordinate (x_0, y_0) in each contour plot represents $Cor((0,0,f,\alpha_1^i),(x_0,y_0,f,\alpha_2^j))$. Each plot is smoothed by a Gaussian of 3 pixel width.

A simple thresholding

$$\tilde{r}(\vec{c}, I) = N(r(\vec{c}, I)) > 0.75 \tag{2}$$

transforms the normalized response to binary values (see figure 1d) indicating the presence or absence of local line segments with a certain orientation and position. Figure 6 shows the correlations of the binarized responses. As in figure 4 collinearity and parallelism are conspicuous relations. But in contrast to the correlation of unmodified Gabor wavelet responses parallelism is a *global property*: The correlation of binarized responses of same orientation is always higher than the correlation of responses of different orientation. In other words, the occurrence of a parallel line segment with 40 pixel distance is more likely than the occurrence of a line segment with different orientation and smaller distance.

We can look at the binarized Gabor wavelet responses as discrete events and we can calculate

$$\mathcal{G}^{(\overset{T}{\equiv},\mathcal{I})}(\vec{c}_1,\vec{c}_2) := \frac{P(\tilde{r}(\vec{c}_2,I)|\tilde{r}(\vec{c}_1,I))}{P(\tilde{r}(\vec{c}_2,I))} = \frac{P(\tilde{r}(\vec{c}_2,I)\cap\tilde{r}(\vec{c}_1,I))}{P(\tilde{r}(\vec{c}_1,I))\cdot P(\tilde{r}(\vec{c}_2,I))} = \frac{\frac{1}{n}\sum_{(\mathcal{I},\equiv)}\tilde{r}(\vec{c}_1,I)\cdot\tilde{r}(\vec{c}_2,I)}{\frac{1}{n}\sum_{(\mathcal{I},\equiv)}\tilde{r}(\vec{c}_1,I)\cdot\frac{1}{n}\sum_{(\mathcal{I},\equiv)}\tilde{r}(\vec{c}_2,I)} \tag{3}$$

segments (see figure 2). The normalization mediates between the continuous wavelet responses and the binary object representations.

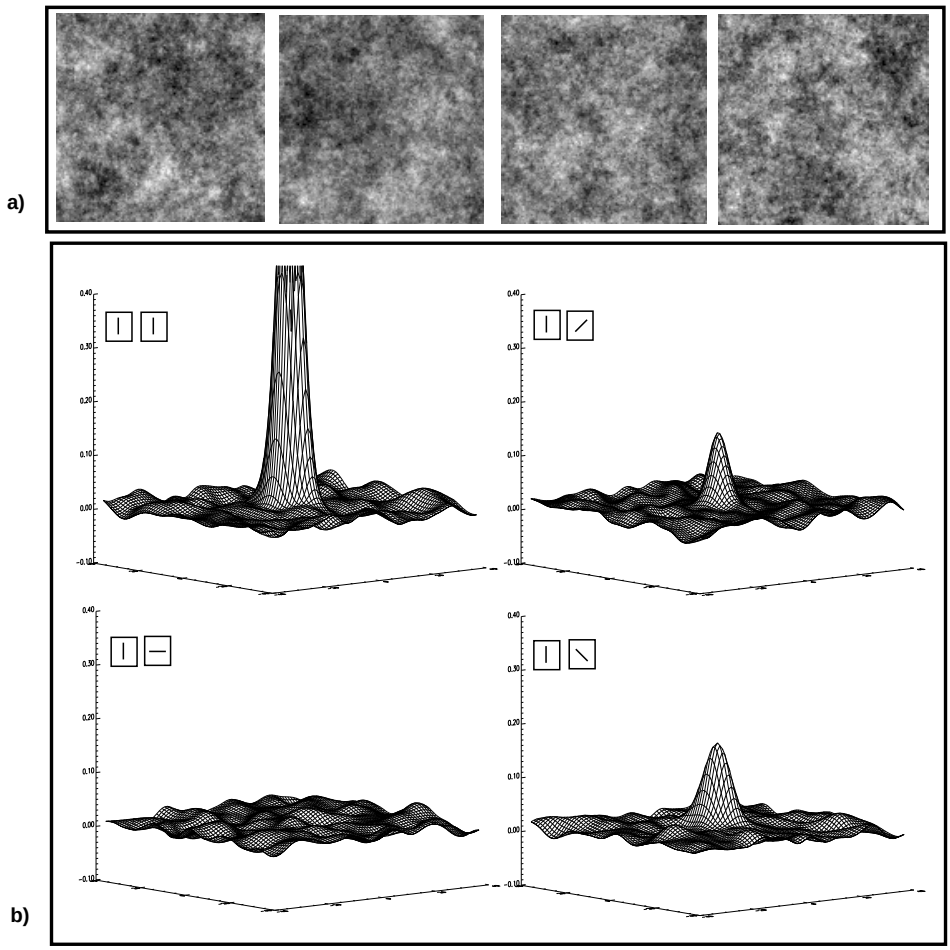

Figure 5: a) $\frac{1}{f}$–noise images. b) $Cor((0,0,f,0),(x,y,f,\alpha_2^j))$, $\alpha_2^j : 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$ for $\frac{1}{f}$–noise images.

which expresses the influence of occurrence of a line segment corresponding to $\vec{c}_1$ to the occurrence of the line segment corresponding to $\vec{c}_2$ (n is the number of evaluated tupels $(\vec{c}_1, \vec{c}_2)$). If $\mathcal{G}(\vec{c}_1, \vec{c}_2)$ is larger than one the occurrence of a line segment corresponding to $\vec{c}_2$ is more likely when a line segment corresponding to $\vec{c}_1$ is present. In case $\mathcal{G}(\vec{c}_1, \vec{c}_2) = 2$ we can say the likelihood has doubled. We call $\mathcal{G}(\vec{c}_1, \vec{c}_2)$ the Gestalt coefficient[3]. Figure 7 shows the result. We recognize that the likelihood of the presence of a certain line segment is strongly influenced by the presence of other line segments. The Gestalt principles "collinearity" and "parallelism" are described by the Gestalt coefficient as statistically significant relations.

The second order statistics of natural images are not rotation invariant. Not only the Gestalt principles collinearity and parallelism are most conspicuous for horizontal and vertical orientation. Also occur horizontal and vertical local lines more often than lines of diagonal orientation. Table 1 shows the mean responses for the non normalized and binarized Gabor wavelet responses of different orientations for natural images and $\frac{1}{f}$–noise.

[3]It can be shown that the Gestalt coefficient $\mathcal{G}^{(\overset{T}{\equiv},\mathcal{I})}$ can be computed from the correlation of the binarized Gabor wavelet responses by the formula

$$\mathcal{G}^{(\overset{T}{\equiv},\mathcal{I})} = c \cdot Cor^{(\overset{T}{\equiv},\mathcal{I})} + 1 \tag{4}$$

where c does only depend on α_1 and α_2.

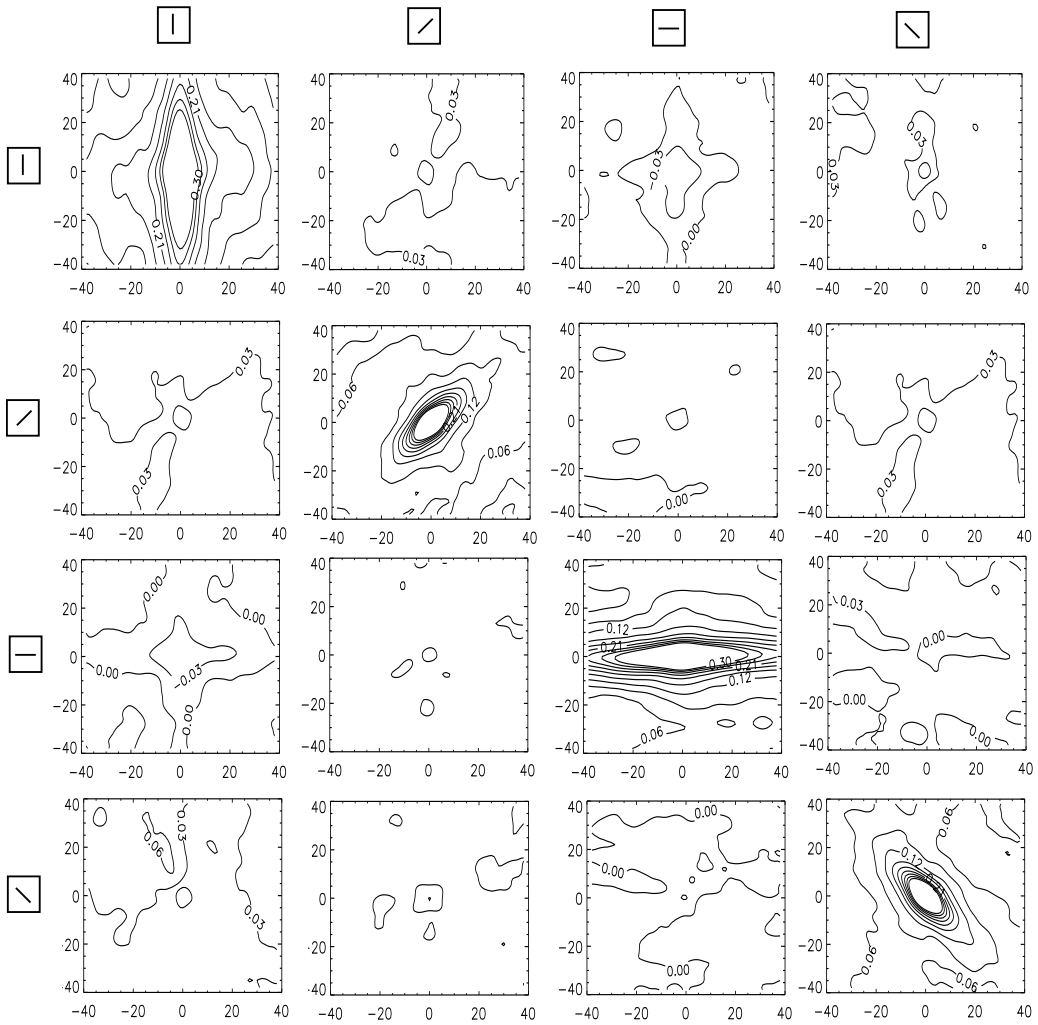

Figure 6: Correlation of binarized Gabor wavelet responses.

α	0		$\frac{\pi}{2}$		π		$\frac{3\pi}{2}$	
natural images / $\frac{1}{f}$ noise	nat	$\frac{1}{f}$	nat	$\frac{1}{f}$	nat	$\frac{1}{f}$	nat	$\frac{1}{f}$
$<r(\vec{c})>$	6047	7705	4529	7719	4871	7705	4529	7704
$<\tilde{r}(\vec{c})>$	0.11	0.02	0.03	0.02	0.07	0.02	0.03	0.02

Table 1: The mean value for non–normalized and binarized Gabor wavelet responses.

To neglect differences resulting from this non-isotropic structure of natural images according to rotation we introduce an additional equivalence relation "$\overset{T\circ R}{\equiv}$" representing translation and rotation invariance:

$$(\vec{c}_1,\vec{c}_2)\overset{T\circ R}{\equiv}(\vec{c}_1{}',\vec{c}_2{}') \Longleftrightarrow (x_1-x_2,y_1-y_2,f_1,f_2,\alpha_1-\alpha_2)=(x_1'-x_2',y_1'-y_2',f_1',f_2',\alpha_1'-\alpha_2')$$

By this equivalence relation we can reduce the number of equivalence classes to $4\cdot 81^2 = 26244$ described by the parameters $\alpha_2^j : 0,\frac{\pi}{2},\pi,\frac{3\pi}{2}$ and $x_2,y_2 : -40,\ldots,40$. In our implementation we realize this equivalence relation by randomly rotation of the pictures used for our statistics (see figure 3c). Figure 8 shows $Cor(\vec{c}_1,\vec{c}_2)$ and $\mathcal{G}(\vec{c}_1,\vec{c}_2)$ for $\overset{T\circ R}{\equiv}$ as contour plots. Figure 9a) shows all four plots in the first line of figure 8 as surface plots. The surface on the top of the other three surfaces corresponds to Gabor wavelets with same orientation.

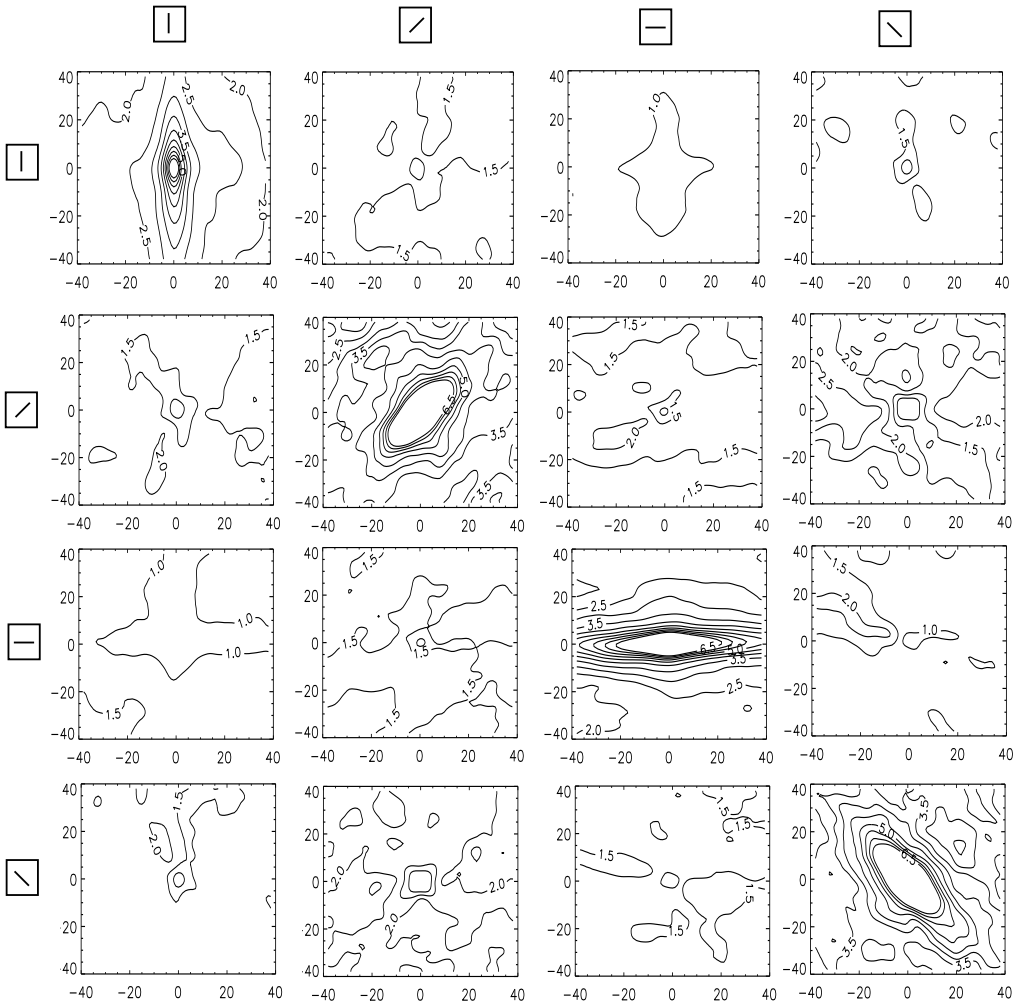

Figure 7: The Gestalt coefficient.

3 Collinearity and Parallelism are Significant Relations

Now we can characterize the Gestalt principles "collinearity" and "parallelism" in terms of the Gestalt coefficient.

Collinearity: Two complex cells $\vec{c}_1, \vec{c}_2$ are collinear if

$$\mathcal{G}^{(\overset{T \circ R}{\equiv}, \mathcal{I})}(\vec{c}_1, \vec{c}_2) > 4.7. \tag{5}$$

Parallelism: Two complex cells $\vec{c}_1, \vec{c}_2$ are parallel if

$$2.2 \leq \mathcal{G}^{(\overset{T \circ R}{\equiv}, \mathcal{I})}(\vec{c}_1, \vec{c}_2) \leq 4.7. \tag{6}$$

Figure 10 shows line segments corresponding to different intervals $[r_1, r_2]$ with $\mathcal{G}(\vec{c}_1, \vec{c}_2) \in [r_1, r_2]$.

4 Conclusion

We showed that statistically significant second order relations in natural images exist and that they represent the Gestalt principles collinearity and paralellism. They can be characterized mathematically by the Gestalt coefficient. The detection of these important relations is facilitated by a

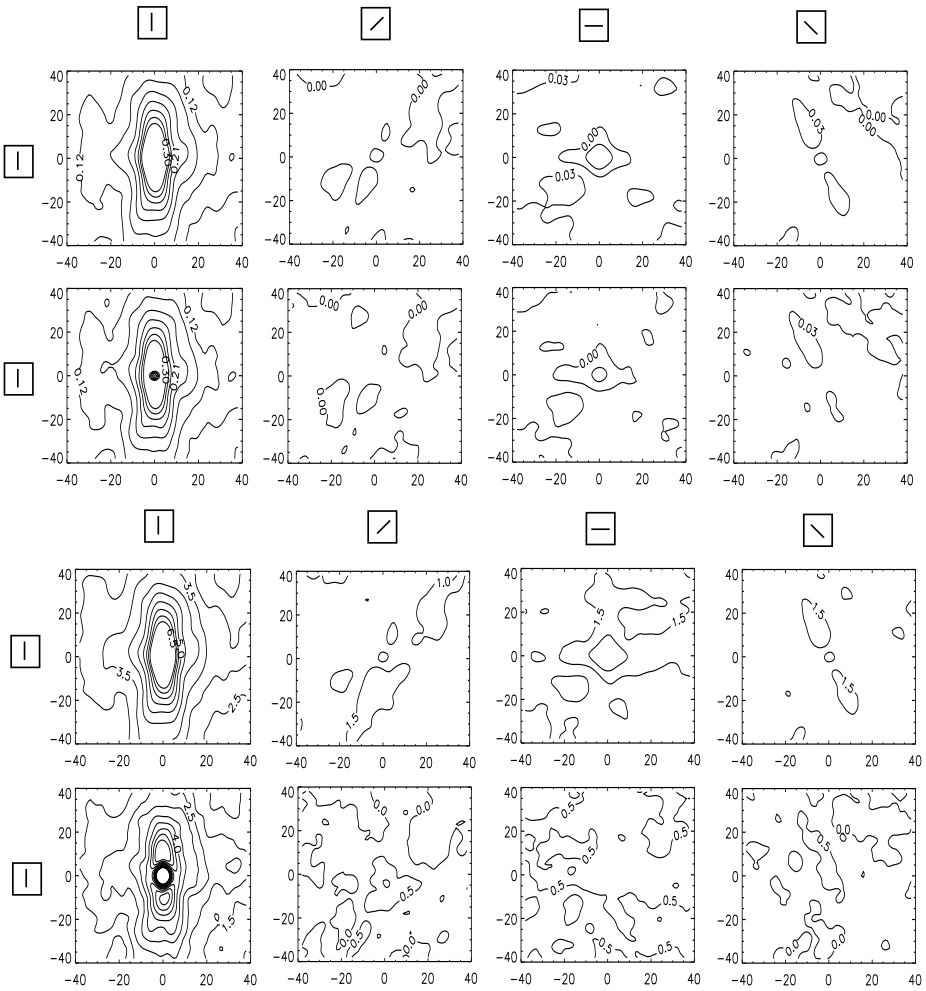

Figure 8: The correlation and the Gestalt coefficient for randomly rotated images as contour plots. 1^{st}–line: $Cor^{(\overset{T\circ R}{\equiv},\mathcal{I})}((0,0,f,0),(x,y,f,\alpha_2^j))$ for $\alpha_2^j : 0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}$ 2^{nd}–line: Correlation in first line minus correlation on $\frac{1}{f}$–images. 3^{rd} and 4^{th}–line: Same as in the 1^{st} and 2^{cd}–line for the Gestalt coefficient.

transformation of the continuous Gabor wavelet reponse to a binary and more complex feature: a local oriented line segment. In the object recognition system developped by Krüger et al. [5] this kind of normalization is utilized to mediate between a binary object representation and the continuous wavelet responses. In a next step we aim to incorporate the Gestalt principles characterized here as grouping mechanisms into this object recognition system (see also [6]).

Acknowledgement: I like to thank Christoph von der Malsburg, Irving Biederman, Rolf Würtz and Michael Pötzsch for fruitful discussions. Many thanks to Rolf Würtz for deriving equation (4).

Appendix: The Normalization Function

For our formalization of the AAC we define

$$\text{Mean}(x,y,I,\mathcal{I}) := \frac{1}{2}\text{Mean}^{local}(x,y,I) + \frac{1}{2}\text{Mean}^{global}(\mathcal{I})$$

as an average response. $\text{Mean}^{global}(\mathcal{I})$ is the average of all Gabor wavelet responses in a large set of images $\mathcal{I}$ and $\text{Mean}^{local}(x,y,I)$ the average of all Gabor wavelet responses at at pixel position

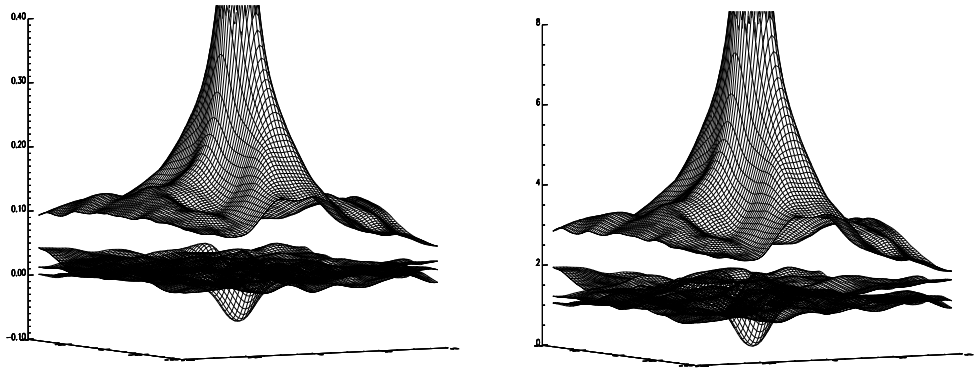

Figure 9: The correlation (left) and the Gestalt coefficient (right) for randomly rotated images as surface plots. The four plots in the first and third line in figure 8 are displayed within one plot.

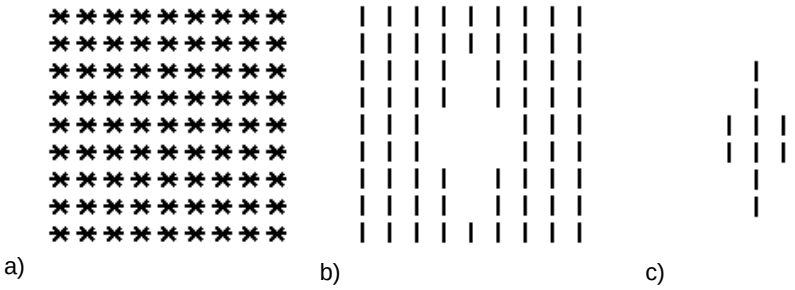

Figure 10: Line segments corresponding to the Gestalt coefficient taking values in the intervals a) $[0, 2.2)$, b) $[2.2, 4.7)$ and c) $[4.7, \infty)$

(x, y) in image I. To improve robustness we neglect the maximal and minimal local response for the calculation of $\text{Mean}^{local}(x, y, I)$. $\text{Max}(x, y, I)$ is the maximum of all Gabor wavelet responses at pixel position (x, y) in image I. We define a sigmoid function (see figure 11) $N^{(x,y,I,\mathcal{I})}(t)$ by

$$N^{(x,y,I,\mathcal{I})}(t) = \begin{cases} 0 & \text{for} \quad t < \theta \cdot \text{Mean}(x, y, I, \mathcal{I}) \\ s \cdot t - s \cdot \theta \cdot \text{Mean}(x, y, I, \mathcal{I}) & \text{for} \quad \begin{cases} t \geq \theta \cdot \text{Mean}(x, y, I, \mathcal{I}) \text{ with} \\ s = \max\{\frac{1}{\text{Max}(x,y,I) - \theta \cdot \text{Mean}(x,y,I,\mathcal{I})}, \frac{1}{2 \cdot \theta \cdot \text{Mean}(x,y,I,\mathcal{I})}\} \end{cases} \end{cases}$$

with $\theta = 1.2$. This normalization function sets all Gabor wavelet responses not significantly above average to zero and applies to the other Gabor wavelet reponses a linear function with a minimal slope of $\frac{1}{2 \cdot \theta \cdot \text{Mean}(x,y,I,\mathcal{I})}$ which transforms the Gabor response into the interval $[0, 1]$.

References

[1] J.G. Daugman, "Uncertainty Relation for Resolution in Space, spatial Frequency, and Orientation optimized by 2D Visual cortical Filters", Journal of the Optical Society of America vol. 2

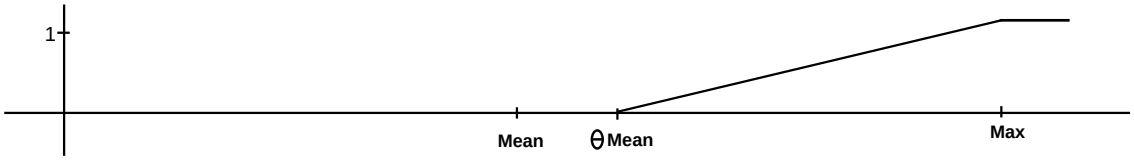

Figure 11: The sigmoid function $N^{(x,y,I,\mathcal{I})}(r(\vec{c}, I))$ transforming the Gabor wavelet reponse to the interval $[0, 1]$.

(7), pp. 1160-1169, 1985.

[2] D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," Journal of the Optical Society of America, vol. 4, no. 12, pp. 2379–2394, 1987.

[3] J.Hummel and I. Biederman, "Dynamic binding in a neural network for shape recognition", Psychological Review, 99, 480-517, 1992.

[4] J.P.Jones and L.A. Palmer,"An evaluation of the two dimensional gabor filter model of simple receptive fields in striate cortex", Journal of Neurophysiology, 58(6):1223–1258, 1987.

[5] N. Krüger and G. Peters, "Object Recognition with Banana Wavelets", Proceedings of the ESANN 97, p: 61–66, 1997.

[6] N. Krüger, M. Pötzsch and G. Peters, "Principles of Cortical Processing Applied to and Motivated by Artificial Object Recognition", accepted for R. Baddeley, P. Hancock and P. Foldiak (ed.) Information Theory and the Brain, Cambridge University Press (in press).

[7] B.A. Olshausen and D. Field, "Natural Image Statistics and efficient Coding", Network, 7: 333–339 (1996).

[8] M.W. Oram and D.I Perrett, "Modeling Visual recognition from neurobiological Constraints", Neural Networks Vol.7, pp:945–972 (1994).

[9] W.D. Ellis (ed.) Gestalt Theory, A source book for Gestalt Psychology, New York 1938.