

On the principle of minimal relative motion – the bar, the circle with a dot, and the ellipse

Zili Liu

Department of Psychology, UCLA, Los Angeles, CA, USA



Beghi, Xausa, & Zanforlin (1991a) and Beghi, Xausa, De Biasio, & Zanforlin (1991b) have presented visual stereokinetic phenomena. When a bar is rotated in the image plane, it appears to be slanted in depth. Likewise, when a circle with an off-centered dot is rotated, a three-dimensional (3-D) cone is perceived. Finally, when an ellipse is rotated in the image plane, an ellipsoid is perceived that is tilted in depth. To explain these phenomena, Beghi et al. (1991a,b) offer an analytic model that assumes that the visual system nullifies the speed differences between all stimulus points. I critique this analytic model, and show that it cannot explain the perceptual phenomena.

Keywords: stereokinesis, structure-from-motion, depth, rigidity, minimal relative motion

Introduction

Stereokinesis is a classical visual illusion. When a 2-D figure is rotated in the image plane, a 3-D structure is vividly perceived (Musatti, 1924). Like most other classical visual illusions, stereokinesis is not yet fully understood. To date, there are two major computational theories that attempt to explain stereokinetic phenomena. One is by Ullman (1979) who assumed that the 3-D percept is the outcome of maximizing rigidity of the stimulus structure. The other is by Zanforlin (1988, 2000) who proposed a minimal relative motion principle and suggested that the rigidity assumption is unnecessary but is often the outcome of his principle. The mathematical validity of this principle relies on the details of its application to each specific phenomenon. In this paper I will analyze three specific examples in two publications by Beghi et al. (1991a,b).

Beghi et al. (1991a,b) demonstrate three perceptual effects. First, when a bar is rotated in the image plane, it appears to be slanted in depth. Second, when a circle with an eccentric dot is rotated, a cone is perceived with the dot being the apex. Third, when an ellipse is rotated in the image plane, the following percept develops gradually over time. The ellipse first appears to deform, but remains in the image plane. Then the ellipse no longer deforms, but instead becomes a circular disk that is tilted in 3-D. Finally, this circular disk becomes a solid ellipsoid that is tilted in depth.

In order to explain these effects, Beghi et al. (1991a,b) offer a mathematical model that equalizes the speeds of all stimulus points. Specifically (p. 426, Beghi et al., 1991a):

“When a set of points have different velocities, the differences can be eliminated [by] adding a depth component. As a result of this minimization, all the points will appear to move with a common or unique velocity that defines the perceived 3-D configuration.”

Although the vector term ‘velocity’ is used here, only differences of the velocity magnitudes, i.e., speeds, are minimized, not of the directions, as will be clear. It is also helpful to quote how Beghi et al. (1991b) developed the mathematics regarding the tilted ellipsoid (p. 434):

“Let us assume that the contour points of the rotating ellipse equalize their velocities ... with respect to the different point on the major axis of the ellipse. ... Each point of the axis has, on the frontal plane, a different velocity depending on its distance from the rotation centre ... But it is possible to equalize the different velocities of all points of the axis and hence the velocities of all the contour points, by displacing the axis in depth by considering it as a rotating line of constant length ... When a ‘z’ component is added to all the points of the rotating line in order to equate their velocities, the line will appear tilted in depth at a well defined angle.”

For each of the three examples in the two papers, Beghi et al. (1991a,b) take two different approaches that seemingly converge to the same solution. I will analyze their mathematical derivations of the model, and will show that the model's predictions are not compatible with the empirical percepts reported in the same papers.

Methods

The Rotating Bar

As shown in Figure 1, when a bar of length l_0 is rotating with an angular speed ω_0 , the speed of an arbitrary point P on the bar, relative to the center C , is (equation before Equation 3 in Beghi et al., 1991a):

$$v_{PC} = \omega_0 l_0 \left| \lambda - \frac{1}{2} \right|, 0 \leq \lambda \leq 1. \quad (1)$$

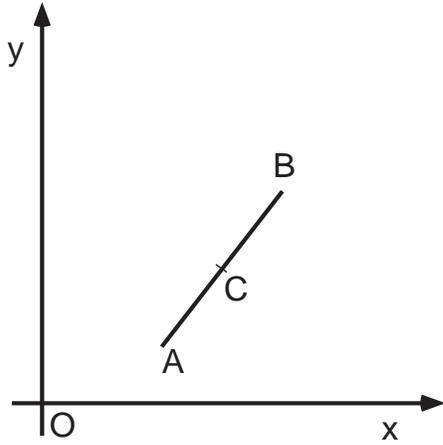


Figure 1. Bar AB is rotating around the origin O in plane Oxy . C is the bar's midpoint.

Obviously, when $\lambda = 1/2$, P coincides with C , and $v_{pc} = 0$. The next step is most critical since Beghi et al. (1991a) apply their principle of minimization. In their words (p. 427) (the two equations below are Equation 3 and the one immediately after in Beghi et al. (1991a)):

“Let us now associate to point P an additional velocity component $v_z(P)$

$$v_{PC}^2 + v_z^2(P) = \text{constant} = v_{\max}^2 = \left(\frac{\omega_0 l_0}{2}\right)^2, \quad (2)$$

from which it follows:

$$v_z(P) = \omega_0 l_0 \sqrt{\lambda - \lambda^2}.” \quad (3)$$

Let us examine $v_z(P)$ as a function of the position λ (in Beghi et al. (1991a), the equation below Equation 3 contains a typo, l_0 should be l_0). Note that

$$v_z(P = C) = v_{z_{\max}} = \pm \frac{\omega_0 l_0}{2}, \quad (4)$$

when $\lambda = 1/2$, and

$$v_z(P = A) = v_z(P = B) = 0, \quad (5)$$

when $\lambda = 0$ and 1 , respectively. That is to say, with the introduction of the additional velocity component in depth, the bar's midpoint C will have the largest depth displacement (the direction of which depends on the sign of $v_z(P=C)$), while its two end points A and B will remain on the image plane. This means that the bar cannot possibly remain rigid, contradicting the reported percept. It is impossible for a rotating bar to have the same speed everywhere and remain rigid.

I believe that the case has already been made at this point regarding the inadequacy of the minimal relative speed principle. I would like to further comment on the two methods, analytic and trajectory, of the minimal speed difference model that give rise to the same value

when computing the depth difference between A and B . Although the true depth difference between A and B should be zero, Beghi et al. (1991a) obtain a different value based on their trajectory derivation. They compute this displacement within a time period $t_0 = \pi/\omega_0$ as follows (Equation 4 of Beghi et al. (1991a), p. 428) (A time period of π/ω_0 , or of the half rotation, is used, because, according to Footnote 2 of Beghi et al. (1991a), “the rotating segment will return to its initial position after a full rotation” (p. 427). It is unclear, however, what triggers v_z to reverse its direction at half rotation.):

$$\Delta z = \pi l_0 \int_0^1 d\lambda \sqrt{\lambda(1-\lambda)} = \frac{\pi^2}{8} l_0. \quad (6)$$

I will now show that this value is in fact the area swept by the bar during this time divided by l_0 , but not the depth difference between A and B . I will assume, with Beghi et al. (1991a), that the bar BCA moves in the positive z direction. Note that v_z is only a function of position λ , not of time t . Within time interval π/ω_0 , this area is (let $\cos\theta = 2\lambda - 1$):

$$\begin{aligned} \int_0^{t_0} dt \int_0^1 d\lambda l_0 v_z(\lambda) &= \pi l_0^2 \int_0^1 d\lambda \sqrt{\lambda(1-\lambda)} \\ &= \frac{\pi l_0^2}{4} \int_{-\pi}^0 d(\cos\theta) \sqrt{1 - \cos^2\theta} = \frac{\pi^2}{8} l_0^2 = l_0 \Delta z. \end{aligned} \quad (7)$$

Beghi et al. (1991a) employ a second method, which assumes that point P moves from B to A within time interval $t_0 = \pi/\omega_0$ with a constant speed $v_0 = l_0/t_0$. (There is a minor inconsistency in Beghi et al. (1991a) with regard to which point (A or B) is the starting point and which one is the end point. Equation 1 of Beghi et al. (1991a) clearly indicates that B is the starting point. Then this is inconsistent with the first equation on p. 428. Regardless of the starting point, the last two equations in Equation 2 ($x_p = \dots, y_p = \dots$) are incorrect.) Specifically, P 's speed relative to the midpoint C , v_p , is calculated first (the 3rd equation from the bottom left on p. 428):

$$v_p^2(t) = \left[\frac{d(x_p - x_C)}{dt} \right]^2 + \left[\frac{d(y_p - y_C)}{dt} \right]^2, \quad (8)$$

then the minimal speed difference principle is applied (the 2nd equation from top right on p. 428):

$$v_p^2(t) + v_z^2(t) = \text{constant} = v_{p_{\max}}^2, \quad (9)$$

where $v_z(t)$ is the z -component of the motion of point P . Finally, the depth displacement $\Delta z = \int_0^{t_0} dt v_z(t)$ is calculated. Although the same displacement is obtained $(\pi^2/8)l_0$, the following needs to be clarified.

The $v_p^2(t)$ derived in Beghi et al. (1991a) is incorrect (the second equation from the bottom left on p. 428). It does not appear to be a typo since follow-up equations are derived from it. It is unclear how, at the end, the z displacement $(\pi^2/8)l_0$ is obtained that is consistent with

the analytic method. In fact, a z displacement of $(\pi^2/4)l_0$ should have been the correct solution following these equations. The correct $v_P^2(t)$ that gives rise to the z displacement $(\pi^2/8)l_0$ should be:

$$v_P^2(t) = (\omega_0 l_0)^2 \left[\frac{1}{\pi^2} + \left(\frac{t}{t_0} - \frac{1}{2} \right)^2 \right]. \tag{10}$$

This can be seen by observing that (I thank Reviewer #1 who pointed this out):

$$x_P - x_C = l_0 \left(\frac{t}{t_0} - \frac{1}{2} \right) \cos(\omega_0 t + \theta_0 + \phi_0), \tag{11}$$

$$y_P - y_C = l_0 \left(\frac{t}{t_0} - \frac{1}{2} \right) \sin(\omega_0 t + \theta_0 + \phi_0).$$

An intuitive way to understand Equation 10 is that P has two orthogonal velocity components relative to C : one along the bar (constant) $\omega_0 l_0 / \pi = v_0$, the other perpendicular to the bar (rotation) $\omega_0 l_0 (t/t_0 - 1/2)$.

The Circle With an Off-centered Dot

As shown in Figure 2, the length of OP in the triangle COP is $\sqrt{r^2 + r_0^2 - 2rr_0 \cos \phi}$, where $r = |CP|$ is the radius of the circle, and r_0 is the length $|CO|$. (Without loss of generality, in Beghi et al. (1991a), the origin O of the Oxy system is positioned half way between the center of the circle, C , and the off-centered dot E . The same convention is adopted here.) Hence, the speed of point P is (the 4th equation from bottom left, p. 429):

$$v(P) = \omega_0 \sqrt{r^2 + r_0^2 - 2rr_0 \cos \phi}. \tag{12}$$

By adding an additional velocity component $v_z(P)$ in depth, Beghi et al. (1991a) obtain (their equations 5 & 6):

$$v^2(P) + v_z^2(P) = \text{constant} = v_{\max}^2(P) = [\omega_0 (r + r_0)]^2. \tag{13}$$

Then (Equation 8 of Beghi et al. (1991a)):

$$v_z(P) = \pm \omega_0 \sqrt{2rr_0(1 + \cos \phi)} \\ = \pm 2\omega_0 \sqrt{rr_0} \left| \cos \frac{\phi}{2} \right|, 0 \leq \phi < 2\pi. \tag{14}$$

The remainder of this section assigns the 3-D coordinates of the circle and dot without mathematical reasoning. So there are no derivations to be analyzed. Still, the introduction of $v_z(P)$ makes the circle non-planar in 3-D due to the nonlinearity of Equation 14 (as in the case of the rotating bar). This contradicts the prediction of the model, e.g., in Figure 1b of Beghi et al. (1991a), and of the empirical perceptual result. Since the alternative method in Beghi et al. (1991a) claims that the circle remains planar in 3-D (p. 430), which contradicts Equation 14, and since the alternative method claims to yield the same result, it is self contradictory.

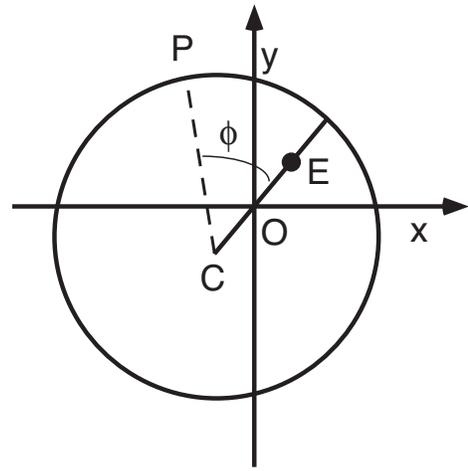


Figure 2. A circle and a dot E are rotating around the origin O in plane Oxy . O is halfway between E and the center of the circle C . P is an arbitrary point on the circle.

The Rotating Ellipse

The tilt in depth

Let us assume that the angular speed of rotation is ω_0 ; and that the length of the major axis is $2a$, as shown in Figure 3. (Without loss of generality, Beghi et al. (1991b) have moved the center of rotation to the center of the ellipse. I use the same convention for the ease of comparison.) Then, on the frontal plane, the speed $v(A)$ of the extreme point A of the ellipse on the major axis is (see Equation 3 of Beghi et al. (1991b)):

$$v(A) = \omega_0 a. \tag{15}$$

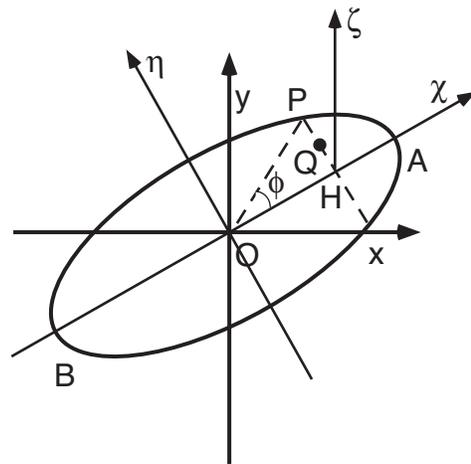


Figure 3. The ellipse is rotating in the plane Oxy around the origin O . AB is its major axis, $Ox\eta\xi$ is its canonical coordinate system. P is an arbitrary point on the ellipse. H is P 's projection onto the major axis, and Q an arbitrary point on PH .

Consequently, the speed $v(H)$ of point H on the major axis is (the equation after Equation 12 in Beghi et al. (1991b)):

$$v(H) = \mu\omega_0 a, \quad 0 \leq \mu \leq 1. \quad (16)$$

Obviously, $v(O) = 0$ where O is the center of rotation and of the ellipse itself.

The next step is most critical because the principle of minimal speed difference is applied. Beghi et al. (1991b) introduce a z velocity component at point H as follows (Equation 13 of Beghi et al. (1991b)):

$$v^2(H) + v_z^2(H) = \omega_0^2 a^2. \quad (17)$$

This leads to (Equation 14 of Beghi et al. (1991b)):

$$v_z^2(H) = \omega_0^2 a^2 (1 - \mu^2), \quad 0 \leq \mu \leq 1. \quad (18)$$

“The velocity component along the z direction would cause a 3-D displacement of the major axis of the ellipse” (Beghi et al. (1991b), p. 436). However, such a displacement takes an odd form. Note that when $\mu = 1$, $v_z^2(H=A) = 0$, and when $\mu = 0$, $v_z^2(H=O) = \omega_0^2 a^2$. This means that the extreme point A of the major axis remains in the plane of $z = 0$, whereas the center of the ellipse O takes the maximum displacement. Hence the problem here is identical to that of the rotating bar.

The displacement in depth

I believe that this critique has been sufficient to establish the inadequacy of the model. As an aside, I would like to further critique the derivation of the apparent length of the semi-major axis of the 3-D ellipsoid in Beghi et al. (1991b, p. 436):

In order to evaluate the displacement along the z direction, let us assume a uniform displacement of point H of the major axis of the ellipse ... from an extreme A toward the opposite extreme, B , ... within a time interval $\Delta t^* = \pi/\omega_0$ and with a velocity $v^* = l_0/\Delta t^*$, where $l_0 = 2a$... [F]rom which it follows

$$\begin{aligned} \Delta z &= \int_0^{\Delta t^*/2} v_z(t) dt = \pi l_0 \omega_0 \int_0^1 v_z(\mu) d\mu \\ &= \dots = \frac{\pi^2}{8} a. \end{aligned} \quad (19)$$

Thus the apparent length of the semimajor axis will be:

$$a^* = a \sqrt{1 + \left(\frac{\pi^2}{8}\right)^2}. \quad (20)$$

Apparently, Δz above is meant to be the displacement from the plane $z = 0$ of the extreme points A and B , which should be zero. Equation 19 above has in fact computed the area swept by the semi-major axis AO ,

divided by a . This can be clearly seen if Equation 19 is rewritten as:

$$\begin{aligned} &\int_0^{\Delta t^*/2} dt \int_0^1 a v_z(H) d\mu \\ &= \omega_0 a^2 \int_0^{\Delta t^*/2} dt \int_0^1 \sqrt{1 - \mu^2} d\mu = a \Delta z. \end{aligned} \quad (21)$$

The shape of the ellipsoid

Finally, it should be noted that the derivation of the ellipsoid itself remains puzzling. As shown in Figure 3 in the canonical reference system $O\chi\eta\zeta$ of the ellipse, the ellipse can be represented as (Equation 4 of Beghi et al. (1991b), in which the “ $p \leq \phi \leq e2\pi$ ” appears to be typo):

$$\chi_P = a \cos \phi, \eta_P = b \sin \phi, 0 \leq \phi < 2\pi, \quad (22)$$

where a, b are the semi-major and semi-minor axis, respectively. Then for a point Q inside the ellipse that satisfies (Equation 5 of Beghi et al. (1991b)):

$$\chi_Q = \lambda \cos \phi, \eta_Q = \lambda b \sin \phi, 0 \leq \lambda \leq 1, \quad (23)$$

it follows that (Equation 7 of Beghi et al. (1991b)):

$$v^2(Q) = \omega_0^2 [a^2 \cos^2 \phi + \lambda^2 b^2 \sin^2 \phi]. \quad (24)$$

Beghi et al. (1991b) then introduce an additional velocity component in the ζ direction (which is the same as the z direction) such that (Equation 8 of Beghi et al. (1991b)):

$$v^2(Q) + v_\zeta^2(Q) = v^2(P) = \omega_0^2 [c^2 \cos^2 \phi + b^2], \quad (25)$$

where $c^2 = a^2 - b^2$. Consequently (Equation 9 of Beghi et al. (1991b)) is

$$v_\zeta^2(Q) = \omega_0^2 b^2 (1 - \lambda^2) \sin^2 \phi. \quad (26)$$

Quoting again from Beghi et al. (1991b): “If we finally associate a third coordinate to Q

$$\zeta_Q = \frac{v_\zeta(Q)}{\omega_0} = \sqrt{b^2 (1 - \lambda^2) \sin^2 \phi}, \quad (27)$$

the following relation holds:

$$\eta_Q^2 + \zeta_Q^2 = b^2 \sin^2 \phi, \quad (28)$$

which is the equation of a circle centered in H .”

This is how Beghi et al. (1991b) derive an ellipsoid (p. 436). However, the introduction of the third coordinate to Q in Equation 27 is unjustified, since $\zeta_Q = v_\zeta(Q)/\omega_0$ is true only if ζ_Q represents the distance of Q from the rotational axis and $v_\zeta(Q)$ is the rotational speed of Q . Neither is true here - recall that ζ and z represent the same direction, that v_ζ is the velocity component in the ζ direction, that the rotational axis is parallel to ζ and z , so that v_ζ is not a rotational velocity component. Moreover, $v_\zeta(Q)$ is supposed to be the ζ velocity component of point Q that is on the Oxy plane. It makes little sense why a ζ

coordinate with a value of $v_z(Q)/\omega_0$ is associated with this point of Q on the Oxy plane.

Conclusions

I have demonstrated that the minimal relative motion principle, or more precisely, the minimal relative speed difference principle, in Beghi et al. (1991a,b) cannot explain the perceptual phenomena reported. It appears that there is no easy fix for the problems. Qualitatively speaking, adding a depth motion component that depends on position but not on time will only keep deforming a figure forever and never give rise to a stationary percept. This contradicts empirical observation. For instance, in the case of the rotating bar, Beghi et al. (1991a) reported: "After a few seconds of inspection, the bar ... appears tilted in depth at a well defined angle" (p. 426). (I thank reviewer #1 for nicely summarizing up this point.)

I would like to emphasize that this paper focuses only on the mathematical aspects of Beghi et al. (1991a,b), but not on the empirical aspects.

Although non-rigidity inevitably results from the principle of minimal speed difference in the examples of these papers, it remains an open question whether rigidity needs to be assumed *a priori* (Braunstein & Andersen, 1986; Ullman 1984) or can still become a natural outcome of some other, perhaps more fundamental, minimization principle as Beghi et al. (1991a,b) have proposed. The recent "slow and smooth" hypothesis in 2-D motion perception (Weiss, Simoncelli, & Adelson, 2002; see also Grzywacz & Yuille, 1991; Heeger & Simoncelli, 1991; Hildreth, 1984; Simoncelli & Heeger, 1992), which is a different minimization principle and can be achieved via local computation, may provide a headway toward a percept that is often rigid in 3-D.

Acknowledgments

This research was supported by National Institute of Health Grant NEI EY-14113 and National Science Foundation Grant IBN-9817979. I thank Bas Rokkers for helpful discussions. Commercial relationships: none.

References

- Beghi, L., Xausa, E., & Zanforlin, M. (1991a). Analytic determination of the depth effect in stereokinetic phenomena without a rigidity assumption. *Biological Cybernetics*, *65*, 425-432. [PubMed]
- Beghi, L., Xausa, E., De Biasio, C., & Zanforlin, M. (1991b). Quantitative determination of the three-dimensional appearances of a rotating ellipse without a rigidity assumption. *Biological Cybernetics*, *65*, 433-440. [PubMed]
- Braunstein, M. L., & Andersen, G. J. (1986). Testing the rigidity assumption: a reply to Ullman. *Perception*, *15*, 641-646. [PubMed]
- Grzywacz, N., & Yuille, A. (1991). Theories for the visual perception of local velocity and coherent motion. In M. Landy and J. A. Movshon (Eds.), *Computational models of visual processing*. Cambridge, Massachusetts: MIT Press.
- Heeger, D., & Simoncelli, E. (1991). Model of visual motion sensing. In L. Harris & M. Jenkin (Eds.), *Spatial Vision in Humans and Robots*. Cambridge: Cambridge University Press.
- Hildreth, E. (1984). *The Measurement of Visual Motion*. Cambridge, MA: MIT Press.
- Musatti, C. L. (1924). Sui fenomeni stereocinetici. *Archivio Italiano di Psicologia*, *3*, 105-120.
- Simoncelli, E., & Heeger, D. (1992). A computational model for perception of two-dimensional pattern velocities [Abstract]. *Investigative Ophthalmology and Visual Science*, *33*, 954.
- Ullman, S. (1979). The Interpretation of Structure from Motion. *Proceedings of Royal Society of London, B*, *203*, 405-426. [PubMed]
- Ullman, S. (1984). Rigidity and misperceived motion. *Perception*, *13*, 219-220. [PubMed]
- Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, *5*, 598-604. [PubMed]
- Zanforlin, M. (1988). Stereokinetic phenomena as good gestalts. *Gestalt Theory*, *10*, 187-214.
- Zanforlin, M. (2000). The various appearances of a rotating ellipse and the minimum principle: a review and an experimental test with non-ambiguous percepts. *Gestalt Theory*, *22*, 157-184.