Multi–Modal Statistics of Edges in Natural Image Sequences

Norbert Krüger and Florentin Wörgötter
University of Stirling, Scotland, norbert{worgott}@cn.stir.ac.uk

Abstract

In this work we investigate the multi-modal statistics of natural image sequences looking at the modalities orientation, color, optic flow and contrast transition. It turns out the statistical interdependencies corresponding to the Gestalt law collinearity increase significantly when we look not at orientation only

1 Introduction

A large amount of research has been focused on the usage of Gestalt laws in computer vision systems (overviews are given in [14, 13]). The most often applied and also the most dominant Gestalt principle in natural images is collinearity [3, 9]. Collinearity can be exploited to achieve more robust feature extraction in different domains, such as, edge detection (see, e.g., [7, 8]) or stereo estimation [2, 13]. In most applications in artificial visual systems, the relation between features, i.e., the applied Gestalt principle, has been defined heuristically based on semantic characteristics such as orientation or curvature. Mostly, explicit models of feature interaction have been applied, connected with the introduction of parameters to be estimated beforehand, a problem recognized as extremely awkward in computer vision. Recently, Geisler et al [6] introduced the idea to overcome heuristic and explicit models by relating feature interaction to the statistics of natural images. The feasibility of this approach becomes strong support from the *measurable interdependencies* of features in visual scenes that turn out to correspond to Gestalt laws [9, 3, 6].

In the human visual system beside local orientation also other modalities such as color and optic flow are computed (see, e.g. [5]). Gestalt principles are

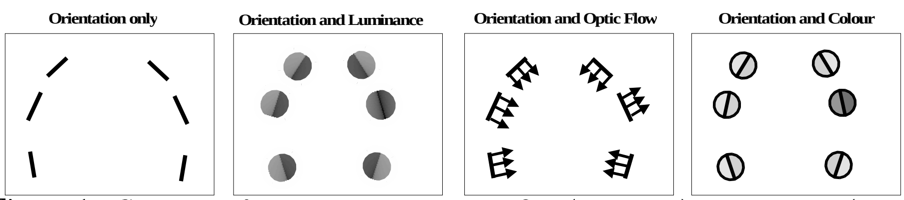

Figure 1: Grouping of entities becomes intensified (left triple) or weakened (right triple) by using additional modalities.

affected by multiple modalities. For example, figure 1 shows how collinearity can be intensified by the different modalities contrast transition, optic flow and color. This paper addresses statistics of natural images in these modalities. As a main result we found that statistical interdependencies corresponding to the Gestalt law "collinearity" in visual scenes become significantly stronger when multiple modalities are taken into account (see section 2).

2 Multi–Modal Statistics in Image Sequences

In the work presented here we address the multi–modal statistics of natural images. We start from a feature space (see also figure 1) containing the submodalities:

Orientation: We compute local orientation o (and local phase p) by the specific isotropic linear filter [4].

Contrast Transition: The contrast transition of the signal is coded in the phase p of the same filter.

Color: Color is processed by integrating over image patches in coincidence with their edge structure (i.e., integrating over the left and right side of the edge separately). Hence, we represent color by the two tuples $(c_r^l, c_g^l, c_b^l), (c_r^r, c_g^r, c_b^r)$ representing the color in RGB space on the left and right side of the edge.

Optic Flow: Local displacements (f_1, f_2) are computed by a well known optical flow technique ([11]).

2.1 Measuring Statistical Interdependencies:

We measure statistical interdependencies by the so called 'Gestalt coefficient' (see also [9]). The Gestalt coefficient is defined by the ratio of the likelihood of an event e^1 given another event e^2 and the likelihood of the event e^1:

$$G(e^1, e^2) = \frac{P(e^1|e^2)}{P(e^1).} \tag{1}$$

For the modeling of feature interaction a high Gestalt coefficient is helpful since it indicates the modification of likelihood of the event e^1 depending on other events. A Gestalt coefficient of one says, that the event e^2 does not influence the likelihood of the occurrence of the event e^1. A value smaller than one indicates a negative dependency: the occurrence of the event e^2 reduces the likelihood that e^1 occurs. A value larger than one indicates a positive dependency: the occurrence of the event e^2 increases the likelihood that e^1 occurs. The Gestalt coefficient is illustrated in figure 2. Further details can be found in [10].

2.2 Second Order Relations Statistics of Natural Images

A large amount of work has addressed the question of efficient coding of visual information and its relation to the statistics of images. Excellent overviews are given in [16, 15]. While many publications were concerned with the statistics on the pixel level and the derivation of filters from natural images by coding principles (see, e.g. [12, 1]), recently statistical investigation for local edge

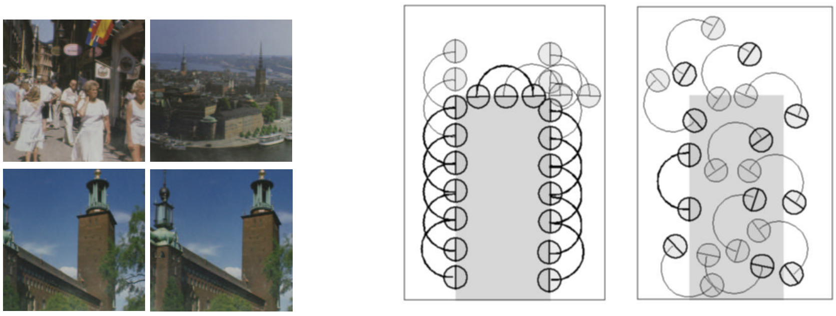

Figure 2: **Left:** Images of the data set (top) and 2 images of a sequence (bottom). **Right:** Explanation of the Gestalt coefficient $G(e^1|e^2)$: We define e^2 as the occurrence of a line segment with a certain orientation (anywhere in the image). Let the second order event e^1 be: "occurrence of collinear line segments two units away from an existing line segment e^2". Left diagram: Computation of $P(e^1|e^2)$. All possible occurrences of events e^1 in the image are shown. Bold arcs represent real occurrences of the specific second order relations e^1 whereas arcs in general represent possible occurrences of e^1. In this image we have 17 possible occurrences of collinear line segments two units away from an existing line segment e^2 and 11 real occurrences. Therefore we have $P(e^1|e^2) = 11/17 = 0.64$. Right diagram: Approximation of the probability $P(e^1)$ by a Monte Carlo method. Entities e^2 (bold) are placed randomly in the image and the presence of the event 'occurrence of collinear line segments two units apart of e^2' is evaluated. (In our simulations we used more than a 500000 samples for the estimation of $P(e^1)$). Only in 1 of 11 possible cases this event takes place (bold arc). Therefore we have $P(e^1) = 1/11 = 0.09$ and the Gestalt coefficient for the second order relation is $G(e^1|e^2) = 0.64/0.09 = 7.1$.

structures have been performed (see, e.g., [9, 3, 6]) and have addressed the representation of Gestalt principles.

Here we go one step further by investigating the second order relations not only in the modality orientation but in our multi–modal feature space

$$e = ((x_1, x_2), o, p, ((c^l_r, c^l_g, c^l_b), (c^r_r, c^r_g, c^r_b)), (f_1, f_2)).$$

In our simulations we collect second order events in bins defined by small patches in the (x_1, x_2)–space and by regions in the modality–spaces defined by the metrics defined for each modality (for details see [10]). Figure 3 shows the Gestalt coefficient for equidistantly separated bins (one bin corresponds to a square of 10×10 pixels and an angle of $\frac{\pi}{8}$ rad). As already been shown in [9, 6] collinearity can be detected as significant second order relation as a ridge in the surface plot for $\Delta o = 0$ in figure 3e. Also parallelism is detectable as an offset of this surface. A Gestalt coefficient significantly above one can also be detected for small orientation differences (figure 3d,f, i.e., $\Delta o = -\frac{\pi}{8}$ and $\Delta o = \frac{\pi}{8}$).

The general shape of surfaces is similar in all following measurements concerned with additional modalities: *we find a ridge corresponding to collinearity and an offset corresponding to parallelism and a Gestalt coefficient close to one for all larger orientation differences.* Therefore, in the following we will only look at the surface plots for equal orientation $\Delta o = 0$. These result shows *that Gestalt laws are reflected in the statistics of natural images: Collinearity and*

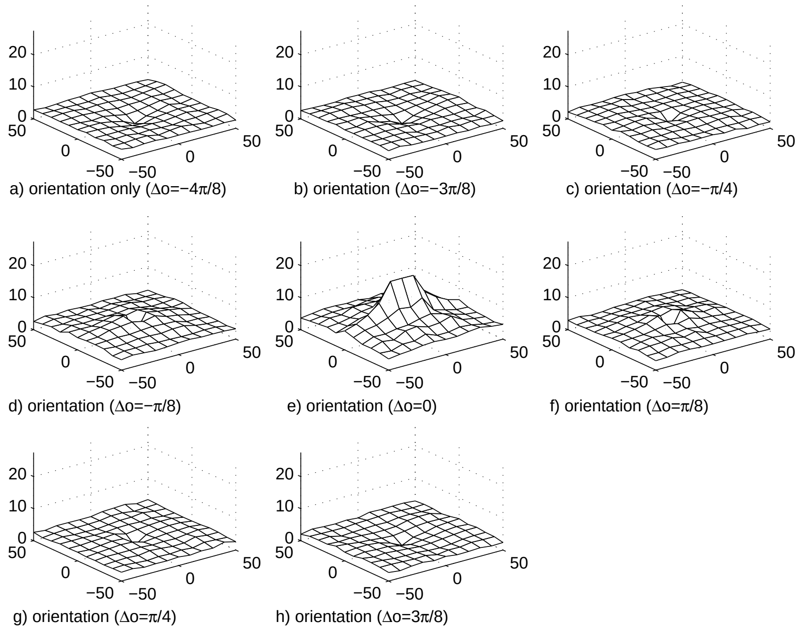

Figure 3: The Gestalt coefficient for differences in position from -50 to 50 pixel in x– and y– direction when orientation only is regarded. Note that the Gestalt coefficient for position (0,0) and $\Delta o = 0$ is set to the maximum of the surface for better display. The Gestalt coefficient is not interesting at this position, since e^1 and e^2 are identical

parallelism correspond to significant second order events of visual low level filters (see also [9]).

2.3 Pronounced Interdependencies by using additional Modalities

Now we can look at the Gestalt coefficient when we also take into account the modalities contrast transition, optic flow and color.
One additional modality: Figure 4b shows the Gestalt coefficient for the events 'similar orientation and similar contrast transition' (the metrics for the different modalities are defined precisely in [10]). In figure 5 the Gestalt coefficient along the x-axes in the surface plot of figure 4 is shown. The Gestalt coefficient on the x-axes correspond to the 'collinearity' ridge. The first column represents the Gestalt coefficient when we look at similar orientation only, while the second columns represent the Gestalt coefficient when we look at similar orientation and similar phase. *We see a significant increase of the Gestalt coefficient compared to the case when we look at orientation only corresponding to the Gestalt law collinearity.* Analogously, we define that two events have 'similar color structure' or 'similar optic flow'. The corresponding surface plot is shown in figure 4c and 4d. The slice corresponding to the collinearity ridge is shown in the third and fourth column in figure 5. An even more pronounced

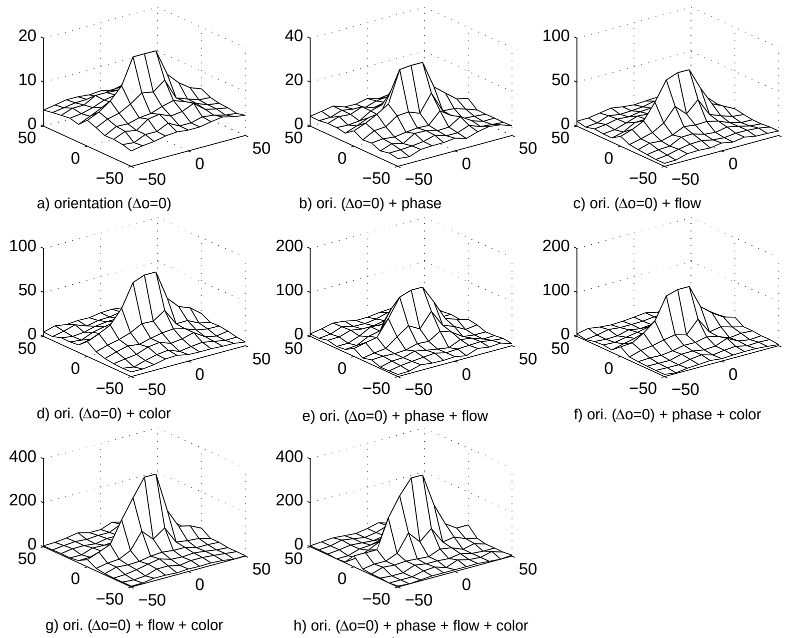

Figure 4: The Gestalt coefficient for $\Delta o = 0$ and all possible combination of modalities.

increase of inferential power for collinearity can be detected.
Multiple additional Modalities: Figure 4 shows the surface for similar orientation, phase and optic flow (figure 4e); similar orientation, phase and color (figure 4f) and similar orientation, optic flow and color (figure 4g). The slices corresponding to collinearity are shown in the fifth to seventh columns in figure 5. We can see that the the Gestalt coefficient for collinear line segments again increases significantly. Most distinctly for the combination optic flow and color (seventh column). Finally we can look at the Gestalt coefficient when we take all three modalities into account. Figure 4h and the eighth column in figure 5 shows the results. Again an increase of the Gestalt coefficient compared to the case when we look at only two additional modalities can be achieved.

Conclusion: In this paper we have addressed the statistics of local oriented line segments derived from natural scenes by adding information to the line segment concerning the modalities contrast transition, color, and optic flow. We could show that statistical interdependencies in the orientation–position domain correspond to the Gestalt laws collinearity and parallelism and that they become significantly stronger when multiple modalities are taken into account.

The results presented here provide further evidence for the assumption that despite the vagueness of low level processes stability can be achieved by *integration of information across modalities.* In addition, the attempt to model the application of Gestalt laws based on statistical measurements, as suggested recently by some researchers (see, [6, 3, 9]) gets further support. Most importantly, the results derived in this paper suggest to formulate the application of Gestalt principles in a multi-modal way.

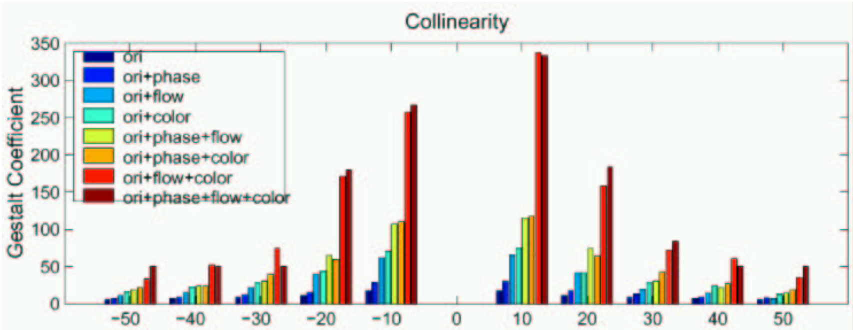

Figure 5: The Gestalt coefficient for collinear feature vectors for all combinations of modalities. The x-axis represents the distance of the collinear line segments in pixel and corresponds to the collinearity ridge in figure 3 and 4. For (0,0) the Gestalt coefficient is not shown, since e^1 and e^2 would be identical.

References

[1] A.J. Bell and T. Sejnowski. Edges are the 'independent components'of natural scenes. *Advances in Neural Information Processing Systems*, 9, 1996.

[2] R.C.K. Chung and R. Nevatia. Use of monucular groupings and occlusion analysis in a hierarchical stereo system. *CVPR*, 1991.

[3] H. Elder and R.M. Goldberg. Inferential reliability of contour grouping cues in natural images. *Perception Supplement*, 27, 1998.

[4] M. Felsberg and G. Sommer. The monogenic signal. *IEEE Transactions on Signal Processing*, 41(12), 2001.

[5] M.S. Gazzaniga. *The cognitive Neuroscience*. MIT Press, 1995.

[6] W.S. Geisler, J.S. Perry, B.J. Super, and D.P. Gallogly. Edge co–occurrence in natural images predicts contour grouping performance. *Vision Research*, 41:711–724, 2001.

[7] G. Guy and G. Medioni. Inferring global perceptual contours from local features. *International Journal of Computer Vision*, 20:113–133, 1996.

[8] F. Heitger, R. von der Heydt, E. Peterhans, L. Rosenthaler, and O. Kübler. Simulation of neural contour meachnisms: representing anomalous contours. *Image and Vision Computing*, 16:407–421, 1998.

[9] N. Krüger. Collinearity and parallelism are statistically significant second order relations of complex cell responses. *Neural Processing Letters*, 8(2):117–129, 1998.

[10] N. Krüger and F. Wörgötter. Multi modal estimation of collinearity and parallelism in natural image sequences. *to appear in Network: Computation in Neural Systems*, 2002.

[11] H.-H. Nagel. On the estimation of optic flow: Relations between different approaches and some new results. *Artificial Intelligence*, 33:299–324, 1987.

[12] B.A. Olshausen and D. Field. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature*, 381:607–609, 1996.

[13] S. Posch. *Perzeptives Gruppieren und Bildanalyse*. Habilitationsschrift, Universität Bielefeld, Deutscher Universitäts Verlag, 1997.

[14] S. Sarkar and K.L. Boyer. *Computing Perceptual Organization in Computer Vision*. World Scientific, 1994.

[15] E.P. Simoncelli and B.A. Ohlshausen. Natural image statistics and neural representations. *Anual Reviews of Neuroscience*, 24:1193–1216, 2001.

[16] C. Zetsche and G. Krieger. Nonlinear mechanisms and higher–order statistics in biologial vision and electronic image processing: review and perspectives. *Journal of electronic imaging*, 10(1), 2001.