Mathematical and perceptual analysis of multiple motions

Michael Dorr, Ingo Stuke,
Cicero Mota, and Erhardt Barth
Institute for Signal Processing
University of Lübeck
Ratzeburger Allee 160, 23538 Lübeck
{dorr,stuke,mota,barth}@isip.mu-luebeck.de
www.isip.mu-luebeck.de

C.M. is affiliated with the University of Amazonas, Brazil and is supported by DAAD under A/99/22641.

Abstract

Motion selectivity is a key feature of visual processing. However, most models of motion perception and motion-selective neurons cannot account for multiple motions. This is partly because the theoretical problems related to multiple motions have not been solved, although useful concepts like 'nulling filters' and 'layers' have been introduced . Our objective is to obtain analytical solutions for multiple motions and to use them to investigate the perception of transparent and occluded motions.

Our theoretical results are an extension of methods based on the minors of the 3x3 structure tensor $\mathbf{J}$. For *n* motions we obtain a *MxM* generalized structure tensor $\mathbf{J}_n$ with *M=6* for two and *M=10* for three motions. With this framework, we can investigate whether the theoretical and computational problems with multiple motions correspond to difficulties in perceiving multiple motions. In the case of one motion, the well-known aperture problem is reflected by the rank of $\mathbf{J}$ being one. In analogy, higher order 'aperture problems' appear in case of multiple motions when the rank of $\mathbf{J}_n$ is *M-2*. We have designed experiments in which subjects viewed between two and four overlaid moving patterns and were asked to indicate the number of motions.

At first, our data confirm previous results according to which only two motions can be recognized reliably . However, even if difficult, our results show that the distinction between 2, 3, and 4 motions is possible. More importantly, when considering the class of three motions, configurations that correspond to a higher-order 'aperture problem' are more often classified as 'two motions'.

In conclusion, our analytical framework for the analysis of multiple motions allows for asking new questions about motion selectivity.

Introduction

Multiple motions can occur due to transparencies and reflections and human observers are able to see multiple motions.

However, Jeff Mulligan reported that "when the number of moving patterns is increased beyond two, subjects are no longer able to perceive all of the patterns simultaneously" (Mulligan, 1992).

Based on new insights in the problem of estimating multiple motions, we have started to reinvestigate some issues related to the perception of multiple motions.

The problem of estimating two motions has been first solved by resolving a six-dimensional eigensystem and a three-dimensional eigensystem to separate the motion vectors (Shizawa and Mase, 1990) . In our framework we do not need to solve such eigensystems and we can extend the method to more than two motions. The approach taken by the group at MIT (e.g., Wang and Adelson, 1994) is more general in that it treats the separation of motions into layers, but is also limited to the use of a discrete set of possible motions and a probabilistic procedure for finding the most likely motions out of the set.

We can conclude that simple analytical solutions for multiple motions have not been proposed.

Intrinsic dimensionality in 3D

$f(x,y,t)$ image-sequence intensity

0: constant in all directions $f(x,y,t) = const.$

1: constant in 2 directions $f(x,y,t) = g(\xi)$

2: constant in one direction $f(x,y,t) = g(\xi,\zeta)$

3: no constant direction $f(x,y,t) = g(\xi,\zeta,\tau)$

The structure tensor $\mathbf{J}$

- Image-sequence intensity $f(x,y,t)$
- Gradient (f_x, f_y, f_t)

$$\mathbf{D}(x,y,t) = (f_x, f_y, f_t)^T (f_x, f_y, f_t)$$

- Structure tensor

smoothing kernel

$$\mathbf{J}(x,y,t) = h(x,y) * \mathbf{D}(x,y,t)$$

$$\mathbf{M} = Minors(\mathbf{J})$$

M_{ij} are the elements of $\mathbf{M}$

11 13 33

$$\mathbf{J} = h * \begin{pmatrix} f_x^{\,2} & f_x f_y & f_x f_t \\ f_x f_y & f_y^{\,2} & f_y f_t \\ f_x f_t & f_y f_t & f_t^{\,2} \end{pmatrix}$$

e.g.

$$M_{11} = \left(h * f_x^{\,2}\right)\left(h * f_y^{\,2}\right) - \left(h * \left(f_x f_y\right)\right)^2$$

Single motion from the minors of $\mathbf{J}$

Spatial patterns that move with constant velocity $\mathbf{v} = (v_x, v_y)$

satisfy the constraint:

$$v_x f_x + v_y f_y + f_t = \alpha(\mathbf{v}) f = \mathbf{V} \nabla f = 0$$

with

$$\mathbf{V} = (v_x, v_y, 1) \qquad \nabla f = (f_x, f_y, f_t)$$

$$f(x,y,t) = f(x - v_x t, y - v_y t) = f(\mathbf{x} - \mathbf{v}t)$$

For the above pattern (rigid motion) one obtains (Barth, 2000):

$$(M_{31}, -M_{21}) / M_{11} = \mathbf{v}$$
$$(M_{23}, -M_{22}) / M_{12} = \mathbf{v}$$
$$(M_{33}, -M_{23}) / M_{13} = \mathbf{v}$$

Analytic solutions for multiple motions

For additive multiple motions

$$f(\mathbf{x},t) = f_1(\mathbf{x}-\mathbf{v}_1 t) + \ldots + f_n(\mathbf{x}-\mathbf{v}_n t)$$

we have:

$$\alpha(\mathbf{v}_1)\ldots\alpha(\mathbf{v}_n) f = 0$$

$$\mathbf{D}_n(x,y,t) = \mathbf{L}^T\mathbf{L} \qquad \mathbf{J}_n \mathbf{V}_n = 0$$

$\mathbf{L}_n$ all partial derivatives of order **n** of f

$\mathbf{V}_n$ all coefficients of the algebraic expansion of

$$\prod_{j=1}^{n} (v_{jx}dx + v_{jy}dy + dt)$$

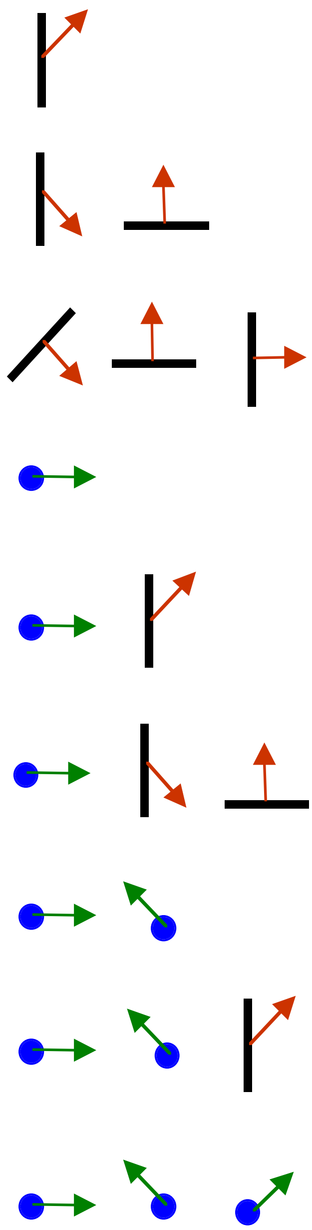
$$\mathbf{J}_n(x,y,t) = h * \mathbf{D}_n(x,y,t)$$

$$\mathbf{V}_{ni} \propto (M_{im}, -M_{im-1}, \ldots, (-1)^m M_{i1})$$

$M_{ij}, i = 1,\ldots,m$ are the minors of $\mathbf{J}_n$

Multiple-motion types

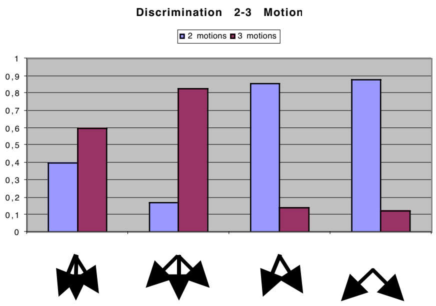
Pattern	Intrinsic dim.	Rank J_1	Rank J_2	Rank J_3
uniform	0	0	0	**0**
one 1D motion	1	1	1	**1**
two 1D motions	2	2	2	**2**
three 1D motions	3	3	3	**3**
one 2D motion	2	2	3	**4**
one 2D + one 1D m.	3	3	4	**5**
two 1D + one 2D m.	3	3	5	**6**
two 2D motions	3	3	5	**7**
two 2D + one 1D m.	3	3	6	**8**
three 2D motions	3	3	6	**9**
3D noise discont.	3	3	6	**10**

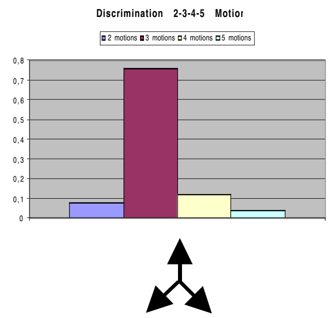


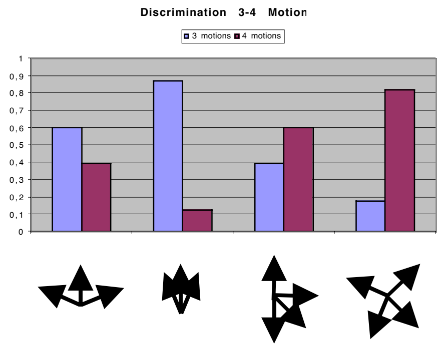
Experimental results

We created sets of randomly distributed grayscale dots that were translated with constant velocity for 50 frames. 2, 3, or 4 of these sets with different directions were then overlaid. The rate of non-zero pixels was always 10%, independent of the number of overlaid motions.

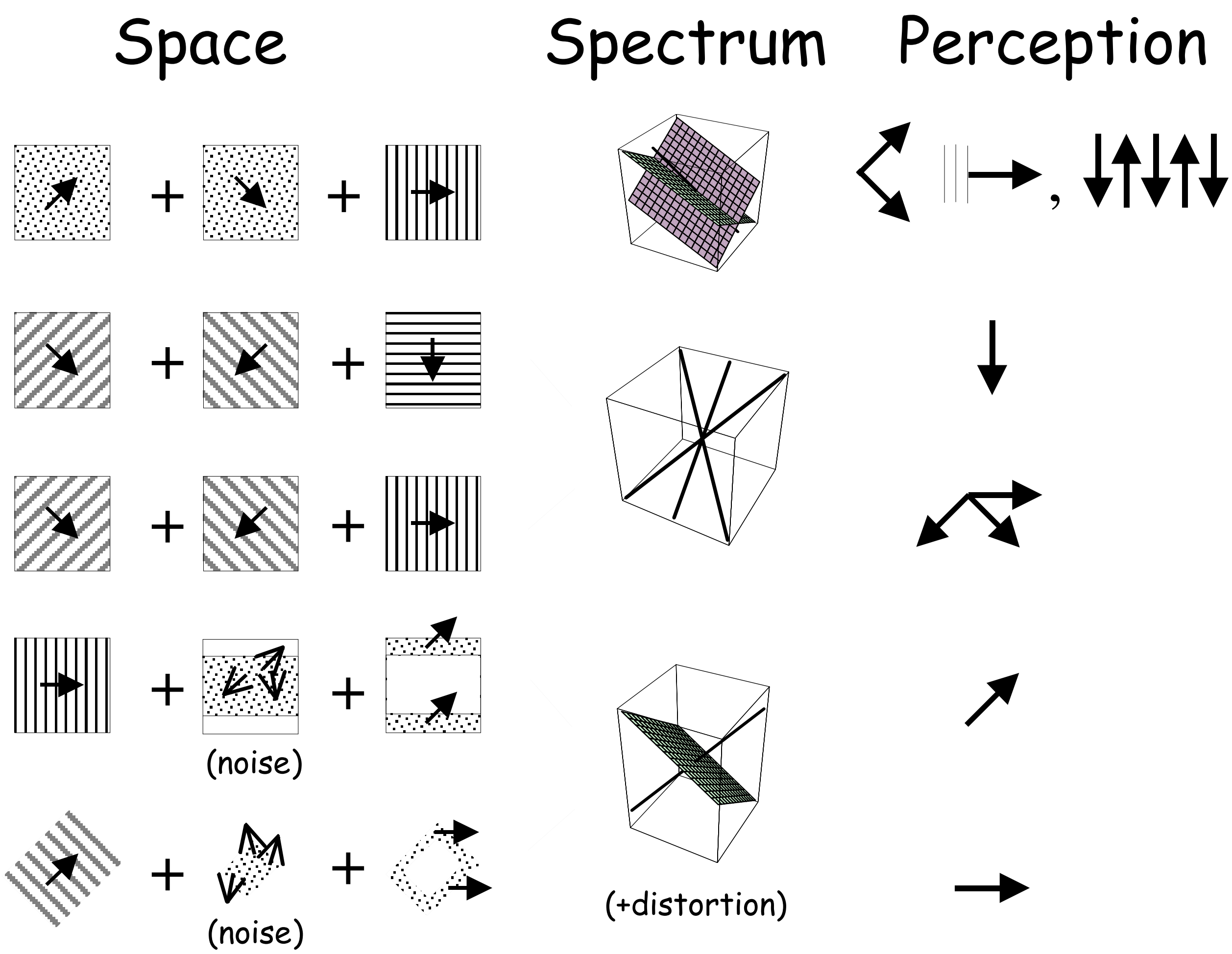
Subjects were presented the resulting movies (duration 666ms, visual angle 8deg) and asked to indicate the number of motions. Their choice was limited to 2 or 3, 3 or 4, or 2 to 5 motions in different experiments.



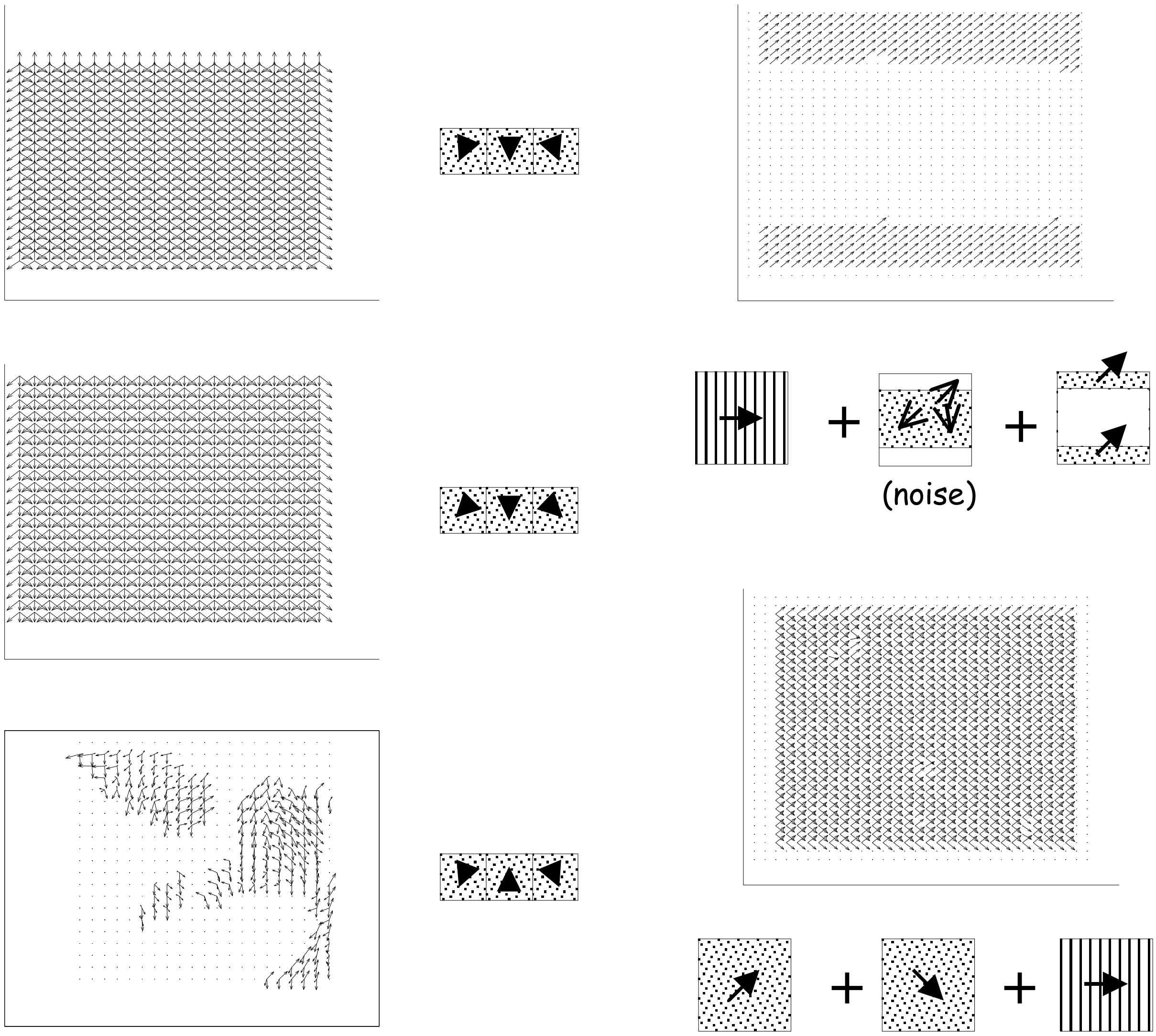




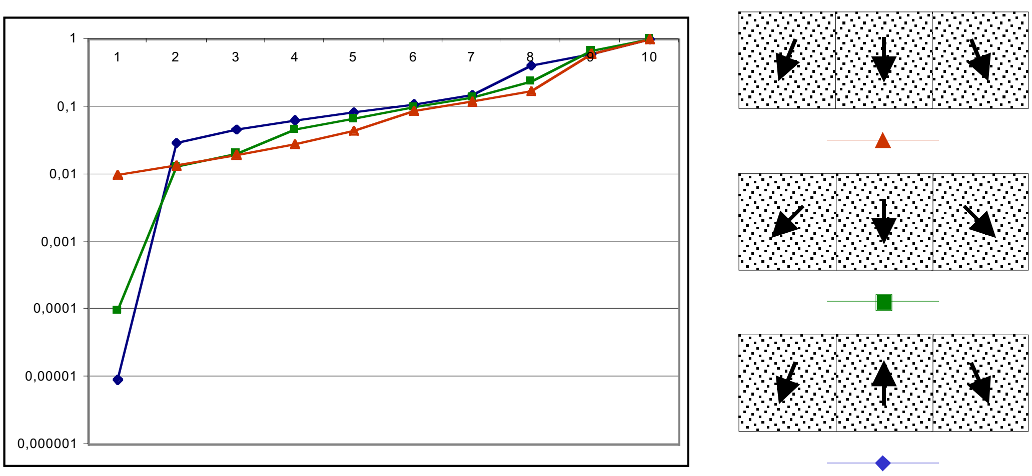
Perceptual phenomena



Simulations



Eigenvalues of J_3 for different directional separations



Conclusions

We have presented a new **theoretical framework** for dealing with multiple motions. In particular, we have shown how the rank of the generalized structure tensor $\mathbf{J}_n$ can be related to different types of multiple motions.

Our **experimental results** show that human subjects can recognize up to four overlaid motions and that the limiting factor in seeing multiple motions is rather the directional separation than the number of motions.

Demo movies have shown new examples of how humans perceive multiple motions, e.g. an illusion similar to the barber-pole illusion that here is induced by overlaid motions.

Computer simulations demonstrate that the difficulties in computing multiple motions are similar to those in seeing them and predict some of the percepts.

Bibliography

E.Barth, „The minors of the structure tensor", in *Mustererkennung 2000, LNCS*, G.Sommer, N.Krüger, and C.Perwass, eds. (Springer, 2000).

J.B.Mulligan, „Motion transparency is restricted to two planes", Inv.Ophth. Vis. Sci. **33**, 1049(1992).

M. Shizawa and K. Mase, „Simultaneous multiple optical flow estimation", in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Atlantic City, June 1990.

J.Y.A. Wang and E.H.Adelson, „Representing moving images with layers", IEEE Transactions on image processing **3**, 625-638 (1994).