SALIENCY FROM MOTION AND FORM COMBINED

Andrea Pavan, Concetta Alberti, Gianluca Campana, Massimiliano Martinelli, Clara Casco
Dipartimento di Psicologia Generale, Università di Padova, Via Venezia 8, 35131, Italy
E-mail: andrea.pavan.1@unipd.it

Abstract

Detectability of an element with a trajectory composed of many small jumps is enhanced relatively to the detectability of individual motion (2 frames), more when the trajectory is straight than when it changes direction by a small amount each frame (Watamaniuk, McKee & Grzywacz, 1995). We asked whether enhancement is higher with line-segments oriented in the direction of motion (iso-motion) or orthogonal to it (ortho-motion). At high curvature (where successive presentation of motion signal are not co-axial), we found higher saliency for ortho-motion, suggesting that saliency in texture segmentation is accounted by the orthogonal velocity component ($V_{\perp}$). However, at small jittering, saliency of iso- and ortho-motion did not differ, suggesting that when the two saliency mechanisms (responding either to co-axiality or to $V_{\perp}$) are both activated, they compete for salience.

Wertheimer's (1923) famous demonstration of "common fate" shows that motion is a powerful cue for figure-ground segmentation: dots that move together group together and segment from the randomly moving dots in the background. Since then, several studies have attempted to define the mechanism responsible for increased saliency of a moving signal as a consequence of its motion in a consistent direction (Braddick, 1974; Casco & Morgan; 1987; Casco, Morgan & Ward, 1989). Watamaniuk, McKee & Grzywacz (1995) showed that a single dot moving on a trajectory can be detected when it is presented amidst noise dots whose movements are indistinguishable from the trajectory-dot on a frame-to-frame basis. The only distinguishing characteristics of the trajectory-dot is that it moves in the same direction over time while the noise dots change their direction of movement randomly each frame. Motion in a consistent trajectory is not the only factor that accounts for the saliency of moving signals. Since at a first stage of motion analysis in V1 most neurons respond to a combination of orientation and direction of motion, both the information from orientation and direction of motion have to contribute to segmentation, at least at early stage of computation. This can occur in different ways.

First, although the signal and background elements have the same orientation, saliency of signal may depend on the relation between its own orientation and the orientation of its motion path: either the same (iso-motion) or different (ortho-motion). Saliency should be higher for iso-motion signals. Indeed, it has been shown that, in static viewing, detection of the presence of a set of similarly oriented target elements in a background of randomly oriented elements is easier when targets are pseudo-aligned and collinear than when they are pseudo-aligned and parallel (Field, Hayes & Hess, 1993). This is due to long-range facilitatory interactions between neurons with iso-oriented and co-axial receptive fields (Gilbert, 1998), when optimally stimulated. Tanaka & Sagi (1999) have demonstrated that this facilitation is also present when the temporal dimension is added. The evidence of facilitation with both simultaneous and delayed elements, lead to the prediction that in an

apparent motion paradigm, there is higher saliency when orientation and direction of motion are collinear (iso-motion) than when they are orthogonal (ortho-motion). This hypothesis is supported by the finding that contour configuration is important in combining the motion detector outputs along a curved contour of discontinued elements (Ledgeway & Hess, 2002).
There is however another factor that may produce an asymmetry in the saliency of iso-motion and ortho-motion signals. This arises from the "aperture problem", which results because any motion unit, with a receptive field of a limited spatial extent, has access only to a motion component ($V_{\perp}$) normal to the orientation of a motion contour that encompasses its limits (Hildreth, 1984). Although line terminators are visible, and several authors (Lorenceau, Giersch & Seriès, 2005) have shown that they are relevant in solving the aperture problem, Casco and co-workers (1999, 2001, 2006) have shown that segregation of a moving target on the bases of orientation contrast is strongly facilitated when the orientation is orthogonal to the direction of motion. These findings predict higher saliency for the ortho-motion than for iso-motion condition since $V_{\perp}$ is higher in the former.
We aimed at assessing the role of these two mechanism in accounting for the saliency of motion signals by using displays of line elements of the same orientation and a paradigm similar to that used by Watamaniuk et al. (1995). We compared saliency of iso-motion and ortho-motion signals with different levels of frame-to-frame spatial jittering of the signal. At high jittering (where iso-oriented elements are not co-axial), we found higher saliency for ortho-motion, suggesting that saliency in texture segmentation is accounted by the orthogonal velocity component (Casco et al., 2001; 2006). However, at small jittering, we found no significant difference in saliency between iso-motion and ortho-motion, suggesting that when the two saliency mechanisms (collinearity and $V_{\perp}$) are both activated (at the same level of processing), they compete for salience.

General Method

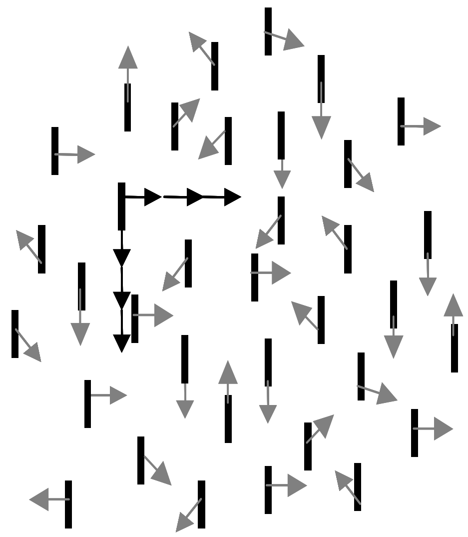
Stimuli

Stimuli, generated by a Cambridge Research System VSG graphics card free-viewed binocularly at 114 cm viewing distance. Stimuli (Figure 1) were random-line kinematograms in which each line moved in a 9-frames sequence (frame duration = 50 ms and frame interval equal to 0 ms), with a constant step size (6 arcmin) and random direction (chosen amongst 12 directions: 0, 45, 90, 135, 180, 255, 270 and 315 deg) except one line (the signal) that also moved of the same constant step size (6 arcmin) but in a fixed direction. Signal motion, whose initial position was chosen randomly within an inner square of the background, was either parallel (iso-motion) or orthogonal (ortho-motion) to its motion direction, in which it could be horizontal (rightwards) or vertical (downwards). A block of 120 trials consisted of 30 repetitions of four conditions randomly presented: a) vertical line-segment moving vertically (iso-motion) b) vertical line-segment moving horizontally (ortho-motion) c) horizontal line-segment moving horizontally (iso-motion) and d) horizontal line-segment moving vertically (ortho-motion). Lines speed were equal to 2°/sec. Elements and background luminance were 21.92 cd/m^2 and 1.46 cd/m^2 respectively.

Procedure

Observers performed a two alternative forced choice task and were asked to discriminate the orientation path of the motion signal. Performance was measured for direction-defined contours of varying straightness. In independent blocks, the straightness of the path was modified by jittering the position of the line-segment signal randomly from frame-to-frame of various amount (± 0, 3, 9, 15, 21 arcmin). Jittering had the effect of varying the straightness

1a



1b

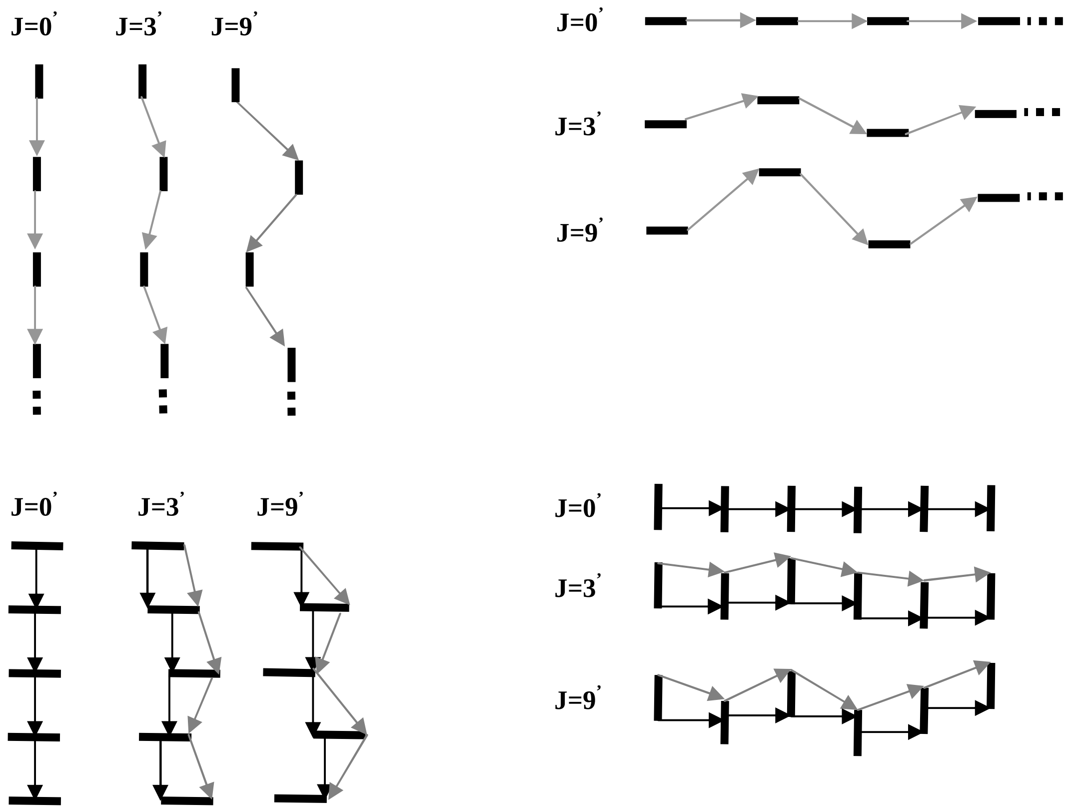


Figure 1. Figure 1a shows a scale schematic representation of a straight trajectory embedded in random-direction motion noise. Spatial positions and vectors are shown for a single frame, for background line segments and four frames, for the motion signal. Figure 1b shows how vectors (and consequent path curvature) change with different levels of jittering (J = 0, 3, 9 arcmin).

of the path (collinearity in the temporal dimension) as well as the orientation of the motion vectors at the terminators. Iso- and ortho-motion however differed because only in the ortho-motion condition $V_\perp$ was present. Three authors and one naïve subjects served as observers.

Results and Discussion

The results of four subjects (the authors and one naïve subject) are shown in Figure 2. For all subjects performance in the iso-motion condition decreases with path angle and is at chance in the large jittering condition. These results confirm the view that, with both static and moving contours, contour configuration is important in that straight and moderately curved contours are detected through combination across space (Field & Hayes, 2003). In the ortho-motion condition instead performance was high regardless of jittering, clearly indicating that the orthogonal motion component rather than the component at the terminators, accounts for performance.

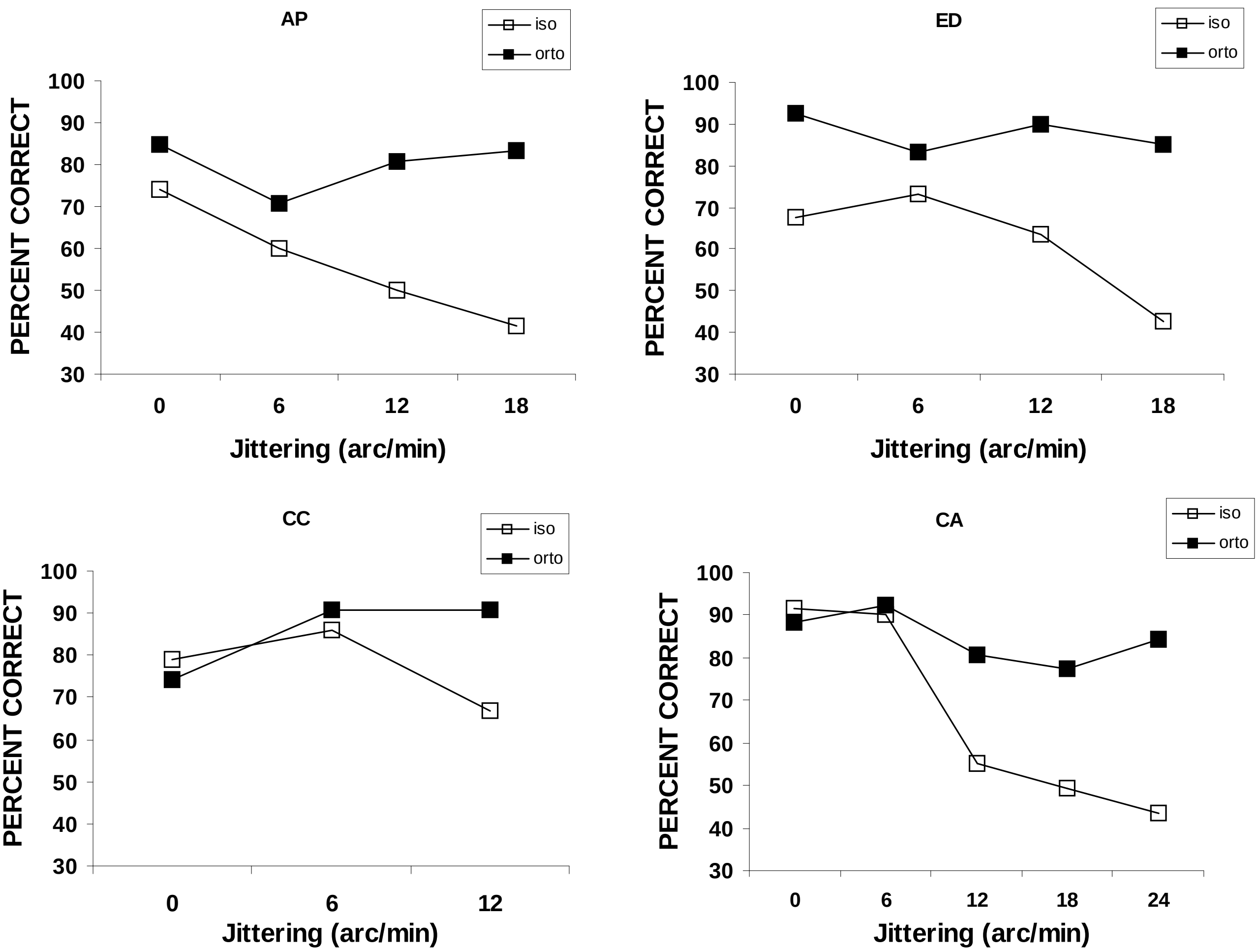


Figure 2. Percent correct performance in contour orientation discrimination is plotted against the amount of jittering (specifying the curvature of the contour) for iso-motion and ortho-motion.

Conclusion

The common fate phenomenon demonstrates that saliency from motion does not depend simply on motion perception resulting from interpretation imposed on figures perceived as segregated from the background at different locations and times (3rd-order motion, Lu & Sperling, 2001) but it results from extraction of the velocity of spatial-temporal variation of light intensity (1st-order motion, Adelson & Bergen, 1985; Watson & Ahumada, 1985; van Santen & Sperling, 1985).

In addition of confirming the role of common fate in accounting for saliency from motion at early level of computation, we isolated another mechanism which accounts for increased saliency of moving objects in addition to perceived speed and motion direction.

In agreement with our previous results (Casco et al., 2001, 2006), we showed that saliency of moving target results from $V_{\perp}$. This implies an integration process that takes together form and motion and responds to $V_{\perp}$.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America (A), 2,* 284-299.

Braddick, O. (1974). A short range process in apparent motion. *Vision Research, 14*, 519-525.

Casco, C., & Morgan, M. J. (1987). Detection of moving local density differences in dynamic random patterns by human observers. *Perception, 16*, 711-717.

Casco, C., Morgan, M. J., & Ward, R. M. (1989). Spatial properties of mechanism for detection of moving dot targets in dynamic visual noise. *Perception, 18*, 285-291.

Casco, C., & Ganis, G. (1999). Parallel search for conjunctions with stimuli in apparent motion. *Perception, 14*, 105-126.

Casco, C., Caputo, G., & Grieco, A. (2001). Discrimination of anorientation difference in dynamic textures. *Vision Research, 41*, 275-284.

Casco, C., Grieco, A., Giora, E., & Martinelli, M. (2006). Saliency from orthogonal velocity component in texture segregation. *Vision Research, 46*, 1091-1098.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local “association field”. *Vision Research, 33,* 173-193.

Field, D., & Hayes, A. (2003). Contour Integration and the Lateral Connections of V1 Neurons. In: Chalupa, L. M. and Werner, J. S. editors. *The Visual Neuroscience, vol. II.* The MIT Press, pp. 1069-1079.

Gilbert, C. D. (1998). Adult cortical dynamics, *Physiological Review, 78*, 467-485.

Hildreth, E. R. (1984). The computation of velocity field. *Proceedings of the Royal Society of London, (B), 221*, 189-220.

Ledgeway, T., & Hess, R. F. (2002). Rules for combining the outputs of local motion detectors to define simple contours. *Vision Research, 42*, 653-659.

Lorenceau, J., Giersch, A., & Seriès, P. (2005). Dynamic competition between contour integration and contour segmentation probed with moving stimuli. *Vision Research, 45*, 103-116.

Lu, Z. L., & Sperling, G. (2001). Three-systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America (A), 18*, 2331-2370.

Tanaka, Y., & Sagi, D. (1998). Long-lasting, long-range detection facilitation. *Vision Research, 38*, 2591-2599.

Van Santen, J. P. H., & Sperling, G. (1985). A temporal covariance model of motion perception. *Journal of the Optical Society of America (A), 1*, 451-473.

Watamaniuk, N. J., McKee, S.P., & Grzywacz, N. M. (1995). Detecting a trajectory embedded in random-direction motion noise. *Vision Research, 35*, 65-77.

Watson, A. B., & Ahumada, A. J. Jr. (1985). Model of human visual motion sensing. *Journal of the Optical Society of America (A), 2*, 322-341.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II. *Psychologische Forschung, 4*, 301-350.