

Categorization of Transparent-Motion Patterns Using the Projective Plane

Cicero Mota⁰, Michael Dorr⁰, Ingo Stuke¹, and Erhardt Barth⁰
University of Lübeck, Germany

Abstract

Based on a new framework for the description of N transparent motions we categorize different types of transparent-motion patterns. Confidence measures for the presence of all these classes of patterns are defined in terms of the ranks of the generalized structure tensor. To resolve the correspondence between the ranks of the tensors and the motion patterns, we introduce the projective plane as a new way of describing motion patterns. Transparent motions can occur in video sequences and are relevant for problems in human and computer vision. We show a few examples for how our framework can be applied to explain the perception of multiple-motion patterns and demonstrate a new illusion.

Keywords: Human and computer vision, multiple transparent motions, generalized structure tensor.

1. Introduction

Motion estimation has many applications in computer vision, e.g., video coding, image tracking, image enhancement, depth recovery, etc. Accordingly, various algorithms for motion estimation are known, see [3, 11] for reviews. However, the problem of motion estimation is always linked to the problem of motion detection and the selection of the appropriate motion model. This is because the assumptions under which the motion parameters can be estimated correctly are rarely fulfilled in real dynamic scenes. In fact, motion estimation is an ill-posed problem [6] and algorithms rely on some sort of local or global regularization of the motion field in order to produce meaningful results. The so-called aperture problem, noise, occlusions, appearing objects, and transparencies in image sequences create situations where motion estimation becomes difficult. Therefore, a correct decision on what local or global motion model to use is as important as the estimation of the motion parameters.

Motion selectivity is also a key feature of biological visual processing and has been studied by recordings of neural responses and by psychophysical experiments. Human observers are able to see and distinguish multiple transparent

motions. A special case is that of overlaid 1D motions, i.e., the case of moving straight patterns. Of particular interest is how human observers resolve the ambiguities that are inherent in this type of patterns [1, 26, 14] and how visual neurons respond to such patterns [19].

Transparent motions are additive or multiplicative superpositions of moving patterns and occur due to reflections, semi-transparencies, and partial occlusions. Different approaches for the estimation of motion vectors for the case of multiple transparent motions are known [21, 8, 9, 13, 25, 27, 18, 24, 23] and meanwhile the non-linear transparent-motions equations introduced by Shizawa and Mase [21, 22] have been solved for an arbitrary number of motions [18]. Nevertheless, the issue of confidence for multiple-motion models has, to our knowledge, only briefly been addressed in [4, 16, 17].

This paper provides a framework for the analysis of image sequences with the occurrence of transparent moving patterns, such that, for example, the motion of two overlaid 1D patterns (e.g. two gratings) can be distinguished from the motion of one 2D pattern (these patterns remain equivalent within traditional theories of only one motion). First, we establish a correspondence between moving patterns and subsets of the projective plane. This is done such that 2D moving spatial patterns correspond to points and 1D spatial patterns correspond to lines of the projective plane. This correspondence is then used to show that different motion patterns correspond to different ranks of the generalized structure tensor J_N , see Table 2.

The purpose of our paper can be understood by analogy with the case of only one motion. Obviously, in case of no image structure, no motion can be determined. In case of 1D spatial structure (e.g. straight edges) the motion is still not defined and this is either solved by not estimating motion at 1D patterns or, in most cases, by estimating only a component of the motion vector that is orthogonal to the orientation of the 1D spatial pattern. For more than one motion, we encounter many more situations that are similar to the aperture problem in the sense that not all motion parameters can be estimated. This generalized aperture problem is therefore more complex.

1.1 Single Motion Estimation Using the Structure Tensor

We first review a method for model selection based on simple confidence measures for the case of only a single motion. This will make our extensions to multiple motions more comprehensible.

⁰Institute for Neuro- and Bioinformatics, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany, {mota, dorr, barth}@inb.uni-luebeck.de

¹Institute for Signal Processing, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany, stuke@isip.uni-luebeck.de

The structure tensor method [7, 15] is a local approach for the estimation of motion vectors. This method relies on the assumption that the intensity or color of a point does not change when the point moves [12]. This assumption leads to the well known Constant Brightness Constraint Equation

$$v_x f_x + v_y f_y + f_t = 0, \quad (1)$$

where $f(\mathbf{x}, t)$ represents the image sequence and $\mathbf{v} = (v_x, v_y)^T$ the motion field. Eq. (1) does not fully constrain \mathbf{v} at a given position \mathbf{x} and therefore \mathbf{v} is estimated under the assumption of being constant in a spatio-temporal region Ω . This assumption is equivalent to the assumption that, for (\mathbf{x}, t) in Ω , the gradient of f lies in a plane whose normal is parallel to $(v_x, v_y, 1)^T$. Estimation is performed by looking for a unitary vector $\mathbf{n} = (n_x, n_y, n_t)^T$ that best represents the normal of such a plane in a least square sense, i.e., a minimal point of the functional

$$E_1(\mathbf{n}) = \int_{\Omega} [\nabla f \cdot \mathbf{n}]^2 d\Omega. \quad (2)$$

Such a vector \mathbf{n} is given, up to a scaling factor, by an eigenvector associated to the smallest, and ideally zero, eigenvalue of the so-called structure tensor:

$$\mathbf{J}_1 = \int_{\Omega} \begin{bmatrix} f_x^2 & f_x f_y & f_x f_t \\ f_x f_y & f_y^2 & f_y f_t \\ f_x f_t & f_y f_t & f_t^2 \end{bmatrix} d\Omega \quad (3)$$

For $d\Omega(\mathbf{x}, t) = \omega(\mathbf{x}, t) d\mathbf{x}dt$, \mathbf{J}_1 can be simply computed as

$$\mathbf{J}_1 = \omega * \nabla f \otimes \nabla f = \omega * (\nabla f \nabla f^T) \quad (4)$$

Note that since E_1 is homogeneous, both \mathbf{n} and $-\mathbf{n}$ are minimal points of E_1 . Actually, $\lambda \mathbf{n}$ minimizes E_1 when the arguments of E_1 are vectors with norm λ . Therefore, we can think of \mathbf{n} as homogeneous coordinates for \mathbf{v} and simply write

$$\mathbf{v} = [v_x, v_y, 1]^T = [n_x, n_y, n_t]^T \quad (5)$$

It follows that the estimation of \mathbf{n} , and therefore \mathbf{v} , is reliable only if $\text{rank } \mathbf{J}_1$ is two. Therefore the goodness of fit for the estimator can be assessed based on the eigenvalues of \mathbf{J}_1 . Note however that even for the ideal case, $\text{rank } \mathbf{J}_1 = 2$, the vector \mathbf{n} does represent motion only if $n_t \neq 0$.

An interesting property of the structure tensor is that, besides allowing for motion estimation, it encodes a local description of the image sequence $f(\mathbf{x}, t)$. Under constant motion \mathbf{v} , the sequence f can be described by

$$f(\mathbf{x}, t) = g(\mathbf{x} - t\mathbf{v}) \quad (6)$$

within Ω . Therefore, a $\text{rank } \mathbf{J}_1 = 0$ corresponds to the motion of regions with constant intensity (\circ) and any motion vector is admissible in this region; $\text{rank } \mathbf{J}_1 = 1$ corresponds to the motion of a straight pattern ($|$), in this case admissible motion vectors are constrained by a line; other moving patterns (\bullet) correspond to the $\text{rank } \mathbf{J}_1 = 2$; and non-coherent motion like noise, popping up objects, etc. correspond to $\text{rank } \mathbf{J}_1 = 3$. Table 1 summarizes these correspondences.

Moving Patterns	rank \mathbf{J}_1
\circ	0
$ $	1
\bullet	2
others	3

Table 1. Different moving patterns and the ranks of the structure tensor: (\circ) constant intensity pattern; ($|$) 1D pattern; (\bullet) 2D patterns.

2. The Generalized Structure Tensor

Our approach is based on the framework for estimating multiple motions as introduced in [18] that we will briefly summarize here. An image sequence consisting of two transparent layers is modeled as

$$f(\mathbf{x}, t) = g_1(\mathbf{x} - t\mathbf{u}) + g_2(\mathbf{x} - t\mathbf{v}), \quad (7)$$

where $\mathbf{u} = (u_x, u_y)$ and $\mathbf{v} = (v_x, v_y)$ are the velocities of the respective layers. In homogeneous coordinates, the basic constraint equation is

$$c_{xx}f_{xx} + c_{xy}f_{xy} + c_{yy}f_{yy} + c_{xt}f_{xt} + c_{yt}f_{yt} + c_{tt}f_{tt} = 0, \quad (8)$$

where $\mathbf{c} = (c_{ij})^T$ is given by

$$c_{ij} = \begin{cases} u_j v_j & \text{if } i = j \\ u_i v_j + u_j v_i & \text{otherwise.} \end{cases} \quad (9)$$

with $u_t = v_t = 1$. As in the single motion case, Eq. (8) implies that the Hessian of f lies in a hyperplane of a six-dimensional space (the space of 3×3 symmetric matrices) whose normal is the symmetric matrix \mathbf{C} with entries c_{ij} if $i = j$ and $c_{ij}/2$ if $i \neq j$. Proceeding in a way similar to the single motion case, \mathbf{c} is estimated as the eigenvector \mathbf{s} related to the smallest eigenvalue of the tensor

$$\mathbf{J}_2 = \int_{\Omega} \begin{bmatrix} f_{xx}^2 & f_{xx}f_{xy} & \cdots & f_{xx}f_{tt} \\ f_{xx}f_{xy} & f_{xy}^2 & \cdots & f_{xy}f_{tt} \\ \vdots & \vdots & \ddots & \vdots \\ f_{xx}f_{tt} & f_{xy}f_{tt} & \cdots & f_{tt}^2 \end{bmatrix} d\Omega. \quad (10)$$

or in short notation

$$\mathbf{J}_2 = \omega * d^2 f \otimes d^2 f = \omega * (d^2 f d^2 f^T), \quad (11)$$

where $d^2 f = (f_{xx}, f_{xy}, f_{yy}, f_{xt}, f_{yt}, f_{tt})^T$. Therefore a reliable estimation of \mathbf{c} is possible only if $\text{rank } \mathbf{J}_2 = 5$. Note however that \mathbf{s} represents the motion vectors of a transparent image sequence only if its last coordinate is different from zero. This condition is necessary but not sufficient. A sufficient condition for \mathbf{s} to represent transparent motion is given in Appendix A.

Moving Pattern	Projective Representation	rank \mathbf{J}_1	rank \mathbf{J}_2	rank \mathbf{J}_3
○	the empty set	0	0	0
	a point	1	1	1
+	2 points	2	2	2
+ +	3 points	3	3	3
•	a line	2	3	4
• +	a line + a point	3	4	5
• + +	a line + 2 points	3	5	6
• + •	2 lines	3	5	7
• + • +	2 lines + a point	3	6	8
• + • + •	3 lines	3	6	9
others	others	3	6	10

Table 2. Different motion patterns (first column) and the ranks of the generalized structure tensors for 1, 2, and 3 motions (table rows). This table summarizes our results by showing the correspondence between the different motion patterns and the tensor ranks that can, in turn, be used to estimate the confidence for a particular pattern, i.e., a proper motion model. Note that the rank of \mathbf{J}_N induces a natural order of complexity for patterns consisting of N additive layers.

The approach described above for two motions can be extended to estimate the motion fields of an additive superposition $f(\mathbf{x}, t)$ of N transparent image layers g_1, \dots, g_N moving with constant but different velocities $\mathbf{v}_1, \dots, \mathbf{v}_N$.

It is known [18] that f and the velocities are constrained by

$$\sum_{j=1}^M c_{I_j} f_{I_j} = 0 \quad (12)$$

where f_{I_j} , $j = 1, \dots, M = \frac{1}{2}(N+1)(N+2)$ are the independent N th-order partial derivatives of the image sequence f , i.e., $I_j = (i_{j_1}, \dots, i_{j_N})$ is an ordered sequence with components in $\{x, y, t\}$ and f_{I_j} is the N th-order partial derivative of f with respect to the components of I_j . The mixed motion parameters c_I are the symmetric function of the coordinates of $\mathbf{V}_n = \mathbf{v}_n + \mathbf{e}_t$, for $n = 1, \dots, N$, and \mathbf{e}_t is the time axis.

The generalized structure tensor for N motions is defined by

$$\mathbf{J}_N = \int_{\Omega} \begin{bmatrix} f_{I_1}^2 & f_{I_1} f_{I_2} & \cdots & f_{I_1} f_{I_M} \\ f_{I_1} f_{I_2} & f_{I_2}^2 & \cdots & f_{I_2} f_{I_M} \\ \vdots & \vdots & \ddots & \vdots \\ f_{I_1} f_{I_M} & f_{I_2} f_{I_M} & \cdots & f_{I_M}^2 \end{bmatrix} d\Omega \quad (13)$$

and can be written in short notation as

$$\mathbf{J}_N = \omega * d^N f \otimes d^N f = \omega * (d^N f d^N f^T), \quad (14)$$

where $d^N f = (f_{I_1}, f_{I_2}, \dots, f_{I_M})^T$. In this case, the vector $\mathbf{c}_N = (c_{I_1}, c_{I_2}, \dots, c_{I_M})^T$ is a null eigenvector of \mathbf{J}_N and, in practice, estimated as the eigenvector \mathbf{s}_N associated to the smallest eigenvalue of \mathbf{J}_N . The velocities are recovered from \mathbf{s}_N by the method described in [18], which is analytical for up to four motion layers. Obviously, the mixed-motion parameters \mathbf{c}_N can be computed only if the null eigenvalue is non-degenerated. In what follows, we will show which transparent moving patterns correspond to other values of the rank of \mathbf{J}_N . As mentioned for two motions, the zero eigenvalue is not a sufficient condition for \mathbf{s}_N to actually represent transparent motions. Cases where this does not happen will be ignored in the following but discussed in the Appendix B.

In analogy to single motions, we will now analyze *generalized aperture problems* as defined by the degree of degeneracy of the eigenvalues of \mathbf{J}_N and reflected in the ranks of \mathbf{J}_N , see Table 2.

The problem of motion estimation has often been studied in the Fourier domain and it is known that additive transparent moving patterns correspond to the additive superposition of Dirac planes through the origin. In the Fourier domain, Eq. (8, 12) correspond to homogeneous polynomials. The study of homogeneous equations is greatly simplified by the use of the projective plane. Therefore, we introduce a projective transform of f below.

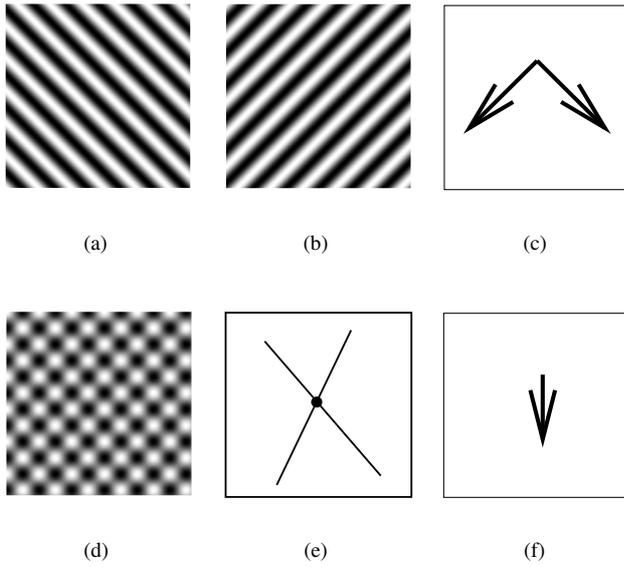


Figure 1. If two gratings of different orientations - as shown in (a) and (b) - are moved in the directions shown in (c), the plaid pattern shown in (d) is seen as moving in the direction indicated in (f) which corresponds to the only coherent velocity that is defined by the intersection of the projective lines as shown in (e).

3. Representation of Multiple Motions in the Projective Plane

Let F be the Fourier transform of f and $\xi = (\xi_x, \xi_y, \xi_t)^T$ the Fourier variable, we define a *projective transform* of f by

$$\mathcal{P}_f[\xi_x, \xi_y, \xi_t] = \frac{1}{\rho} \int_{-\infty}^{+\infty} |F(s\xi_x, s\xi_y, s\xi_t)| ds, \quad (15)$$

where $\rho = (\xi_x^2 + \xi_y^2 + \xi_t^2)^{\frac{1}{2}}$. Note that the right-hand side of the above equation does not depend on the length of ξ . Since planes and lines of an Euclidean space correspond to lines and points of the projective plane (see Appendix B), this transform allows us to think of the motion layers of $f(x, t)$ as points and lines of the projective plane. Besides the reduction of dimension, the projective plane establishes a natural duality between lines and points that is not present in Euclidean geometry. This is because a (projective) line ℓ is exactly described by an equation of the form

$$ax + by + cz = 0. \quad (16)$$

Thus any line ℓ corresponds to a *dual* point $[a, b, c]$ and vice-versa.

To illustrate the usefulness of the framework, we show how to geometrically determine the velocity of a given 2D

moving pattern: the moving pattern is mapped to a plane in the Fourier domain, from where it is further projected to the projective plane where it is a Dirac line. Finally, the velocity is found by applying the duality, here denoted with \mathcal{D} , to the Dirac line. The process is schematically shown below:

$$\text{moving 2D pattern} \xrightarrow{\mathcal{F}} \text{plane} \xrightarrow{\mathcal{P}} \text{line} \xrightarrow{\mathcal{D}} \text{velocity.}$$

In the case of a moving 1D-pattern $g(x) = \tilde{g}(\mathbf{a} \cdot \mathbf{x})$, e.g. a spatial grating, the Fourier transform reduces to a line, and its projective transform to a point. The duality operation will determine the set of admissible velocities for the grating which is a line in the projective plane:

$$\text{moving 1D pattern} \xrightarrow{\mathcal{F}} \text{line} \xrightarrow{\mathcal{P}} \text{point} \xrightarrow{\mathcal{D}} \text{line of admissible velocities.}$$

As another example, we show how to determine the *coherent motion* of superimposed gratings (plaids) [1, 19]: the set of admissible velocities for each layer is a line, the intersection of these two lines is the only admissible velocity for both layers, that is, the coherent velocity for the plaid. Further examples will be given in Section 4.

We summarize the main points below (for further details see Appendix B):

- The projective transform of transparent motions is the superposition of Dirac lines in the projective plane (in case of moving 2D patterns).
- The dual point to each Dirac line in the projective plane is the velocity of the respective layer.
- A moving 1D pattern corresponds to a Dirac point in the projective plane. In this case any admissible velocity for the grating is a point on the line that is dual to the Dirac point in the projective plane.
- Dirac lines intersect at an ideal point if and only if the corresponding patterns move in the same direction (with different speeds).
- The ideal line corresponds to a static pattern.

The projective transform and its properties establish a one-to-one correspondence between different motion patterns and subsets of the projective plane (points and lines). Furthermore, these distinct configurations in the projective plane are in a one-to-one correspondence to the rank of \mathbf{J}_N . Table 2 summarizes these correspondences and details of how these correspondences have been established are given in the Appendix A. Further benefits of the projective-plane representation of motion will become evident in the next section.

4. Applications to Some Perceptual Phenomena

For the case of only one motion, the aperture problem has a high significance for the visual perception of motion. As argued before, the motion of a 1D pattern is ambiguous from

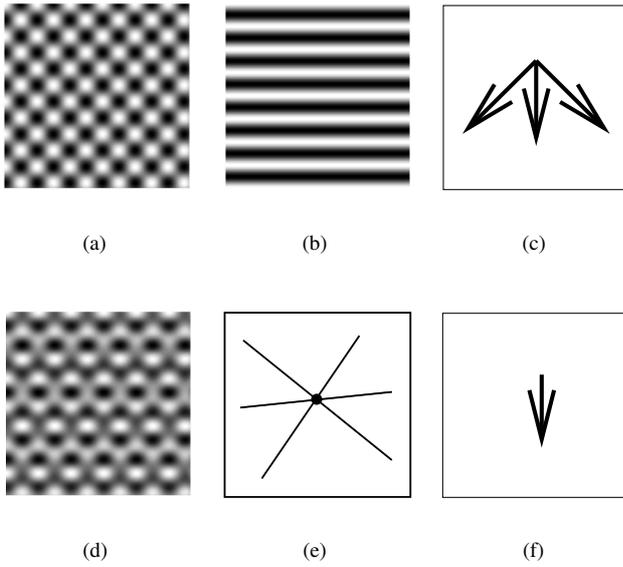


Figure 2. Coherent motion of three superimposed gratings. To the superposition of two gratings (a) a third grating shown in (b) is added. The physical motions of the three gratings are as shown in (c) and the lines of admissible velocities for each grating in (e). The percept is that of a coherent pattern as shown in (d) moving in the direction indicated by the arrow in (f). The coherent percept of one motion corresponds to the intersection of the lines in only one point.

a theoretical point of view, and so are the percepts in the sense that they depend on the motion of the so-called terminators, i.e. the ends of the 1D patterns.

Similar effects appear with superimposed gratings that can induce motion percepts that are different from the directions orthogonal to the individual gratings. For example, two gratings, one moving down and to the left, the other one moving down and to the right, are perceived as a single pattern moving downwards under most experimental conditions - see Fig. 2. On the other hand, three moving gratings can give rise to three mutually exclusive percepts [1].

We are now going to explain these phenomena using our theoretical framework presented above. We will also show that our framework predicts an illusion for the superposition of a grating with a random dot field and then give some experimental data for this illusion.

Finally, we will also give some data for the discrimination of multiple motions. We will show that it is, in principle, possible to distinguish between 2, 3, and 4 overlaid motions. It seems that the limiting factor is not the number of motions but rather the angular separation of motion vectors, which is, in turn, related to the rank of J_n . Preliminary results have

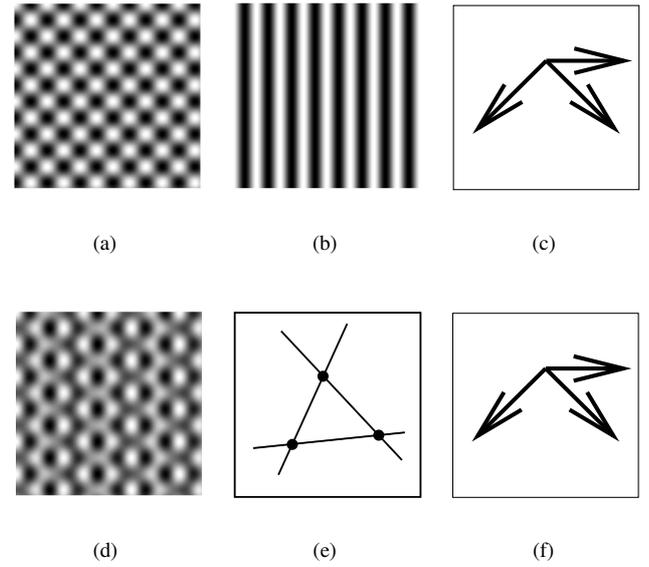


Figure 3. Incoherent motion of three superimposed gratings. The sub-figures are according to those in Fig. 2. However, the directions of motions are now changed such that the lines of motion in the projective plane do not intersect in a single point (e). This makes the motions undefined and causes the percept to change dramatically such that a coherent motion is not perceived. Observers can see either of the single motions indicated in (f) (the other two motions are seen either individually or grouped to a plaid motion).

been presented in [10].

4.1 Two 1D Transparent Moving Gratings

In the projective plane, two moving gratings correspond to the {line, line} case - see Table 2. According to the theory, the perceived motion should correspond to the intersection point U of the two lines and indeed it does - see Fig. 1.

4.2 Three 1D Transparent Moving Gratings

In the case of three moving gratings, a percept of one coherent pattern only arises when all three lines intersect in the same point. This is, for example, the case for the configuration shown in Fig. 2. On the other hand, a configuration as shown in Fig. 3 has no unique percept: human observers see the three 1D patterns as moving individually or see combinations of one 1D pattern and a 2D plaid pattern.

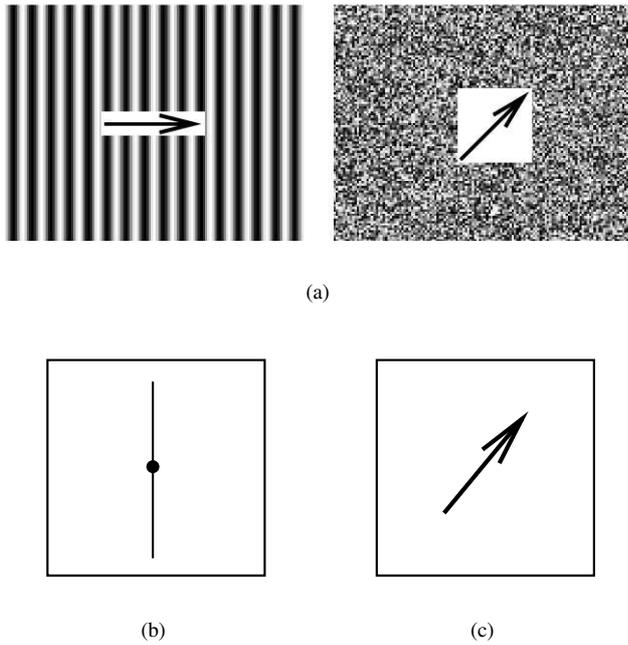


Figure 4. Stimulus generation for the 2D-over-1D entrainment (a). Admissible velocities for the grating (line) and for the 2D stimulus (point) are perceived as single motion (c).

4.3 Entrainment Effect for 2D Patterns Over 1D Patterns

A spatial field of dots superimposed on a grating (Fig. 4) corresponds to the {line, point} case. If the point falls on the line, the grating should seem to move in coherence with the random dots. To test this hypothesis, we generated sinusoidal gratings of frequency $\xi = 1/8$, orientation $\psi = k\pi/4$, $k = 1, \dots, 8$, and viewing angle size $10^\circ \times 10^\circ$. These were translated perpendicular to their orientation ($\phi_g = \psi \pm \pi/2$) with a velocity of $v_g = 1.6^\circ/s$. Mean brightness of the screen was 10 cd/m^2 . Then, a 2D dot pattern with same brightness distribution was overlaid to the grating and translated with direction $\phi_r = \phi_g \pm \pi/4$ and velocity $v_r = v_g/\sqrt{2}$, so that one component of the motion vector always coincided in the grating and the moving dot pattern. 15 of these stimuli were presented to 7 human subjects for 1.6 seconds. After presentation of each stimulus, subjects had to rotate an arrow to indicate the direction of the grating they had perceived. The deviation of subjects' responses from the true direction of the grating is given in Fig. 7(a). If the dot pattern had exerted no influence on the percept for the grating at all, a single peak at 0° could be expected. Analogously, a single peak at 45° would indicate that subjects always perceived a single coherent pattern. Note that the small peak at 135° actually depicts cases of 45° deviation since it can be attributed to the induced motion phenomenon (the same effect that makes us see the

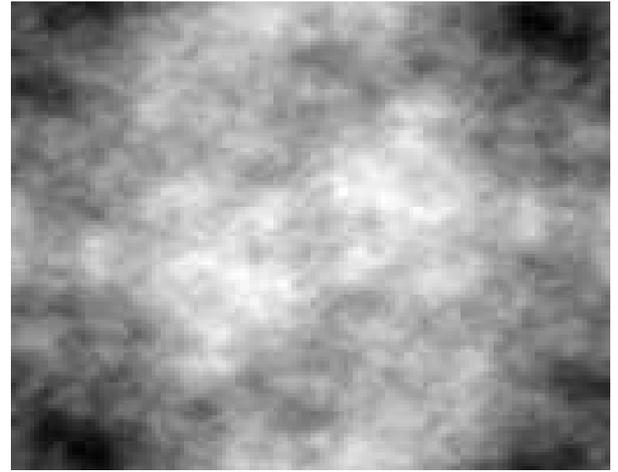


Figure 5. Example of a 1/f noise stimulus

platform moving while sitting in a moving train).

4.4 Entrainment Effect and the Barberpole Illusion

The shape of an aperture through which a grating is seen can strongly influence motion perception. This phenomenon is called the barberpole illusion. For example, the straight lines in Fig. 6 seem to change their direction along their path behind the aperture [26]: the bar moves as indicated by the arrows and the perceived motion is indicated by the dashed line.

To show that the entrainment effect is able to override the barberpole illusion, we designed the stimuli illustrated in Fig. 6. We masked the moving grating by an aperture perpendicular to the orientation of the grating. This should strengthen the percept of motion in a direction orthogonal to the grating. As an additional modification, only the terminators of the grating were overlaid with a random dot field that moved in one coherent direction. Because this led to the rise of new terminators at the boundary of the coherent random dot field, the remaining middle of the stimulus was overlaid with a white-noise pattern, which had the same density and brightness as the coherent noise pattern. Nevertheless, the entrainment effect seen in Fig. 7(b) is still qualitatively similar to that in Fig. 7(a) which shows that the effect dominates over the influence of the aperture.

4.5 Discrimination of Multiple Transparent Motions

The perception of multiple overlaid motions has also been investigated by counting the number of layers that a person can discriminate and it has been argued that it is impossible to discriminate more than two transparent motions [20]. To analyze the nature of this apparent bottleneck we performed the following experiments.

Stimuli consisted of 253 ms long, $10^\circ \times 10^\circ$ sized image sequences made of either 2, 3, or 4 translated 1/f-noise pat-

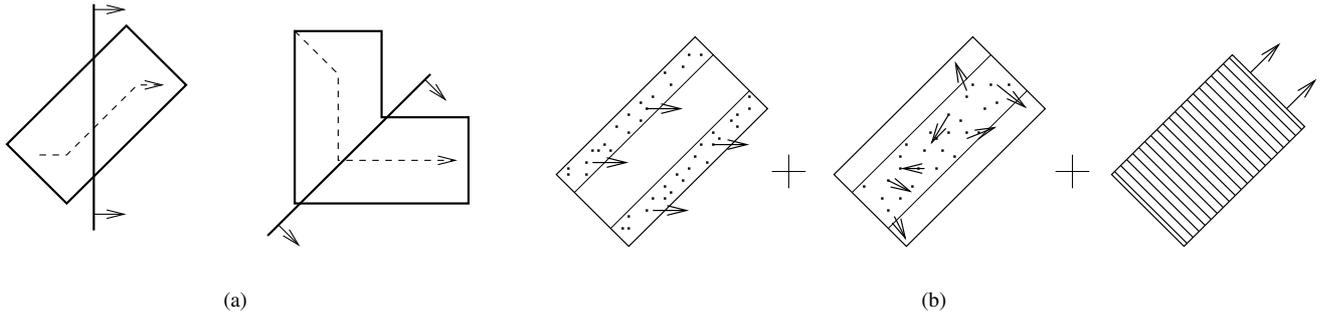


Figure 6. Barberpole illusion (a). Stimulus with aperture orientation perpendicular to that of the 1D grating, random noise in the center, and a random dot field moving coherently in the horizontal plane (b).

terns that were overlaid. $1/f$ -noise images are characterized by an hyperbolically-shaped spectrum, e.g. a high proportion of low-frequency content and few high-frequency components. This property has been chosen to resemble statistical properties of natural images [2]. For an example, see Fig. 5. Note that when several of these images are overlaid, this characteristic is preserved; therefore, one cannot detect the number of overlaid motions from still images alone. These patterns were translated with $12^\circ/s$. The directional separation of the motion vectors was $15, 30, \dots, 180^\circ$ respectively. Subjects then had to indicate, in a 5 alternatives forced choice paradigm, whether they had perceived 1, 2, 3, 4, or 5 motions. In addition to the experiments described above, we performed an additional set of experiments that differed only by the fact that we used random dot patterns instead of the $1/f$ patterns.

Results can be seen in Fig. 8 and 9. Note that overall the discriminability increases with the angular separation of the motions. Also note, however, that the difficulty of the discrimination task increases with the number of motions. Nevertheless, three motions can be well discriminated with sufficient angular separation. We therefore suggest that the angular separation is the main limiting factor, which is, of course, in turn limited by the number of motions. As shown in Fig. 10, this effect can be predicted qualitatively in terms of confidence measures based on the generalized structure tensor. The confidence measure is obtained as the inverse slope of the line fitted to the distribution of the logarithm of the $M - 1$ largest eigenvalues of the generalized structure tensor. The inverse slope values are shown normalized to the range $[0, 1]$. Similar results would be obtained for the $1/f$ patterns.

5. Discussion

We have presented a method for categorizing transparent-motion patterns in terms of the ranks of the generalized structure tensors. Based on our results, the confidence for a particular pattern can be evaluated computationally by either determining the rank J_N or by using the minors of the structure

tensors [18]. For example, we can discriminate the case of two superimposed 1D patterns (moving plaid) and a 2D pattern moving in the direction of the coherent motion of the plaid pattern.

Our results can be seen as an extension of the concept of *intrinsic dimension* [28, 5]. In the current framework, the intrinsic dimension corresponds to the rank of J_1 . As shown in Table 2, by introducing the generalized structure tensor, we can further differentiate the signal classes of a given (integer) intrinsic dimension. In some sense, we thereby define fractional intrinsic dimensions.

Although motion estimation is a key component of many computer-vision and image processing systems, the motion models are often too simple and fail with realistic data. Our results provide (i) new means for increasing the complexity of the motion models and (ii) measures for determining the confidence for a particular model. We should note that the framework can be applied to make explicit the correspondence between the ranks of J_N , for a value of N larger than 3, and the different moving patterns.

The theory presented in this work provides a conceptual understanding of the difficulties in the estimation of multiple transparent motions, which are due to the generalized aperture problem. We have used the projective transform to establish a correspondence between the rank of the generalized structure tensor and different transparent moving patterns. Note, however, that the generalized structure tensor is derived by integration of the derivatives of the image sequence in a local neighborhood and as such can be used for the estimation of transparent motion in that neighborhood, including situations where the motion vectors may vary over space. For the estimation of N transparent motions, N th-order partial derivatives are involved. The ubiquitous presence of noise can be compensated by prefiltering the sequences with proper kernels, see [18]. Such prefiltering is equivalent to the use of more general filters instead of the derivatives and can thus break the unfavorable relationship between the order of differentiation and the sensitivity to noise.

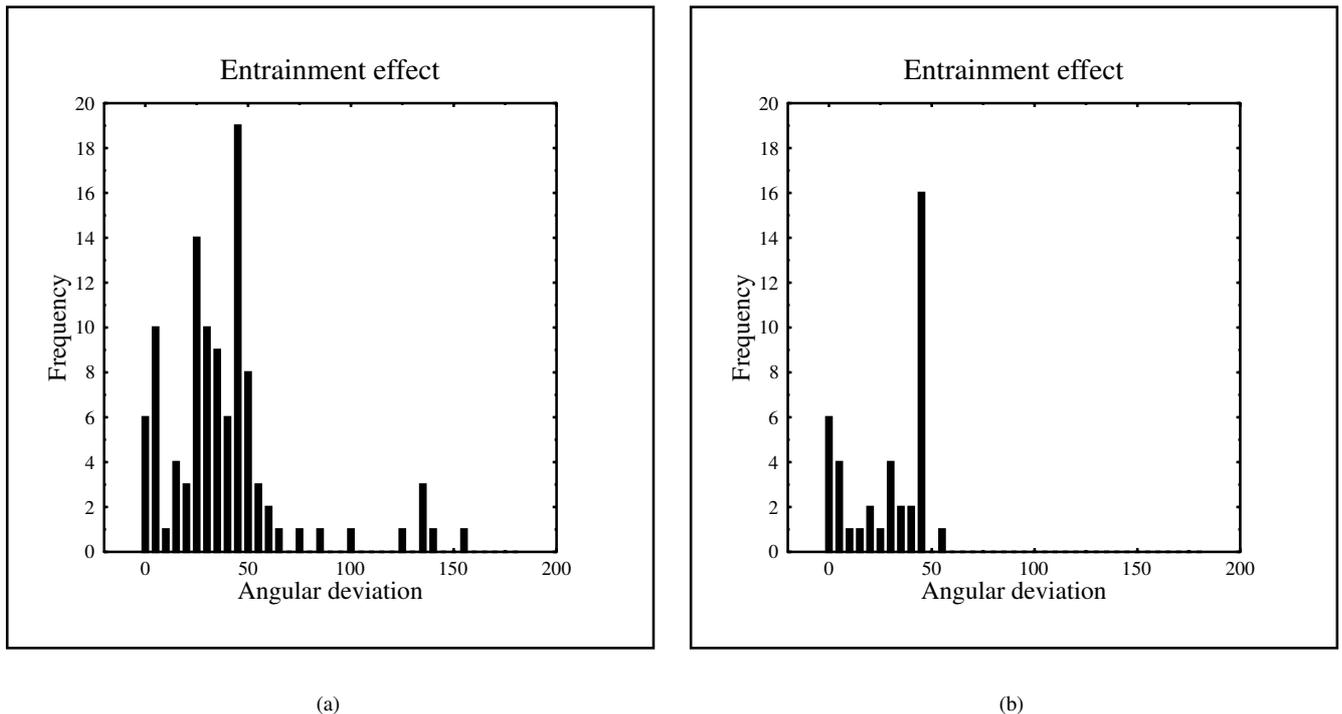


Figure 7. Data illustrating the entrainment effect of a 2D pattern over a 1D grating. No aperture (a). Aperture orientation perpendicular to that of the 1D grating (b).

Finally, we have also shown how our results can be used to explain some phenomena in biological vision. In particular, the concept of the projective plane proved useful for describing and visualizing different visual percepts. Furthermore, we demonstrated new illusionary percepts that are in accordance with the ambiguities that one would expect from the theory.

6. Acknowledgment

Work is supported by the *Deutsche Forschungsgemeinschaft* under Ba 1176/7-2.

7. References

- [1] E. H. Adelson and J. A. Movshon. Phenomenal coherence of moving visual patterns. *Nature*, 300(5892):523–5, 1982.
- [2] R. M. Balboa and N. M. Grzywacz. Power spectra and distribution of contrasts of natural images from different habitats. *Vision Research*, 43(24):2527–37, 2003.
- [3] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, 1994.
- [4] E. Barth, M. Dorr, I. Stuke, and C. Mota. Theory and some data for up to four transparent motions. *Perception*, 30 (Supplement):36, 2001.
- [5] E. Barth and A. B. Watson. A geometric framework for nonlinear visual coding. *Optics Express*, 7:155–85, 2000. <http://www.opticsexpress.org/oearchive/source/23045.htm>.
- [6] M. A. Bertero, T. Poggio, and V. Torre. Ill-posed problems in early vision. *Proceedings of IEEE*, 76(8):869–89, 1988.
- [7] J. Bigün, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with application to texture analysis and optical flow. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 13(8):775–90, 1991.
- [8] M. J. Black and P. Anandan. The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104, Jan. 1996.
- [9] T. Darrell and E. Simoncelli. Nulling filters and the separation of transparent motions. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 738–9, New York, June 14–17, 1993. IEEE Computer Press.
- [10] M. Dorr, I. Stuke, C. Mota, and E. Barth. Mathematical and perceptual analysis of multiple motions. In H. H. Bülthoff, K. R. Gegenfurtner, H. A. Mallot, and R. Ulrich, editors, *TWK 2001: Beiträge zur 4. Tübinger Wahrnehmungskonferenz*, page 173. Knirsch Verlag, Kirchentellinsfurt, Germany, 2001.

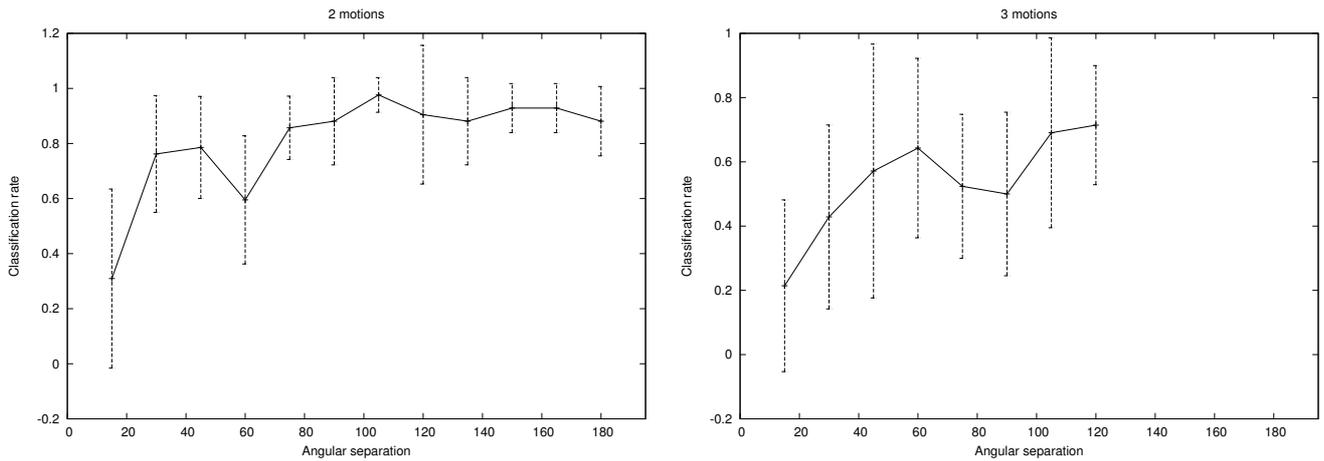


Figure 8. Discrimination of multiple motions with $1/f$ patterns.

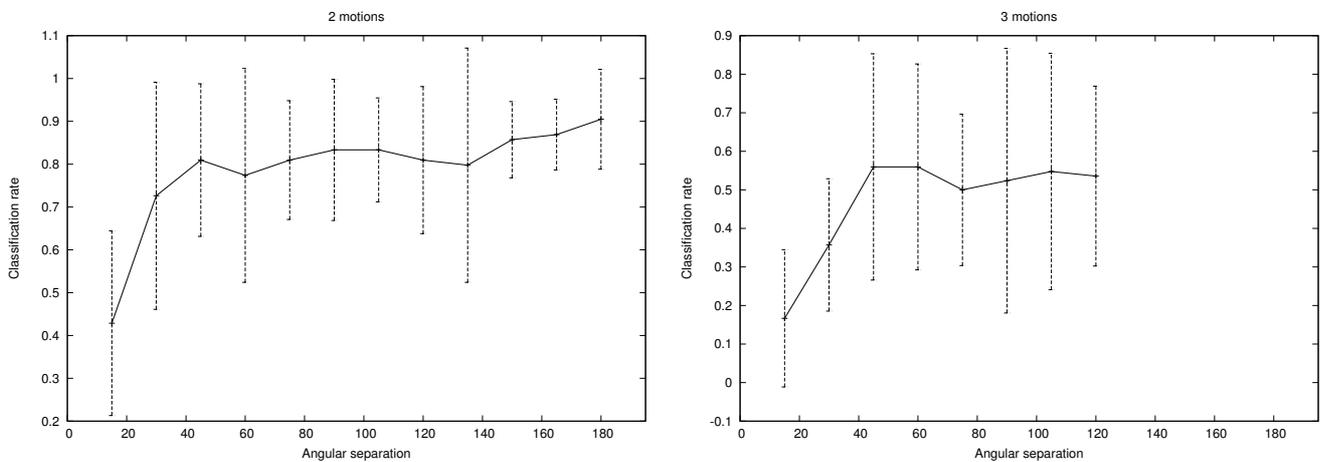


Figure 9. Discrimination of multiple motions with random-dot patterns.

- [11] H. Haußecker and H. Spies. Motion. In B. Jähne, H. Haußecker, and P. Geißler, editors, *Handbook of Computer Vision and Applications*, volume 2, pages 309–96. Academic Press, 1999.
- [12] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1–3):185–203, Aug. 1981.
- [13] M. Irani, B. Rousso, and S. Peleg. Computing occluding and transparent motions. *International Journal of Computer Vision*, 12(1):5–16, Feb. 1994.
- [14] K. Langley. Computational models of coherent and transparent plaid motion. *Vision Research*, 39:87–108, 1999.
- [15] B. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc. DARPA Image Understanding Workshop*, pages 121–30, 1981.
- [16] C. Mota, M. Dorr, I. Stuke, and E. Barth. Categorization of transparent-motion patterns using the projective plane. In W. Dosch and R. Y. Lee, editors, *Proc. ACIS 4th Int. Conf. Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*, pages 633–9, Lübeck, Germany, Oct. 16–18, 2003.
- [17] C. Mota, M. Dorr, I. Stuke, and E. Barth. Analysis and synthesis of motion patterns using the projective plane. In *Human Vision and Electronic Imaging Conference. 16th Symposium on Electronic Imaging Science and Technology*, San Jose, CA, Jan. 19–22, 2004. To appear.
- [18] C. Mota, I. Stuke, and E. Barth. Analytic solutions for multiple motions. In *Proc. IEEE Int. Conf. Image Processing*, volume II, pages 917–20, Thessaloniki, Greece, Oct. 7–10, 2001. IEEE Signal Processing Soc.
- [19] J. A. Movshon, E. H. Adelson, M. S. Gizzi, and W. T. Newsome. The analysis of moving visual patterns. In C. Chagas, R. Gattas, and C. Gross, editors, *Study group on pattern recognition mechanisms*, volume 54, pages 117–151. Pontificiae Academiae Scientiarum Scripta Varia, Vatican City, 1985.

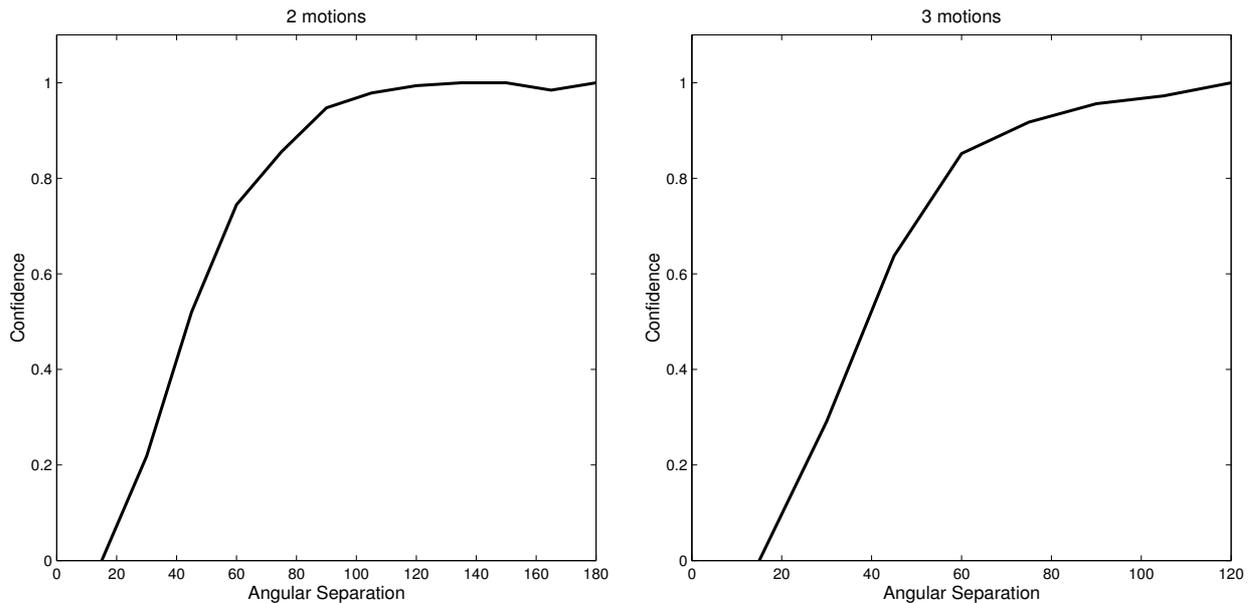


Figure 10. Simulation results obtained for the same patterns as the experimental results shown in Fig. 9.

- [20] J. B. Mulligan. Motion transparency is restricted to two planes. *Inv Ophth Vis Sci (suppl)*, 33:1049, 1992.
- [21] M. Shizawa and K. Mase. Simultaneous multiple optical flow estimation. In *IEEE Conf. Computer Vision and Pattern Recognition*, volume I, pages 274–8, Atlantic City, NJ, June 1990. IEEE Computer Press.
- [22] M. Shizawa and K. Mase. A unified computational theory for motion transparency and motion boundaries based on eigenenergy analysis. In *IEEE Conf. Computer Vision and Pattern Recognition*, pages 289–95, Maui, HI, June 1991. IEEE Computer Press.
- [23] I. Stuke, T. Aach, E. Barth, and C. Mota. Estimation of multiple motions using block-matching and markov random fields. In S. Panchanathan and B. Vasudev, editors, *Visual Communications and Image Processing 2004, IS&T/SPIE 16th Annual Symposium Electronic Imaging*, San Jose, California, Jan. 18–22, 2004. To appear.
- [24] I. Stuke, T. Aach, C. Mota, and E. Barth. Estimation of multiple motions: regularization and performance evaluation. In B. Vasudev, T. R. Hsing, A. G. Tescher, and T. Ebrahimi, editors, *Image and Video Communications and Processing 2003*, volume 5022 of *Proceedings of SPIE*, pages 75–86, May 2003.
- [25] D. Vernon. Decoupling Fourier components of dynamic image sequences: a theory of signal separation, image segmentation and optical flow estimation. In H. Burkhardt and B. Neumann, editors, *Computer Vision - ECCV'98*, volume 1407/II of *LNCS*, pages 68–85. Springer Verlag, Jan. 1998.
- [26] S. Wuerger, R. Shapley, and N. Rubin. On the visually perceived direction of motion by Hans Wallach: 60 years later. *Perception*, 25:1317–67, 1996.
- [27] W. Yu, K. Daniilidis, S. Beauchemin, and G. Sommer. Detection and characterization of multiple motion points. In *IEEE Conf. Computer Vision and Pattern Recognition*, volume I, pages 171–7, Fort Collins, CO, June 23–25, 1999. IEEE Computer Press.
- [28] C. Zetzsche and E. Barth. Fundamental limits of linear filters in the visual processing of two-dimensional signals. *Vision Research*, 30:1111–7, 1990.

A. The Null-Eigenvector of J_N and the Motion Vectors

The null-eigenvector of J_N represents N transparent-motion velocities if it is possible to solve the vectorial equation

$$c(v_1, \dots, v_N) = s_N \quad (17)$$

for v_1, \dots, v_N . In the case of two motions, this equation can be written in matrix form as

$$S_2 = \frac{s_{tt}}{2} [(u_x, u_y, 1)^T (v_x, v_y, 1) + (v_x, v_y, 1)^T (u_x, u_y, 1)], \quad (18)$$

where S is the matrix with entries s_{ij} if $i = j$ and $s_{ij}/2$ if $i \neq j$. Therefore, s_2 represents two transparent motions if and only if

$$\begin{aligned} \det S_2 &= 0 \\ \det S_2^{11} + \det S_2^{22} + \det S_2^{22} &< 0, \\ s_{tt} &\neq 0 \end{aligned} \quad (19)$$

where $\det S^{jj}$ are the diagonal minors of S , e.g., $\det S^{33} = s_{xx}s_{yy} - s_{xy}^2/4$. Eq. (18) can be easily extended to more than two motions but we could not find any expressions for the resulting S_N analogous to those in Eq. (19). Nevertheless, it has been shown in [18] that the solutions of Eq. (17) can be expressed as the roots of a complex polynomial whose coefficients are explicitly given in terms of s_n . If v_1^c, \dots, v_N^c are the roots of this polynomial, a necessary and sufficient condition for s_N to represent transparent-motion vectors is therefore

$$s_N = c(v_1^c, \dots, v_N^c). \quad (20)$$

B. The Projective Plane

The projective plane is the set of all directions in the three-dimensional Euclidean space. These directions (the points of the projective plane) can be represented by homogeneous coordinates $[x, y, z]$. A pair of homogeneous coordinates $[x, y, z], [x', y', z']$ represents the same point if and only if $(x, y, z) = \lambda(x', y', z')$ for some non-zero factor λ . A point with coordinates $[x, y, 0]$ is called an *ideal point* and the set of ideal points is called the *ideal line*.

A point (x, y) in a Euclidean plane corresponds naturally to a projective point by the identification $(x, y) = [x, y, 1]$. Therefore, we can think of the projective plane as the union of the plane $z = 1$ and the ideal line.

Relevant Properties of the Projective Plane

Below we summarize the properties of the projective plane that are useful for our analysis of moving patterns:

- Dimension reduction: lines and points of the projective plane correspond to planes and lines through the origin of the three-dimensional space respectively;
- Duality: each line ℓ of the projective plane is associated to a dual point V by the corresponding orthogonality of planes and lines in the three-dimensional Euclidean space and vice-versa;
- No parallelism: any two lines of the projective plane do intersect;
- Two projective lines intersect at an ideal point if and only if their dual points and e_t are aligned.

C. The Rank of J_N

From the discussion in Section 3, we have seen that the set of admissible velocities of a moving layer g is the dual space to the support of \mathcal{P}_G . This dual set is called the *phase space* for the velocities of g . In what follows, we will suppose that no pair of layers forming f moves with collinear velocities and none of the layers is static. This means that the lines supporting two non-degenerated Dirac lines always intercept at a finite (non-ideal) point.

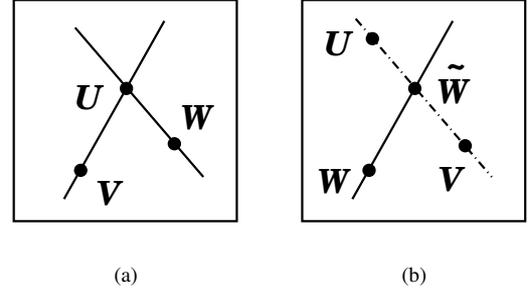


Figure 11. Admissible velocities of overlaid-motions patterns in the projective plane: (a) two overlaid 1D patterns, U is the coherent velocity, $c(u, u), c(u, v), c(u, w), c(v, w)$ are independent null-eigenvectors of J_2 ; (b) same for one 1D pattern and two 2D patterns, $c(u, v, w)$ and $c(u, v, \tilde{w})$ are independent null-eigenvectors of J_3 .

The mixed-motion parameters vectors $c_N = c(v_1, \dots, v_N)$ can be interpreted as elements of the space of symmetric N -tensors (here denoted by \mathcal{S}_N). Therefore, if $\beta = \{U, V, W\}$ is a basis for the three-dimensional Euclidean space, the set $\{c(v_1, \dots, v_N) : V_n \in \beta, \text{ for } n = 1, \dots, N\}$ is a basis for \mathcal{S}_N . For example, $\{c(u, u), c(u, v), c(u, w), c(v, v), c(v, w), c(w, w)\}$ is a basis for \mathcal{S}_2 . We will use this relationship between basis of \mathbb{R}^3 and \mathcal{S}_N to construct a maximal number of elements in the kernel of J_2 and J_3 . By ‘kernel of J_N ’ we denote the set of vectors that correspond to the zero eigenvalues of J_N .

The Rank of J_2

For two moving layers, the non-trivial possibilities for the phase space of the velocities are a $\{\text{line, line}\}, \{\text{point, line}\}, \{\text{point, point}\}$.

line, line: Choose a basis $\beta = \{U, V, W\}$ of \mathbb{R}^3 such that U is the intersection of the two lines, and V and W belong to each of these lines, see Fig. 11(a). Now it is clear that $c(u, u), c(u, v), c(u, w)$ and $c(v, w)$ are elements in the kernel of J_2 . Since these vectors are linearly independent, we can conclude that $\text{rank}(J_2) \leq 2$.

line, point: Choose U as the point and V, W in the line. The vectors $c(u, v), c(u, w)$ are null-eigenvectors of J_2 and therefore $\text{rank}(J_2) \leq 4$.

point, point: Choose U, V as the two points and W freely. The only element in the kernel of J_2 is $c(u, v)$, therefore $\text{rank}(J_2) \leq 5$.

We found the above bounds to the $\text{rank}(J_2)$ given two moving patterns. Since it is possible to reach these bounds,

they are actually tight. Note that two moving patterns do not produce rank 1 or 3. These ranks are actually produced by a single moving object. The phase space for the two velocities, in this case, is {line, plane} or {point, plane}. We analyze the first case below, the other is similar.

line, plane: Choose U, V as points in the line and W out of it. The only element that does not belong to the kernel of J_2 is $c(w, w)$ and therefore $\text{rank}(J_2) = 1$.

The Rank of J_3

For three moving patterns, the non-trivial possibilities for the phase spaces of the velocities are a {line, line, line}, {point, line, line}, {point, point, line}, and {point, point, point} which correspond to the values 3, 6, 8, and 9 of the rank of J_3 . Since the analyses of these cases are very similar, we consider only the two last cases.

point, point, line: Choose U, V as the points and W in the line, see Fig. 11(b). In principle it appears that only the element $c(u, v, w)$ belongs to the kernel of J_3 . Also note that any two lines intersect in the projective plane. Let \tilde{W} be the intersection of the given line with the line determined by U and V . Now, if we assure that W does not coincide with \tilde{W} , we find the second independent symmetric tensor in the kernel of J_3 , that is, $c(u, v, \tilde{w})$. We conclude that $\text{rank}(J_3) \leq 8$. Since these are all the possibilities, except maybe for degenerate cases, the bound 8 is tight.

point, point, point: Choose U, V, W as these points. Only $c(u, v, w)$ belongs to the kernel of J_3 . Hence, $\text{rank}(J_3) = 9$ except for degenerate cases.

Similar to the case J_2 , three moving patterns do not fill all the possibilities for the rank of J_3 . The gaps are filled by single or two moving patterns. These correspond to ranks 1,4 and 2,5,7 respectively. Table 2 summarizes the possibilities for the ranks of J_N for $N = 1, 2, 3$.



Cicero Mota is on leave from University of Amazonas, Brazil and is currently with the Institute for Neuro- and Bioinformatics, University of Luebeck, Germany. His research interests are in computer vision and information theory. Dr. Mota studied mathematics at the University of Amazonas and at the Institute for Pure and Applied Mathematics – IMPA (Ph.D. 1999) in Brazil. He has

been with the Institute for Signal Processing, University of Luebeck (visiting scientist) and with the Institute for Applied

Physics, University of Frankfurt, both in Germany. His research has been supported by the Brazilian agencies CNPq and CAPES, and by the German agencies DAAD and DFG.



Michael Dorr is currently working towards his degree in Computer Science, Bioinformatics, and Biomathematics at the University of Luebeck, Germany. His research interests are the psychophysics of motion estimation and the modeling of eye movements and attention. In 2001, he has been awarded the "Student Poster Prize" at the Tuebingen Perception Conference for [10].



Ingo Stuke is currently a scientist at the Institute for Signal Processing, University of Luebeck, Germany. His main research interests are on the field of motion and orientation estimation for image processing and computer vision applications. From 1995 to 2001 he studied computer science at the University of Luebeck.



Erhardt Barth is currently a scientist at the Institute for Neuro- and Bioinformatics, University of Luebeck, Germany. His main interests are in the areas of human and computer vision. He studied electrical and communications engineering at the University of Brasov and at the Technical University of Munich (Ph.D. 1994). He has been with the Universities of Melbourne and Mu-

nich, the Institute for Advanced Study in Berlin, the NASA Vision Science and Technology Group in California, and the Institute for Signal Processing in Luebeck. In May 2000 he received a Schloessmann Award.