Movement Detectors of the Correlation Type Provide Sufficient Information for Local Computation of the 2-D Velocity Field

W. Reichardt[1], *R.W. Schlögl*[2], *and M. Egelhaaf*[1]

[1]Max-Planck-Institut für Biologische Kybernetik,
Spemannstrasse 38, D-7400 Tübingen, Fed. Rep. of Germany
[2]Max-Planck-Institut für Biophysik,
Kennedyallee 70, D-6000 Frankfurt/Main, Fed. Rep. of Germany

The projection of the velocity vectors of objects moving in three-dimensional space on the image plane of an eye or a camera can be described in terms of a vector field. This so-called *2-D velocity field* is time-dependent and assigns the direction and magnitude of a velocity vector to each point in the image plane. The 2-D velocity field, however, is a purely geometrical concept and does not directly represent the input site of a visual information processing system. The only information available to a visual system is given by the time-dependent brightness values as sensed in the image plane by photoreceptors or their technical equivalents. From spatio-temporal coherences in these changing brightness patterns motion information is computed. This poses the question about whether the spatio-temporal brightness distributions contain sufficient information to calculate the correct 2-D velocity field. Here we show that the 2-D velocity field generated by motion parallel to the image plane can be computed by purely local mechanisms.

In the literature on both biological motion processing and computer vision there is often stated that the 2-D velocity field cannot be computed by any local mechanism [1-6]. This conclusion is mainly based on approaches that implicitly regard a moving contour as nothing but a series of straight line segments which are each seen through a small aperture by some local motion analysing mechanism (see Fig.1). Information on the local curvature of the contour is, thus, not taken into account. From these mathematical approximations it is then concluded that all that a local mechanism can do is to determine the component of the local velocity vector perpendicular to the contour line, i.e. in the direction of the brightness gradient. If $F(x,y,t)$ =

Springer Series in Synergetics Vol. 42: **Neural and Synergetic Computers**
Editor: H. Haken

F[x+s(x,y,t); y+r(x,y,t)] represents the brightness of the moving pattern as a function of the spatial location x,y and time t, where s(x,y,t) and r(x,y,t) denote the time-dependent displacement of the pattern in the x- and y-direction, respectively, the mapping of the 2-D velocity vectors $v(x,y,t) = [ds(x,y,t)/dt; dr(x,y,t)/dt]$ on their components along the brightness gradient vectors $v^{\perp}(x,y,t)$ can be represented by the transformation

$$v^{\perp} = 1/(F_x^2+F_y^2) \cdot \begin{bmatrix} F_x^2 & F_xF_y \\ F_yF_x & F_y^2 \end{bmatrix} \cdot v \quad . \qquad (1)$$

(The subscripts denote the partial derivatives of F with respect to x or y). Since this transformation is not one-to-one, an infinite number of velocity vectors is mapped onto the same $v^{\perp}$. This ambiguity is commonly refered to as the *aperture problem* [1-5]. Using this type of representation of motion information, the correct 2-D velocity field, therefore, cannot be measured locally. Instead, the correct 2-D velocity of moving objects or pattern segments can only be computed in a further stage of analysis by combining the motion measurements from different locations and taking some global constraints into account [3-7].

These conclusions, however, should not be generalized, as is often done in the literature [2-5], to motion detection schemes which yield different representations of motion information. In particular, the so-called aperture problem is just a by-product of a specific mathematical approximation of the spatio-temporal geometry of a moving contour. In contrast we will show in the following that the correct 2-D velocity field can, in principle, be calculated by purely local mechanisms without reference to additional global constraints.

Our approach to these problems differs from the aforementioned ones. It is based on a movement detection scheme, the so-called *correlation-type of movement detector*, which has originally been derived from experiments on motion vision in insects [8,9], but in the meantime has been shown to account for certain aspects in motion vision of other species including man [10-13]. The visual field is assumed to be covered by a two-dimensional array of local movement

detectors which evaluate a kind of spatio-temporal cross-correlation between the light intensity fluctuations at two neighbouring points in space. More specifically, each movement detector has two spatially displaced input stages and consists of two mirror-symmetrical subunits (Fig.2). The input signal of one branch of each subunit is delayed in some way and multiplied with the instantaneous signal of the neighbouring input channel. The final output of the detector is then given by the difference between the outputs of the two subunits. Of course, a single movement detector senses only those components of motion which result in intensity changes along the orientation of its axis. Therefore, the outputs of a pair of *differently oriented* detectors at each retinal location are combined to a vector in order to obtain a two-dimensional representation of local motion (see Fig.3). The total of these local vectors, thus, represents a vector field which indicates the direction and magnitude of the local motion measurements.

Although computed by the movement detectors from the temporal modulations at their input stages, the field of local motion measurements can be related mathematically to the corresponding 2-D velocity field. This transformation can be described best on the basis of a formal approach which is characterized by a transition from an array of movement detectors with a discrete spatial sampling base to a continuous field of detectors with the distance between the neighbouring retinal inputs being infinitesimally small [14-17]. With the pattern function $F(x,y,t)$ and the 2-D velocity vectors $\mathbf{v}(x,y,t)$ the local response vectors $\mathbf{v}^*(x,y,t)$ of the movement detection system are given, in a first approximation, by the following transformation [16,17]

$$\mathbf{v}^* = T \cdot \mathbf{v} \,. \tag{2}$$

T represents a tensor which is proportional to the detector delay ϵ and has elements depending in a non-linear way on the pattern brightness function F and its first and second partial derivatives with respect to x and y

$$T = -\epsilon \begin{bmatrix} F_x^2 - F \cdot F_{xx} & F_x F_y - F \cdot F_{xy} \\ F_y F_x - F \cdot F_{yx} & F_y^2 - F \cdot F_{yy} \end{bmatrix} \,. \tag{3}$$

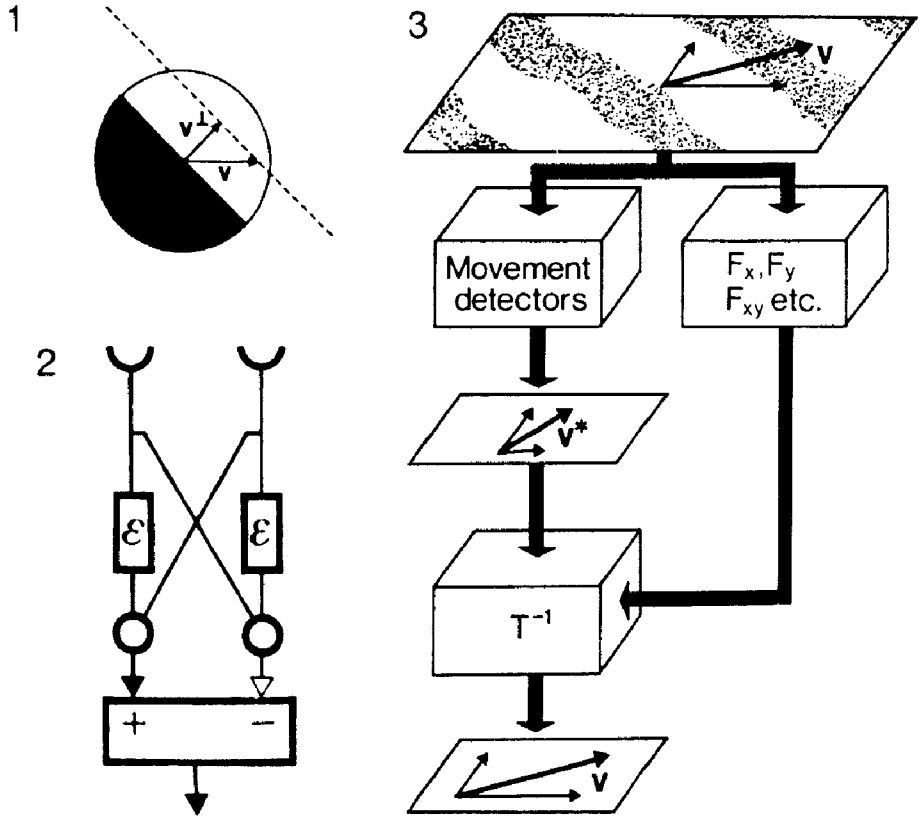

Fig. 1 An illustration of the so-called aperture problem in motion computation. A straight line segment moved with the velocity $\mathbf{v}$ is viewed through a local aperture. The only component of motion that can be computed is oriented perpendicular $\mathbf{v}^{\perp}$ to the orientation of the segment.

Fig. 2 A local movement detector consisting of two spatially displaced input stages and two mirror-symmetrical subunits. The input signal of one branch of each subunit is delayed by a brief time interval ϵ and multiplied with the undelayed signal of the neighbouring input channel. The output of the detector is given by the difference of the outputs of the two subunits. Therefore, it forms a movement direction selective device.

Fig. 3 A possible procedure for computing the correct 2-D velocity locally. A contrast element of a brightness pattern F moves according to the vector $\mathbf{v}$ and a pair of motion detectors reponds. Their outputs form the two components of the response vector $\mathbf{v}^*$. $\mathbf{v}^*$ and $\mathbf{v}$ are related by a two-dimensional tensor T, as described in the text. If the elements of the tensor T which contain spatial derivatives of the brightness pattern function are computed in parallel to the movement detector outputs and if the matrix of the tensor can be inverted, $\mathbf{v}$ may be computed.

By comparing the transformation of the 2-D velocity field described by equations (2) and (3) and equation (1), respectively, it is obvious that $\mathbf{v}^*$ usually deviates from the direction of the brightness gradient. This difference depends in a characteristic way on the curvature of the brightness function of the moving pattern. Moreover, the local response vectors, in general, also do not coincide with the correct 2-D velocity vectors. The occurrence of the second partial derivatives of the pattern brightness function with respect to the spatial coordinates might be surprising at first sight, since (in its discrete form) a movement detector samples the visual surround at only two spatial locations (see Fig.2). However, at least three points are necessary for an approximation of a second derivative. Due to the memory-like operation of the delay in one branch of each detector subunit (see Fig.2) three independent points of the pattern brightness function are simultaneously represented.

Because of the characteristic dependence of T on the curvature of the pattern brightness function, the map of $\mathbf{v}$ on $\mathbf{v}^*$ given by the transformation (2) is one-to-one for most stimulus patterns. This is the case, if the determinant of T does not vanish. T can then be inverted and equation (2) solved for $\mathbf{v}$

$$\mathbf{v} = T^{-1} \cdot \mathbf{v}^* . \tag{4}$$

In this way the correct 2-D velocity field can be calculated by using only local information about the pattern (see Fig.3).

There is only one special class of brightness pattern functions for which T cannot be inverted at any spatial location and, consequently, leads to ambiguous local motion measurements. This is the case if the determinant of T vanishes. These pattern functions can be analyzed most conveniently by the substitution $F(x,y,t) = e^{q(x,y,t)}$, which is possible for all $F \geq 0$ [17]. In other words, $q(x,y,t)$ represents the logarithm of the brightness pattern function $F(x,y,t)$. Using this substitution the tensor in equation (3) reads

$$T = \epsilon \cdot e^{2q} \begin{bmatrix} q_{xx} & q_{xy} \\ q_{yx} & q_{yy} \end{bmatrix} \tag{5}$$

and its determinant vanishes if the following condition is satisfied:

$$\det T = q_{xx}\, q_{yy} - q_{xy}^2 = 0 \,. \tag{6}$$

Again, the subscripts denote partial derivatives with respect to x and y. The only solutions of this partial differential equation for x and y are spatial brightness distributions the logarithm of which represents so-called *developable surfaces* [18]. Intuitively, a developable surface is one that can be cut open and flattened out. Or more precisely, at any location on a developable surface one can find a tangent that lies in the surface and has the same surface normal for all of its points. Spatial brightness distributions which, on a logarithmic scale, can be described as *cylindrical*, *conical* or *tangent* surfaces are the only possible examples of developable surfaces. For a cylindrical surface the tangents are all parallel, for a conical surface the tangents intersect in a common point and for a tangent developable surface the surface makes contact with one point of an arbitrarily given space curve. In

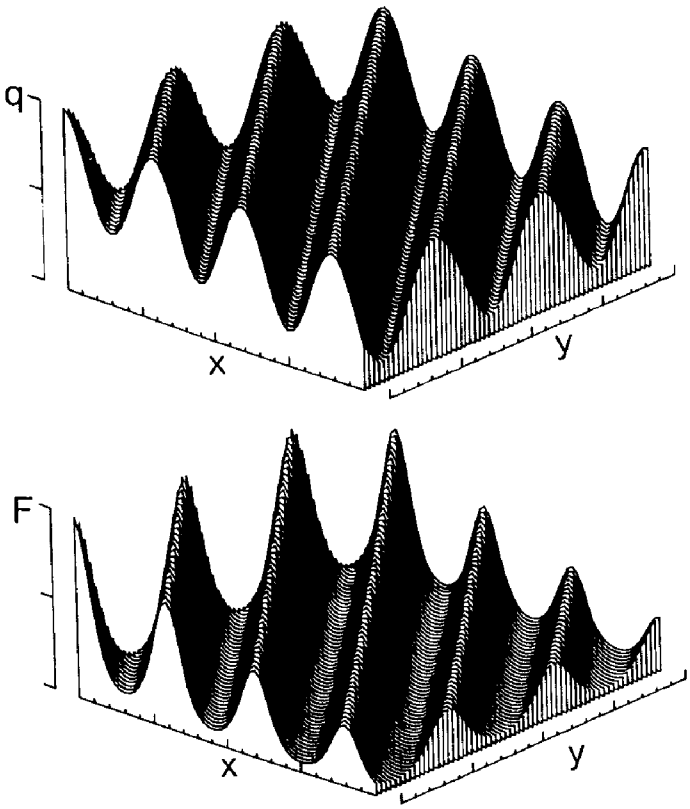

Fig. 4 An example of a developable surface of the *cylinder* class. The upper part of the figure shows the surface q = q(x,y), whereas the lower part the corresponding surface F = F(x,y).

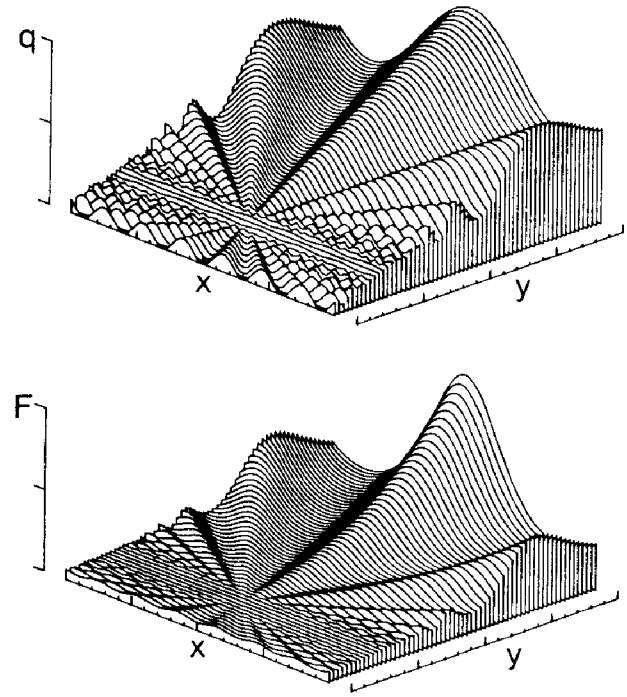

Fig. 5 An example of a developable surface of the *cone* class. The upper part of the figure shows the surface $q = q(x,y)$ and the lower part the corresponding surface $F = F(x,y)$.

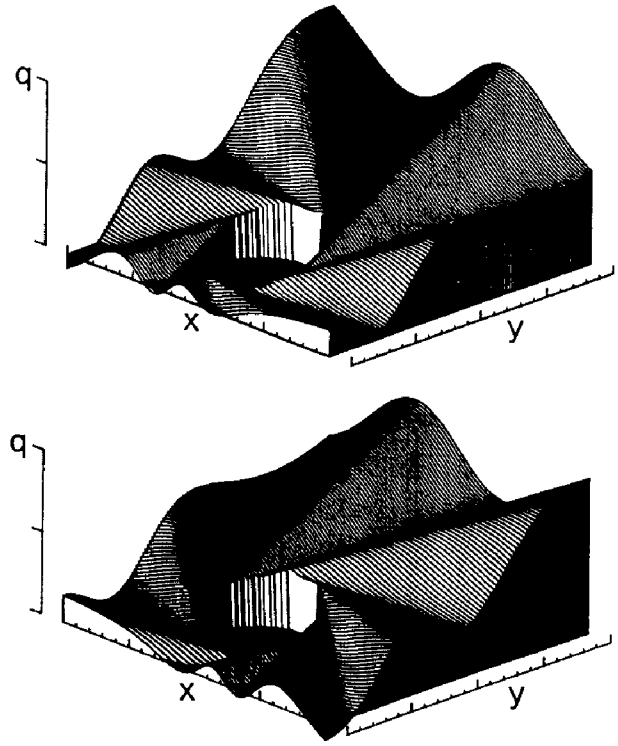

Fig. 6a An example of a developable surface of the *tangent* class.
a) The upper part of the figure shows the outer sheet and the lower part, the inner sheet of the surface $q = q(x,y)$.

Figs. 4 to 6 examples of each of the three classes of developable surfaces are given.

Natural brightness patterns usually cannot be represented *globally* by developable surfaces and, therefore, do not solve equation (6) for all x and y. However, since we focus here on local mechanisms, the possibility that the determinant of T vanishes only *locally* has also to be taken into account. If this happens for particular values x,y and t, equation (2) cannot be solved at just this location and time. This suggests that apart from certain locations, the correct 2-D velocity field can usually be recovered. Only at these locations one is confronted with equivalent ambiguities in the local motion measurements as described above in connection with the so-called aperture problem. On the basis of the motion detection scheme used here these ambiguities are usually restricted to small segments of natural brightness patterns. Spatial integration over an array of local motion detectors is a simple means to overcome this remaining problem in most cases. It should be noted that these conclusions are solely based

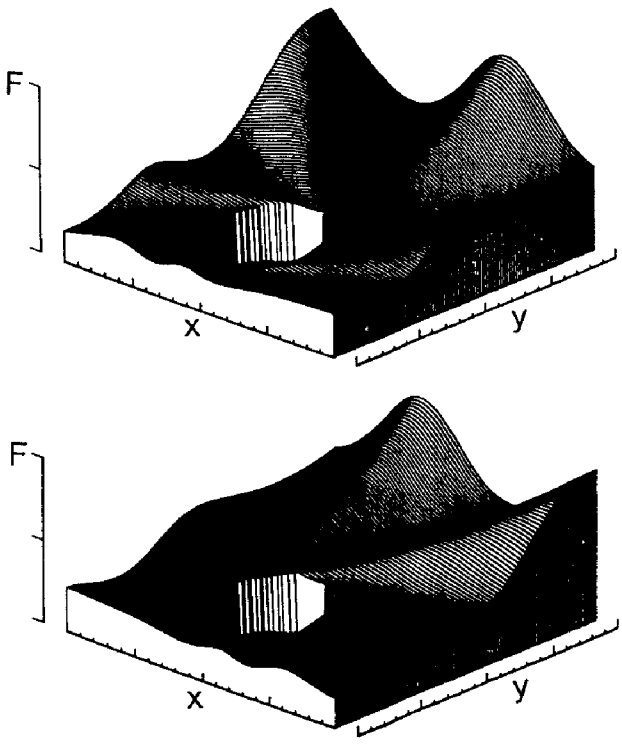

Fig. 6b An example of a developable surface of the *tangent* class. The upper part of the figure shows the corresponding outer sheet and the lower part the corresponding inner sheet of the surface F = F(x,y).

on a mathematical analysis. It was not intended here to address the problems which might arise when equation (4) is tried to be solved numerically.

Of course, the correct 2-D velocity field can only be computed from the local motion measurements by using equation (4), if the elements of T are explicitly known. These have to be derived from separate measurements in parallel to motion detection. It should be emphasized that while we have, in the correlation-type of biological movement detectors, a physiologically established and technically plausible implementation of the mechanism that yields $\mathbf{v}^*$, it is beyond the scope of this article to propose algorithms which yield the elements of the tensor. A technical solution is comparatively simple, since algorithms approximating spatial derivatives are in common use in computer vision [1,19].

In conclusion, there is no principle reason why it should generally be impossible to compute the correct 2-D velocity field for moving stimulus patterns on the basis of local mechanisms alone. The correlation-type of movement detectors as derived from biological systems form an appropriate basis to accomplish this task. This is because it yields local motion measurements which usually contain sufficient information on the relevant properties of the stimulus pattern.

References

1. Horn, B.K.P.: Robot Vision. MIT Press, Cambridge, Mass., 1987)
2. Ullman, S.: Trends in Neurosci. 6, 177 (1983)
3. Hildreth, E.C.: Proc. R. Soc. Lond. B. 221, 189 (1984)
4. Hildreth, E.C., C. Koch.: Ann. Rev. Neurosci. 10, 477 (1987)
5. Koch, C.: In Neural Computers, ed. R. Eckmiller, Ch. von der Malsburg (Springer, Berlin, Heidelberg 1988) p. 101
6. Adelson, E.H., J.A. Movshon: Nature 300, 523 (1982)
7. Movshon, J.A., D.H. Adelson, M.S. Gizzi, W.S. Newesome: In Pattern Recognition Mechanisms, ed.C. Chagas, R. Gattass, C. Gross (Springer, Berlin, Heidelberg, New York, Tokyo 1985) p. 117

8. Hassenstein, B., W. Reichardt: Naturforsch. 11b, 513(1956)
9. Reichardt, W.:In Principles of Sensory Communication, ed W.A. Rosenblith (John Wiley, New York 1961) p. 303
10. van Doorn, A.J., J.J. Koenderink: Exp.Brain.Res. 45, 179 (1982)
11. van Doorn, A.J., J.J. Koenderink: ibid. 45, 189 (1982)
12. van Santen, J.P.H., G. Sperling: J. Opt. Soc. Am. A 1, 451 (1984)
13. Wilson, H.R.: Biol. Cybern. 51, 213 (1985)
14. Reichardt, W., A. Guo: ibid. 53, 285 (1986)
15. Egelhaaf, M., W. Reichardt: ibid. 56, 69 (1987)
16. Reichardt, W.: J. Comp. Physiol. A 161, 533 (1987)
17. Reichardt, W., R.W. Schlögl: in preparation (1988)
18. Courant, R., D. Hilbert: Methods of Mathematical Physics, Vol.II (Interscience Publishers, New York, London 1962)
19. Ballard, D.H., C.M. Brown: Computer Vision. (Prentice-Hall, Inc., Englewood Cliffs, New Jersey 1982)