Categorization of Transparent-Motion Patterns Using the Projective Plane

Cicero Mota[1,2], Michael Dorr[1], Ingo Stuke[2], and Erhardt Barth[1]

[1]Institute for Neuro- and Bioinformatics
[2]Institute for Signal Processing
University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany
{mota, dorr, barth}@inb.uni-luebeck.de, stuke@isip.uni-luebeck.de

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 describe the perception of multiple-motion patterns and predict a new illusion.

1. Introduction

Many problems in computer vision rely on motion estimation, but standard motion models fail in case of transparent motions. Transparent motions are additive superpositions of moving patterns and occur due to reflections, semi-transparencies, and occlusions. Different approaches for the estimation of motion vectors for the case of multiple transparent motions are known [7, 3, 4, 8]. The non-linear transparent-motions equations introduced by Shizawa and Mase [7] have been solved for an arbitrary number of motions [5]. However, the problem of motion estimation is always linked to the problem of motion detection. This is because the assumptions under which the motion parameters can be estimated correctly are rarely fulfilled in real dynamic scenes. Therefore, a correct decision on what local or global motion model to use is often more important and difficult to obtain than the estimation of the motion parameters. As we shall see, the strength of our approach lies not only in providing new solutions for the motion parameters, but also confidence measures for different classes of motion patterns.

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. This is known as the aperture problem and 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 and has, to our knowledge, not been addressed before.

Motion selectivity is 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 these type of patterns [1] and how visual neurons respond to such patterns [6]. This paper will provide a framework for the analysis of these motion 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. To accomplish this, first we establish a correspondence between moving patterns and subsets of the projective plane. This is done such that 2D moving patterns correspond to points and 1D moving structures correspond to lines of the projective plane. This correspondence is then used to show that different moving patterns correspond to different ranks of the generalized structure tensor $\boldsymbol{J}_N$, see Table 1.

2. The generalized structure tensor

Our approach is based on the framework for estimating multiple motions as introduced in [5] that we will briefly

	1D	1D+1D	1D+1D+1D	2D	2D+1D	2D+1D+1D	2D+2D	2D+2D+1D	2D+2D+2D
J_1	1	2	3	2	3	3	3	3	3
J_2	1	2	3	3	4	5	5	6	6
J_3	1	2	3	4	5	6	7	8	9

Table 1. Different motion patterns (table columns) and the ranks of the generalized structure tensors for 1, 2, and 3 motions (table rows). The intrinsic dimension is equal to the rank of J_1. 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. Observe that the rank of J_N induces a natural order of complexity for patterns consisting of N additive layers.

summarize here. Suppose that an image sequence f is the overlaid superposition of N image layers $g_1, \ldots, g_N$ moving with constant but different velocities $v_1, \ldots, v_N$ respectively:

$$f(\boldsymbol{x}, t) = \sum_{n=1}^{N} g_n(\boldsymbol{x} - t\boldsymbol{v}_n) \tag{1}$$

In such ideal case, it is known [5] that f and the velocities are constrained by

$$\sum_I f_I c_I = 0 \tag{2}$$

where $I = (i_1, \ldots, i_N)$ is an ordered sequence with components in $\{x, y, t\}$, f_I represents the Nth-order partial derivative of f with respect to the components of I, the mixed motion parameters c_I are the symmetric function of the coordinates of $\boldsymbol{V}_n = \boldsymbol{v}_n + \boldsymbol{e}_t$, for $n = 1, \ldots, N$, and $\boldsymbol{e}_t$ is the time axis.

The generalized structure tensor is defined by

$$\boldsymbol{J}_N = \omega * \left[(f_I)(f_I)^T \right] \tag{3}$$

where ω is a convolution kernel. In such an ideal setup, the vector $\boldsymbol{c}_N = (c_I)$ is a null eigenvector of $\boldsymbol{J}_N$ and, in practice, estimated as the eigenvector associated to the smallest eigenvalue. The velocities are recovered from $\boldsymbol{c}_N$ by the method described in [5], which is analytical for up to four motion layers. Obviously, the mixed-motion parameters can be computed only if the null eigenvalue is non-degenerated. For a single motion, the degeneracy of the null eigenvalue of $\boldsymbol{J}_1$ is known to be equivalent to the aperture problem. In what follows, we will show that *generalized aperture problems* are equivalent to the degeneracy of the eigenvalues of $\boldsymbol{J}_N$ and are thus reflected in the ranks of $\boldsymbol{J}_N$.

3. The projective plane

The projective plane is the set of all directions in the three-dimensional Euclidean space. These directions can be represented by non-null vectors but this representation is not unique since two collinear vectors represent the same direction. The representation can be made unique by choosing to represent directions by vectors having their last non-null component always equal to 1. A subset of these vectors have their last component equal to zero. The points represented by such vectors are called ideal points and the set of all ideal points is the ideal line. Therefore, we can think of the projective plane as the union of the plane $\xi_t = 1$ and the ideal line, where $\xi_x \boldsymbol{e}_x + \xi_y \boldsymbol{e}_y + \xi_t \boldsymbol{e}_t$ are points of the three-dimensional Euclidean space.

3.1. Relevant properties of the projective plane

We now summarize the properties of the projective plane that will be useful for the 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 $\boldsymbol{V}$ 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 $\boldsymbol{e}_t$ are aligned.

3.2. Representation of multiple motions in the projective plane

It is well known that if a spatio-temporal signal f is the superposition of moving layers, its Fourier transform is the superposition of Dirac planes through the origin of the Fourier space. The normal to each of these planes carries the motion vector of the respective layer. In this section,

these results will be interpreted in terms of the projective plane.

For this reason we introduce the following projective transform. Let F be the Fourier transform of f (the transform, of course, would apply to any function). We define a *projective transform* of F by

$$\mathcal{P}_F(\boldsymbol{P}) = \int_{\ell_{\boldsymbol{P}}} |F(\boldsymbol{\xi}, \xi_t)| \, d\boldsymbol{\xi} d\xi_t \qquad (4)$$

where $\boldsymbol{P}$ is a point of the projective space and $\ell_{\boldsymbol{P}} = \{\lambda \boldsymbol{P}; \lambda \in \mathbb{R}\}$ is the line directed by $\boldsymbol{P}$. This transform enables us to associate to each moving pattern a line or a point in the projective plane and simplifies the geometric reasoning.

To illustrate the usefulness of the framework, we show how to geometrically determine the velocity of one 2D moving pattern: the moving pattern is mapped to a plane in the Fourier domain, where it is further projected onto the projective plane resulting in a Dirac line. The velocity is picked by applying the duality, here denoted with $\mathcal{D}$, to the Dirac line. The process is schematically shown below:

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

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

$$\text{moving grating} \circ\!\!\xrightarrow{\mathcal{F}}\!\!\bullet \text{ line} \circ\!\!\xrightarrow{\mathcal{P}}\!\!\bullet \text{ point} \circ\!\!\xrightarrow{\mathcal{D}}\!\!\bullet \begin{smallmatrix}\text{admissible}\\\text{velocities}\end{smallmatrix} \quad .$$

We summarize the main points below:

- The projective transform of F is the superposition of Dirac lines in the projective plane (in case of 2D patterns);

- The dual point to each Dirac line in the projective plane is the velocity of the respective layer;

- For a 1D pattern, its Dirac line in the projective plane further reduces to a Dirac point. In this case any admissible velocity for the grating is a point of the line dual to the Dirac point in the projective plane;

- Dirac lines intersect at an ideal point if and only if their corresponding spatial velocities are collinear;

- A Dirac line supported by the ideal line corresponds to a static pattern.

As a further example, we show how to determine the *coherent motion* of superimposed gratings (plaids) [1]: 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.

3.3. The patterns in the projective plane and the rank of $\boldsymbol{J}_N$

Up to now we have derived a correspondence between different motion patterns and subsets of the projective plane (points and lines). The problem of determining the rank of $\boldsymbol{J}_N$ is equivalent to the problem of finding the largest number of independent null-eigenvectors. In this framework, it is a matter of choosing the right points in the support of Dirac lines in the projective plane. We give the detail in the appendix. Table 1 summarizes the correspondence between motion patterns and the ranks of $\boldsymbol{J}_1, \boldsymbol{J}_2$ and $\boldsymbol{J}_3$.

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 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.

4.1. Two 1D transparent moving gratings

In the projective plane, two moving gratings correspond to the {line, line} case. According to the theory, the perceived motion should correspond to the intersection point $\boldsymbol{U}$ 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
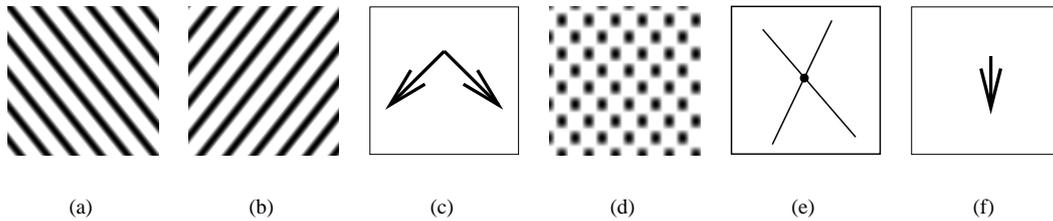
(a) (b) (c) (d) (e) (f)

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 lines as shown in (e).
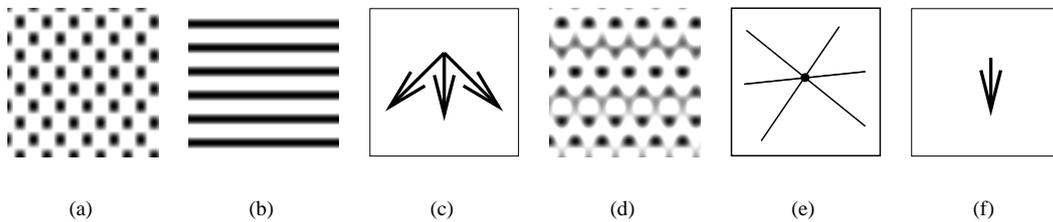

(a) (b) (c) (d) (e) (f)

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.

or see combinations of one 1D pattern and a 2D plaid pattern.

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 $\psi = k\pi/4, k = 1, .., 8$, and viewing angle size $10° \times 10°$. These were translated perpendicular to their orientation ($\phi_g = \psi \pm \pi/2$) with a velocity of $v_g = 1.6°/s$. 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. 5(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.

The shape of an aperture through which a grating can be seen can also strongly influence motion perception. To compare the strength of this so-called barberpole illusion with that of the effect described above, we constructed stimuli as above, but covered by an aperture perpendicular to the orientation of the grating. Nevertheless, results in Fig. 5(b) are qualitatively similar to those in Fig. 5(a).

5. Discussion

We have presented a way of categorizing transparent-motion patterns in terms of the ranks of the generalized structure tensors. Based on these 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 [5]. For example, we can discriminate
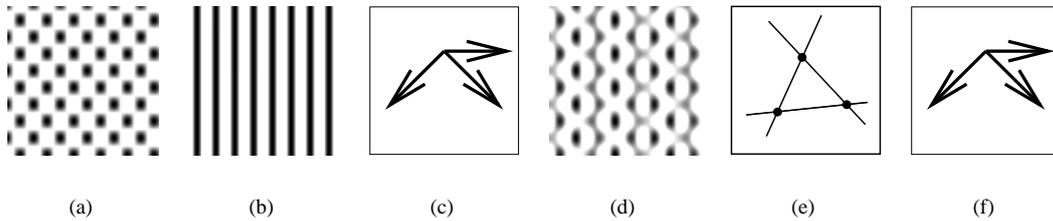
Figure 3. Incoherent motion of three superimposed gratings. The sub-figures are according to those in Fig. 2. However, the directions of motions are now changed such that the lines of motion in the projective plane do not intersect in a single point (e). This makes the motions undefined and causes the percept to change dramatically such that a coherent motion is not perceived. Observers can see either of the single motions indicated in (f).
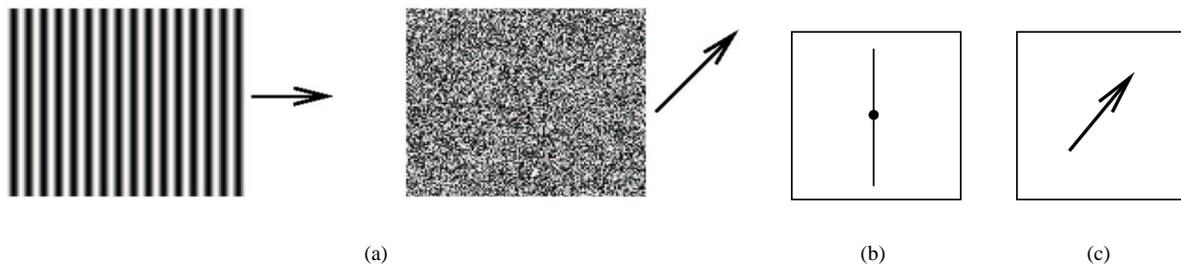

Figure 4. Stimulus generation for the 2D-over-1D entrainment (a). Admissible velocities for the grating (line) and for the 2D stimulus (point) are perceived as single motion (c).

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* [9, 2]. In the current framework, the intrinsic dimension corresponds to the rank of J_1. As shown in Table 1, 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 an N larger than 3, and the different moving patterns.

We have also shown how our results can be used to describe some phenomena in biological vision. In particular, the concept of the projective plane proved useful for describing and visualizing different visual percepts. Fur-

thermore, we predicted new illusionary percepts that are in accordance with the ambiguities that one would expect from the theory.

6. Appendix: The rank of J_N

From the discussion in Section 3.2, 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.

The mixed-motion parameters vectors $c_N = c(v_1, \ldots, v_N)$ 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 $\{c(v_1, \ldots, v_N) : V_n \in \beta,$ for $n = 1, \ldots, N\}$ 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
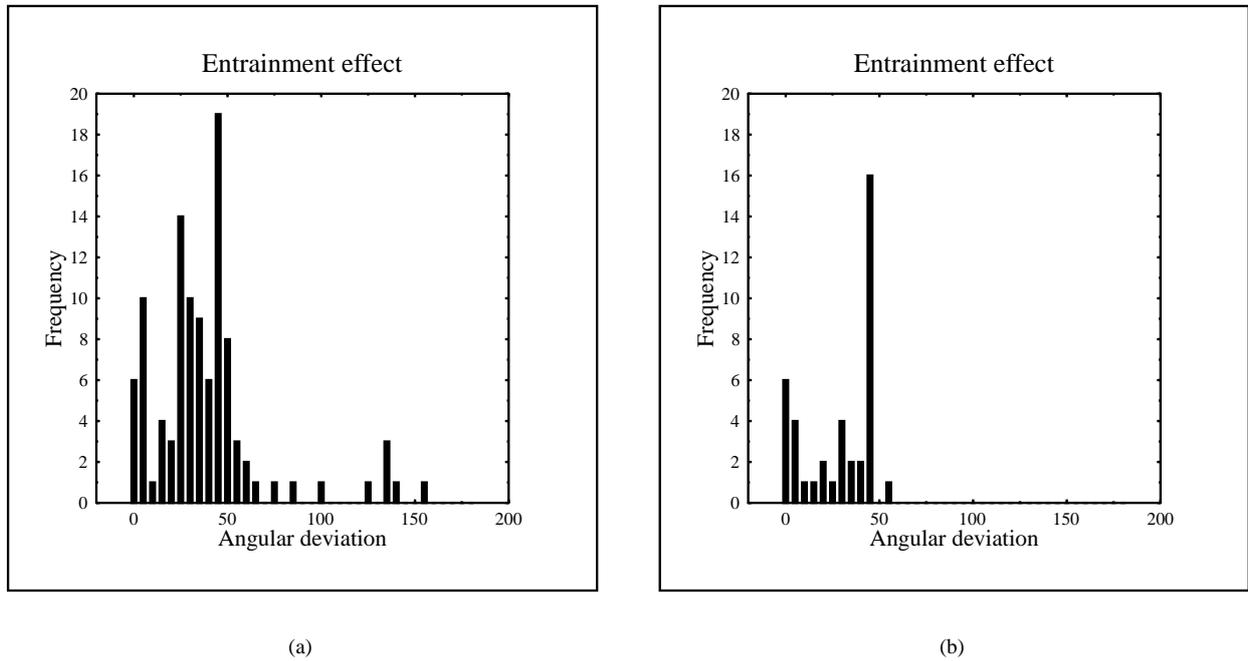
(a)

(b)

Figure 5. 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).
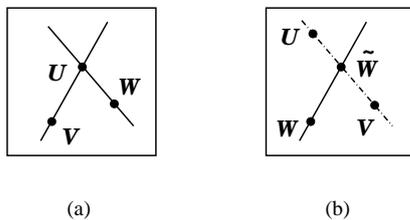

(a)   (b)

Figure 6. Admisible 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.

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.

6.1. The rank of J_2

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

line, line: Choose a basis $\beta = \{U, V, W\}$ of $\mathbb{R}^3$ such that U is intersection of the two lines, and V and W belong to each of these lines, see Fig. 6(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 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 rank$(J_2) = 1$.

6.2. 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. 6(b). In principle it appears that only the element $c(u, v, w)$ belongs to the kernel of J_3. To reveal another one, 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 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, 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 1 summarizes the possibilities for the ranks of J_N for $N = 1, 2, 3$.

7. Acknowledgment

8. References

[1] E. H. Adelson and J. A. Movshon. Phenomenal coherence of moving visual patterns. *Nature*, 300(5892):523–5, 1982.

[2] E. Barth and A. B. Watson. A geometric framework for nonlinear visual coding. *Optics Express*, 7:155–185, 2000. http://www.opticsexpress.org/oearchive/source/23045.htm.

[3] 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.

[4] 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.

[5] 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.

[6] J. A. Movshon, E. H. Adelson, M. Gizzi, and W. T. Newsome. The analysis of moving visual patterns. In C. Chagas, R. Gattas, and C. G. Gross, editors, *Study group on pattern recognition mechanisms*, pages 117–151. Pontifica Academia Scientiarum, Vatican City, 1985.

[7] 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.

[8] 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.

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