Optic Flow and the Metric of the Visual Ground Plane

JACK M. H. BEUSMANS*

Received 9 April 1997; in revised form 15 August 1997

A theory is developed in which the optic flow of an observer translating over the ground plane determines the metric of egocentric visual space. Optic flow is used to operationalize the equality of spatial intervals not unlike physicists use time to compare spatial intervals. The theory predicts empirical matching ratios for collinear, sagittal intervals to within 2\% of the mean (eight subjects, standard error also 2\%). The theory predicts that frontoparallel intervals on the ground plane will match sagittal intervals if their relative image motions match, which was found empirically. It is suggested that the optic flow metric serves to calibrate static depth cues such as angular elevation and binocular parallax. © 1998 Elsevier Science Ltd. All rights reserved.

INTRODUCTION

Visual space refers to the spatial structure a (human) visual system derives from and imposes on the external environment from which it receives its input. In the past, both the internal geometric properties of visual space as well as its geometrical relationships to the external world have been studied, leading to intrinsic and extrinsic geometries of visual space (Norman, Todd, Perotti & Tittle, 1996). The metric of external space itself is, of course, Euclidean for most practical purposes of humans, and it is clear that the external metric of visual space is non-Euclidean as the perceived length of a line segment depends on its orientation with respect to the observer (Norman et al., 1996). And at the beginning of this century, it was demonstrated by “alley” experiments that the intrinsic metric of visual space is generally also not Euclidean (see Graham, 1965; and Indow, 1991, for reviews). It was found that two lines of luminous points in an otherwise dark environment, forming an alley in front of the observer and adjusted such that they are perceptually parallel, do not appear to be equidistant along their length. This is only possible in a non-Euclidean visual space, and the particular outcome of the alley experiments is consistent with a hyperbolic geometry. These results led to the development of mathematical theories of large-scale visual space, the best known of which is perhaps Luneburg’s theory of binocular visual space (Luneburg, 1950; Blank, 1953, 1958, 1978). The particular geometry proposed by Luneburg is problematic because of its reliance on binocular convergence to encode egocentric distance, and it has found only limited experimental support (Heelan, 1983; Wagner, 1985; Cutting, 1986; Indow, 1991); his assumptions of constant negative curvature and homogeneity, for example, may not hold in practice (Foley, 1972, 1980; Battro, di Pierro Netto & Rozestraten, 1976; Indow, 1991).

The intrinsically non-Euclidean nature of visual space is not limited to the reduced-cue conditions that are typically used in experiments on alleys and binocular visual space. Lappin, Koenderink and van Doorn (1996) found that observers standing in an open field in broad daylight have a visual space which is elliptic within a few meters from the observer and hyperbolic farther away. Norman et al. (1996) also found that near space is elliptic. The many experiments on perceived egocentric distance, \(d_p\), in open fields also are consistent with a hyperbolic space (Gilinsky, 1951; Baird, 1970; Da Silva, 1985; Wiest & Bell, 1985; Philbeck, Loomis & Beall, 1997). Although the data are highly variable, a consistent finding is that perceived distance is a power function of real distance, \(d_p = k(d/d_c)^n\), with \(n < 1\) and \(d_c\) a constant depending on the observer; that is, distance is increasingly underestimated. A similar pattern is sometimes found when subjects are asked to blindly walk towards targets at various distances in front of them (Loomis, Da Silva, Fujita & Fukusima, 1992; Philbeck et al., 1997; Philbeck & Loomis, 1997). Such a nonlinear relationship between perceived and actual distance means that visual space is intrinsically elliptic for \(d < d_c\) and hyperbolic for \(d > d_c\). A linear compression of distance, i.e., \(n = 1\) and \(k/d_c < 1\) (as proposed by Wagner, 1985), would yield a visual space whose intrinsic metric is Euclidean.

*Cambridge Basic Research, Nissan Research and Development, Inc., 4 Cambridge Center, Cambridge, MA 02142, U.S.A. [Tel: (617) 374 9670; Fax: (617) 374 9697; Email: beusmans@pathfinder.cbr.com].
The question now is why visual space has this one particular structure and not another. The answer will lie in the kind of visual cues or combinations of cues that are used to construct visual space and the precise manner in which they are used. The most thoroughly investigated cue has undoubtedly been binocular disparity, partially because of its obvious relation to depth in near space and partially because it is amenable to mathematical analysis. But as mentioned above, experimental support for its role in far space has been limited. Experimental support for other cues that could be used in far space, such as the size of familiar objects, angular distance from the horizon, aerial perspective, and texture gradients has generally been rather equivocal and does not seem promising for a quantitative account of visual space (for comprehensive reviews see Sedgwick, 1986; Cutting & Vishton, 1995; Cutting, 1997). Some of the other well-known possible cues to distance such as accommodation and binocular parallax or convergence are limited to near visual space (Gogel, 1977; Kaufman, 1974). And even there, recent experiments on the wallpaper illusion suggest that convergence is not sufficient to perceive distance (Logvinenko & Belopolskii, 1994).

A fundamental problem with most, if not all, of these cues is that other information is needed for their calibration, if they are to provide metric information in far visual space. In near visual space, immediate feedback about distances is available through grasping hand movements and other interactions between the observer and the environment, but in far visual space, no such feedback exists under static conditions. That such continuous feedback is a necessary and normal component of sensorimotor systems is demonstrated by their fast adaptation to sensory and motor rearrangements (Welch, 1986; Moehl, 1989; Rieser, Pick, Ashmead & Garing, 1995; Wolf, Voss, Hein & Heisenberg, 1992). The idea to be explored in the present paper is that image motion provides the feedback essential for constructing a metric in far visual space and that this metric is used to calibrate static distance and depth cues.

Historically, motion parallax associated with lateral head movements has been studied most extensively. It has been found that motion parallax can be used to judge egocentric distance in near space (Ferris, 1972; Johansson, 1973; Nakayama & Loomis, 1974; Gogel & Tietz, 1979) and exocentric distances, i.e., distances between objects in the scene (Gibson, Gibson, Smith & Flock, 1959; Hagen & Teghtsoonian, 1981; Hell & Freeman, 1977; Hell, 1979; see Sedgwick, 1986, for a comprehensive review). One of the few studies not using lateral head movements to induce motion parallax is the Eriksson (1974) study in which observers judged distances while walking back and forth over a distance of 1.5 m. Motion parallax was found to dominate the cues of relative size of similar figures and of object height in the image. The relative expansion of visual angles associated with forward motion can also be used to estimate time-to-collision and hence distance (Lee, 1976; Regan, Kaufman, & Lincoln, 1986; Kaiser & Mowafy, 1993). Just like perceived distance, perceived time-to-collision follows a power law, although the exponent is somewhat smaller (0.72 over a range of 10 sec; van der Horst, 1991). This suggests that optic flow and egocentric distance may be closely and perhaps even causally related. In the above experiments, typically only a few isolated dots or objects in otherwise total darkness were used so that the associated image motion is too sparse to form a structured field. Dense fields of moving dots have been used more recently and lead to vivid percepts of object surface shape (Rogers & Graham, 1979; Braunein & Tittle, 1988). This suggests that large, dense optic flow fields could play a role in the perception of large-
scale visual space, and, in particular, the visual ground plane.

The main idea is that the optic flow generated by an observer translating over the ground plane can be used to compare spatial intervals. This is illustrated in Fig. 1, which is an adaptation of one of Gordon’s figures (Gordon, 1965). It shows what Gordon called the “positional field”: the projection of a grid of lines on the ground plane, which are separated by some fixed distance, here 3 m. It illustrates vividly the effect of linear perspective, namely that fixed spatial intervals subtend smaller angles with increasing distance to the observer. Superimposed on the positional field is the velocity or optic flow field of an observer moving at 3 m/sec. At this speed, the lines perpendicular to the observer’s direction of motion would map onto each other after a 1 sec period; that is, the line at a distance of 4.5 m would have moved to a distance of 1.5 m in front of the observer. If the observer could compare these two positional fields, i.e., compute the displacement field over this 1-sec period, then she would be able to invert the effect of linear perspective and veridically perceive the spatial intervals along the direction of travel. However, humans observers do not have direct access to the displacement field itself but can only measure its low-order approximations, the velocity or optic flow field being the first-order approximation. If optic flow vectors are used to equate spatial intervals, then spatial intervals will be more and more underestimated as their distance to the observer increases. Consider the velocity vector at 4.5 m in front of the observer; it reaches only about halfway to 1.5 m, so that it corresponds to an objective distance of approx. 1.5 m (ignoring nonlinearities in the local positional field). In contrast, the velocity vector at 13.5 m spans the entire 3 m interval. Thus, using optic flow vectors a 3 m interval at 13.5 m will be equated with a 1.5 m interval at 3.5 m; in other words, farther intervals are underestimated compared with closer ones. Perceived egocentric distance based on these intervals would be a power function of actual distance with an exponent that is less than 1.

The proposal is that the metric provided by optic flow becomes incorporated into the representation of the ground plane and is used to calibrate other cues to spatial distance such as elevation in the visual field. Thus, the metric derived from optic flow will manifest itself even when the observer is standing still. The remainder of the paper elaborates on the proposed use of optic flow to metricize visual space. The predictions of the optic flow theory were tested psychophysically by having eight subjects, standing in an open field in daylight, match spatial intervals in far visual space (between 3 and 30 m). An apparatus was used that allowed subjects to quickly adjust one interval to match another one without interference from the experimenter. The data matched the predictions well (Experiments 1 and 2), thus providing correlational evidence that optic flow is instrumental in determining the metric of the visual ground plane.

SPATIAL METRIC THROUGH OPTIC FLOW

The two coordinate systems that will be used to specify points in the plane and to describe and analyze the optic flow field are illustrated in Fig. 2. The Cartesian coordinate system is centered at the observer’s feet (Oc). Its x-direction is defined by the observer’s direction of motion v, which is illustrated as a vector from the observer’s cyclopean eye, Oc. The z-direction is perpendicular to the ground plane, in the direction of gravity. The plane spanned by x (or v) and z is the so-called sagittal plane. Spatial intervals along the direction of x will be called sagittal, and intervals along y will be called frontal. The intersection of the sagittal plane and the ground plane will be called the sagittal line. Upper-case letters refer to points and objects in the environment and lower-case letters refer to their projection onto the viewing sphere or image.

The observer’s cyclopean eye is at a height h above the ground plane, and it is the origin, Oc, of the spherical coordinate system of the viewing sphere. The elevation, θ, of a point P on the ground plane is the angle between the direction of P and the ground plane: θ = 〈(P, Oc), (P, Oc)〉, which is also its angular distance from the horizon. Thus, the horizon has an elevation of 0, and the observer’s feet have an elevation of −90 deg. The azimuth, φ, of P is the angle between the projection of P’s direction on the ground plane and the x-direction: φ = 〈(P, Oc), x〉. The spherical coordinate system (θ,φ) will also be the basis of the proposed representation of the ground plane (subsection: Representation of the ground plane).

Optic flow, v(p), at a point p in the image has components in the local θ and φ directions:

\[ v_\theta(p) = -v/h \sin^2 \theta \cos \phi \]  \hspace{1cm} (1)
\[ v_\phi(p) = v/h \sin \theta \sin \phi \]  \hspace{1cm} (2)

where \( v_\theta \) is the image motion component in the direction of increasing \( \theta \), \( v_\phi \) is the component in the direction of increasing \( \phi \), and \( v \) is the observer’s speed (Koenderink & van Doorn, 1981).

Collinear, sagittal intervals

Figure 3(A) illustrates how two collinear, sagittal intervals \( T_2T_1 \) and \( T_3T_2 \) can be equated. The optic flow
predicts that $t_3$ will move to $t_2$ and that $t_2$ will move to $t_1$ over some period $dt$. Let $\alpha_{ij}$ indicate the visual angle between $t_i$ and $t_j$; then $\alpha_{32}/v_0(t_3) = \alpha_{21}/v_0(t_2) = dt$, or equivalently, $\alpha_{32}/\alpha_{31} = v_0(t_2)/v_0(t_3) \approx 1 + 2\alpha_{32}/\sin\theta_3$. The intervals $T_3T_2$ and $T_2T_1$ will be called matching intervals when they satisfy this condition. Note that the ratio of two matching intervals does not depend on the observer’s speed; eye height is still important as it determines the elevation $\theta$ of an interval’s endpoints. To give a numerical example, suppose $v = 3.2$ m/sec (walking speed), $h = 1.6$ m, and $T_3 = 9$ m ($\theta_3 = -10$ deg) and $T_2 = 6$ m ($\theta_2 = -15$ deg), then $v_0(t_3) = 3.5$ deg/sec and $dt = 5/3.5 \approx 1.4$ sec; during that time $\alpha_{32}$ expands to the size of $\alpha_{21}$, which is an increase by a factor of 2. Returning to Fig. 3(A), $T_3T_2$ is smaller than $T_2T_0$ because the optic flow at $t_2$ only “reaches” $t_1$ and not $t_0$.

Image velocities are not the only way to compare intervals; the expansion or divergence of $t_3t_2$ could be used as well. For example, the expansion at the center of $\alpha_{32}$ during the time $dt = \alpha_{32}/v_0(t_3)$ can be used to predict the interval $t_2t_1$. These two methods lead to slightly different metrics because image motion increases nonlinearly with elevation (equation (1)).

Because $v_0$ increases with $\theta$, the visual angle of equivalent intervals increases towards the observer. However, because this increase is quadratic, i.e., faster than linear, and a simple first-order approximation of the flow field is used for the predictions, the closer equivalent intervals will be systematically underestimated in terms of their 3-D size. If higher-order predictions were used, equivalent intervals would become more veridical; the results of Experiment 1 suggest that the human visual system uses a first-order prediction, i.e., simple optic flow.

Specifically, let $v_0 = \theta$ be a hypothetical linear flow field and consider the fate of $t_3$ and $t_2$ during the period $dt = |\alpha_{32}/v_0(t_3)| = -(\theta_3 - \theta_2)/\theta_3$, that is, the time during which $t_3$ will move to $t_2$ at its present velocity. The predicted location of $t_2$ is: $t_2(t + dt) = \theta_2 + v_0 dt = \theta^2_2/\theta_3$. In other words, the visual system predicts that $t_3(t + dt) = t_2(t) = \theta_2$ and that $t_2(t + dt) = \theta^2_2/\theta_3$. In fact, however, it takes less time to move from $t_3$ to $t_2$ because image velocity increases with $\theta$. A closed form equation for the location of $t_3$ as a function of time is easily obtained from the differential equation $v_0 = d\theta/dt = \theta^2/\theta_3$. The actual time to reach $t_2$ is therefore $dt' = \log(\theta_2/\theta_3)$ and during this time $t_2$ actually moves to $t_2'(t + dt') = \theta_2 \exp(\log(\theta_2/\theta_3)) = \theta^2_2/\theta_3$. Thus, the predicted and actual relationships between the two
targets \( t_1 \) and \( t_2 \) over time are identical and no distortion of space would result.

Now, consider the actual flow field, \( v_0 = -v/h \sin^2 \theta \), associated with translation over a ground plane. The predicted location of \( t_2 \) is now \( t_2(t + dt) = t_2 + (\theta_2 - \theta_1) \sin^2 \theta_2 \sin^2 \theta_3 \). The corresponding differential equation is \( d\theta /dt = -v/(h \sin^2 \theta) \), whose closed form solution is \( t_3(t) = \arccot(\cot(t_1(0)) + tv/h) \). The actual time \( dt' \) is then \( dt' = (\cot(\theta_3) + \cot(\theta_2))h/v \), and \( t_3'(t + dt') = \arccot(2\cot(\theta_3) - \cot(\theta_2)) \). Numerical simulations show that \( [t_2(t + dt')] < [t_3(t + dt')], \) that is, the predicted location falls short of the actual location. This means that the 3-D size of closer intervals will be systematically underestimated.

If perceptually equivalent intervals are used as yardsticks to measure perceived distance from the observer, then the perceptual yardstick will correspond to larger and larger 3-D intervals as its distance from the observer increases. In other words, perceived distance would be progressively underestimated. Numerical simulations show that perceived distance, \( d_p \), with an optic flow metric would vary with real distance approximately as a power law: \( d_p = d^n \), with \( n < 1 \) (Fig. 4). The particular numerical value of \( n \) depends on the size of the yardstick; the smaller the yardstick the closer the optic flow prediction approximates veridicality and the closer \( n \) will be to unity. If a series of yardsticks is constructed starting with a 15 deg yardstick between \(-45\) and \(-60\) deg elevation, then \( n \) is approx. 0.79. Starting off with a yardstick of only 5 deg between \(-45\) and \(-50\) deg elevation, yields a value for \( n \) of approx. 0.93. Thus, the optic flow metric can explain why distances are generally, but not always, underestimated during perceptual judgments and blindfolded walking.

**Sagittal and frontal intervals**

Sagittal and frontal intervals cannot be brought into correspondence and compared directly during observer translation. They can, however, be compared on the basis of their relative image motions. Along the sagittal line, the magnitude of relative image motion becomes locally associated with 3-D interval size. That is, if the interval \( z_{32} \) were to increase by a factor of two, the associated relative motion, \( v_0(t_2) - v_0(t_3) \), would also increase by a factor of two to a first-order approximation, and so would the distance \( z_{32} \). In fact, \( v_0(t_2) - v_0(t_3) \approx (v/h) \sin \theta_3 \). Thus, the proposal is not to simply use relative motion as a cue to distance; this would not work because relative motion within a sagittal interval \( t_3t_2 \) of a fixed size increases as its elevation decreases (that is, as \( T_3T_2 \) gets closer to the observer). Moreover, relative motion is also proportional to the observer’s speed \( v \) (which could be factored out by normalizing by local image motion). The important point is that relative motion and interval size are always proportional, even though the exact value of the constant of proportionality may vary over time.

The proposal is to use the relative image motion in the sagittal and frontal directions as a measure for comparing spatial intervals. Figure 3(B) shows an example of a sagittal \((T_2T_1)\) and a frontal interval \((T_2T_3)\). Frontal intervals will be approximated by iso-elevation lines on the viewing sphere (i.e., lines of constant \( \theta \); see Fig. 1). This approximation is valid for distances beyond a few meters and for small frontal intervals; for example, the elevation of a frontal interval at 15 m varies from \(-6.1\) deg \((\phi = 0\) deg) to \(-5.7\) deg \((\phi = 20\) deg), and at 2 m it varies from \(-38.7\) to \(-36.9\) deg. In Experiments 1 and 2 distances were always larger than 1.5 m and frontal intervals were smaller than 20 deg for intervals closer than 10 m, and smaller than 10 deg for farther intervals. The image velocities associated with sagittal and frontal intervals are then [Fig. 3(B)]:

\[
\begin{align*}
v_0(t_1) &= -v/h \sin^2 \theta_1 \\
v_0(t_2) &= v_0(t_3) = -v/h \sin^2 \theta_2 \\
v_0(t_1) &= v_0(t_2) = 0 \\
v_0(t_3) &= -v/h \sin \theta_3 \sin \phi_3.
\end{align*}
\]

Equating the motion of \( t_1 \) relative to \( t_2 \) with the motion of \( t_1 \) relative to \( t_2 \) yields:

\[
\begin{align*}
v_0(t_3) - v_0(t_2) &= v_0(t_1) - v_0(t_2) \\
v/h \sin \theta_3 \sin \phi_3 &= v_0(t_1) - v_0(t_2).
\end{align*}
\]

These relative image motions can be quite large. For example, if the elevation of \( T_2 \) and \( T_3 \) is \(-10\) deg and \( \phi_3 = 10\) deg and assuming a walking speed of 3.2 m/sec, then frontal velocity \( v_0(t_3) = 3.5 \) deg/sec. The same
FIGURE 5. Top view of the experimental set-up for matching sagittal intervals.

relative speed along the sagittal line is obtained when \( \theta(t_1) \approx -15 \text{ deg} \).

Assuming that \( \sin \phi_3 \approx \phi_3 \) and letting \( \Delta \theta = \theta_2 - \theta_1 \) and \( \Delta \phi = \phi_3 - \phi_1 = \phi_3 \), we obtain:

\[
\Delta \phi = 2 \cos \theta_2 \Delta \theta - (\cos^2 \theta_2 / \sin \theta_2)(\Delta \theta)^2.
\]

For small \( \Delta \theta \) and assuming \( \cos \theta_2 \approx 1 \),

\[
\Delta \phi \\
\Delta \theta \approx 2.
\]

Thus, matching a frontal interval to a given sagittal interval by equating the relative motions at the intervals’ endpoints yields a frontal interval that subtends a visual angle approximately twice that of the sagittal interval. This relationship between relative expansion in sagittal and frontal directions had been noted earlier by Koenderink & van Doorn (1981). As \( \theta_2 \) gets larger, that is, the intervals get closer to the observer, the ratio \( \Delta \phi / \Delta \theta \) decreases. For example, if \( T_2 \) is one eye height in front of the observer so that \( \theta_2 = -45 \text{ deg} \), \( \Delta \phi = 2 \cos 45 \text{ deg} \Delta \theta \approx 1.4 \Delta \theta \); for two eye heights, \( \Delta \phi = 2 \cos 27 \text{ deg} \Delta \theta \approx 1.8 \Delta \theta \). Conversely, as \( \theta_2 \) gets smaller equation (10) can no longer be used as a valid approximation, and equation (9) must be used. It shows that \( \Delta \phi \) will increase rapidly as \( \sin \theta_2 \) becomes much smaller than \( \Delta \theta \).

The comparison of the sagittal and frontal intervals on the basis of relative motion does not depend on the observer’s speed; eye height plays some role because it determines the elevation of objects in the image.

**Representation of the ground plane**

The metric in the sagittal direction (\( M_{s} \); subsection: Collinear sagittal intervals) and the metric for comparing sagittal and frontal directions (\( M_{f} \); subsection: Sagittal and frontal intervals) are combined into one coherent observer-centered representation of the ground plane: \( G = (M_{s}, M_{f}) \). The metric \( M_{s} \) is used for all possible directions from the observer (i.e., all possible \( \phi \) with \(-90 \text{ deg} \leq \theta \leq 0 \text{ deg}\)); that is, the representation is isotropic and does not depend on \( \phi \) (azimuth). Also note that \( G \) does not depend on observer speed and that observer eye height enters only indirectly through its relationship with elevation.

The optic flow metric is used to calibrate static cues to distance and depth on the ground plane, such as binocular disparity and elevation in the visual field. Thus, the flow metric will be apparent whenever an observer perceives a ground plane even when not moving at that particular moment. As an aside, the elevation in the visual field may not necessarily be measured with respect to the horizon, which can be difficult to locate (Bingham, 1993; Sedgwick, 1986). An alternative is to measure elevation with respect to the direction of the observer’s feet, so that the horizon has an elevation of 90 deg. A priori, it seems more sensible to use the visual angle subtended by the observer’s feet (or some other fixed direction associated with the observer, such as the hood of the car he is driving) and some distant target on the ground plane rather than the angle between that target and the horizon, the projection of points at infinity. This would explain why estimates of the distance of a lead car are affected by the amount of roadway visible between a lead and following car, when the elevation of the lead car is kept constant (Evans & Rothery, 1976).

**EXPERIMENT 1: EQUATING COLLINEAR, SAGITTAL INTERVALS**

**Materials and methods**

Subjects had to match the 3-D size of two collinear, sagittal intervals (Fig. 5), that were demarcated by three bright orange markers \( T_1, T_2, \) and \( T_3 \). \( T_1 \) and \( T_2 \) were rectangles (7 cm wide, 10 cm high) and \( T_3 \) was a triangle (10 cm high, base of 15 cm); \( T_3 \) was larger than the other two targets in order to keep it visible at the larger distances used in the experiment. Subjects were instructed to match the 3-D size of the closer interval \( T_1T_2 \) to that of the farther interval \( T_3T_2 \); that is, they were instructed to match the intervals in terms of their objective, 3-D physical size. Marker \( T_2 \) was fixed throughout the experiment. A system of pulleys allowed the subjects to adjust the closer interval by moving \( T_1 \) along a thin line approx. 14 cm above the ground and extending from \( T_3 \) towards the observer. This made it possible to match the two intervals very quickly and without interference from the experimenter. Note that,
strictly speaking, this procedure differs from a bisection task in which the middle target T2 is moved back and forth, thereby simultaneously changing both the nearer and the farther interval. During the match, the experimenter was standing off to the side, so as not to introduce size or distance cues.

Across trials, the distance of the observer to marker T2 was varied between 3 and 20 m and the farther interval (T2T3) was varied between 0.5 and 12 m, resulting in 14 combinations (Table 1). This block of 14 trials was repeated three times. In half the trials, T1 and T2 were very close to begin with so that subjects had to increase the interval size. In the other half, T1 and T2 were very far apart initially so that subjects had to decrease the interval’s size. After making the match, the experimenter measured interval T1T2 and removed T3. Next, the subject was instructed to walk to a new location marked by a golf tee for the next trial (golf tees were numbered by their distance in meters from T2; they were only visible when looking straight down at them from above; the grass was approx. 5–10 cm long); note that the subject’s walking between trials was not intended to induce optic flow and is not crucial to the experiment.

Subjects viewed the markers binocularly and received no feedback about their matching performance. Subjects had to maintain collinearity with respect to the targets throughout the experiment; that is, they were not permitted a side view after making a match. They could, however, make slight head and body movements. The experiment lasted between 30 and 40 min per subject.

Subjects. Eight subjects (ages 20-38 yr), who, except for the author (JB), were naive as to the purpose of the experiment.

Results

Figure 6(A) summarizes the main results in terms of the actual 3-D size of matching nearer and farther sagittal intervals. Data are expressed as the ratio of the nearer and farther intervals and are plotted as a function of the size of the farther interval. Data points with the same T2 distance are connected and labeled according to distance (labels 3, 5, A, F and K indicate 3, 5, 10, 15, and 20 m distance to T2).

It is clear from Fig. 6(A) that subjects did not make veridical judgments in most conditions but tended to equate the farther interval with a smaller nearer interval. The error in judgment increases with egocentric distance and interval size, up to approx. 35% for an interval of 12 m at 20 m. Two nearby, small intervals were overestimated by approx. 10% (the 0.5 m farther interval at 5 m was equated with a closer interval of 0.56 m, and the 1.5 m farther interval at 10 m was equated with an interval of 1.64).

### Table 1: Locations of the targets T2 and T3 comprising the farther intervals

<table>
<thead>
<tr>
<th>Distance to T2</th>
<th>3</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.5</td>
<td>1.5</td>
<td>3.0</td>
<td>4.0</td>
<td></td>
</tr>
<tr>
<td>1.0</td>
<td>1.0</td>
<td>3.0</td>
<td>6.0</td>
<td>8.0</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>4.5</td>
<td>10.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

For each distance T2, the distance between T2 and T3 is listed (all distances in meters).
Figure 6(B) and (C) show the results that would have been obtained if the two intervals had been equated on the basis of optic flow (B) or binocular disparities (C). It is clear that matches were not made by simply equating disparities, which would have led to much larger underestimates. In contrast, optic flow predictions are quite close to the empirical data.

To further quantify how well the optic flow metric
predicts the data, the subsequent analysis will be in terms of the visual angles subtended by the nearer and farther interval for each subject separately. The reason for treating subjects separately is that interval sizes vary somewhat with eye height, as can be seen in Fig. 7 (e.g., compare farther interval sizes for GK and JA, the tallest

FIGURE 8. Ratios of actual and predicted intervals vs farther interval size. Predictions are based on optic flow using expansion of the farther interval. Note that the scale of the ordinates is smaller than in Fig. 7.
and shortest subject; see Table 3 for eye heights). It is clear from Fig. 7 that subjects did not simply equate the visual angles of the intervals.

To evaluate how well the optic flow metric can predict the size of the nearer interval, ratios of the actual and predicted nearer interval size, $\rho_{pe}$, are plotted vs farther interval size ($z_{23}$) in Fig. 8. A regression on the ratios from all subjects yields $\rho_{pe} = 0.94 - 0.2 \cdot 10^{-3}z_{23}$. ($R^2 = 0.008$, $F(1,112) = 0.93$, $p(F) = 0.337$) for the optic flow metric using image velocities (see subsection: Collinear, sagittal intervals). Using expansion, $\rho_{pe} = 1.04 - 0.1 \cdot 10^{-3}z_{23}$, ($R^2 = 0.0003$, $F(1,112) = 0.35$, $p(F) = 0.555$). As neither ratio was significantly affected by the size of $z_{23}$ in any of the subjects, a one-way analysis of variance between the two groups was performed (that is, all the 8 [subjects] × 14 [conditions] = 112 ratios derived from the optic flow “expansion” prediction were combined into one group and the ratios from the optic flow “velocities” were combined in another group). The mean ratio, $\rho_{pe}$, of the optic flow “velocities” group was 0.91 ± 0.02 (SEM); for the “expansion” group, $\rho_{pe} = 1.02 ± 0.02$; the two means were significantly different.

In summary, a metric based on the expansion of intervals due to observer-induced optic flow can predict to within 2% (or one standard error) of the mean when two collinear, sagittal intervals will appear to have the same 3-D size.

Discussion

Although the optic flow metric predicts the matching data well, other cues may do the same. Two obvious candidates that can yield quantitative predictions are binocular parallax (or binocular convergence) and binocular disparity. In both cases, however, it will be clear that one has to resort to ad hoc criteria for matching spatial intervals in terms of binocular parallax or disparity.

The binocular parallax, $\gamma$, of target $T_1$ is approximately $1/d'$, where $d'$ is the distance from the subject’s eye position to target $T_1$, and $d'$ is the interpupillary distance (taken to be 6.3 cm). As the closest target was 2 m in front of the subject and the smallest eye height was 1.50 m, the maximum binocular parallax in this experiment was $\gamma = 1.49$ deg. Parallaxes for $T_2$ ranged from 1.11 deg (3 m) to 0.19 deg (20 m) for an eye height of 1.50 cm.

The two sagittal intervals $T_1T_2$ and $T_2T_3$ could be matched by equating the ratios of the parallaxes associated with each interval: $\gamma_2/\gamma_1 = \gamma_3/\gamma_2$, which is to say $d_2'/d_1' = d_3'/d_2'$. An equivalent monocular cue is to compare the visual angles subtended by the targets themselves. Assuming the targets have some physically identical dimension (say width $w_1$), the visual angles, $\alpha_i$, subtended by that dimension can be compared: $\alpha_1/\alpha_1 = \alpha_2/\alpha_2$, so that again $d_2'/d_1' = d_3'/d_2'$. This matching scheme predicts the data fairly well but not as well as the optic flow metric. A regression on the combined data from all subjects yields a ratio of actual to predicted near intervals as a function of farther interval size of $\rho_{pe} = 1.21 - 0.3 \cdot 10^{-3}z_{23}$, ($R^2 = 0.017$, $F(1,112) = 1.89$, $p(F) = 0.172$). As with the optic flow metric, interval size ($z_{23}$) does not significantly affect the ratio; averaged over all subjects and conditions, the mean ratio is $1.16 ± 0.02$ (SEM), which is significantly different from the corresponding optic flow values ($\rho_{pe} = 0.91 ± 0.02$ and $\rho_{pe} = 1.02 ± 0.02$; all means are different at $P < 0.0001$ level). This scheme is of course hypothetical and leaves aside the question whether binocular parallax is a sufficiently reliable and strong signal, in particular in far visual space. The experimental evidence suggests that it is not. Gogel (1977) found that two points in a dark room, one at 1.5 m and one at 7.5 m, having binocular parallaxes of 4.6 and 0.9 deg, respectively, could be assigned the correct depth order. But points at 3 and 6 m, with parallaxes of 2.3 and 1.15 deg, could not be correctly ordered in depth. Since the differences in parallax in the current experiment were less than 1.3 deg, it is unlikely that interval matches were based on them. Also note that the standard deviation in the vergence signal of a moving observer is 0.67 deg (Steinman, Cushman & Martins, 1982), again suggesting that it is unlikely to be a useful cue in far visual space. Finally, Richards and Miller (1969) found that 10 out of 25 subjects did not use convergence information in making depth judgments in near space (<2 m).

The binocular disparity, $\delta$, of two targets $T_1$ and $T_2$ that are farther away than a few meters is approximated closely by their difference in binocular parallax, $\delta = \gamma_1 - \gamma_2 = 1/d_1' - 1/d_2' = 1/\Delta d/d_1'd_2'$. Disparities in this experiment were small, ranging from a minimum of 0.03 deg with $T_2 = 20$ m and $T_3 = 24$ m to a maximum of 0.22 deg (13') with $T_2 = 3$ m and $T_3 = 4$ m, with a mean of 0.09 deg (5.3') averaged over the 14 configurations. Interval matches were not made by simply equating the disparities of the closer and farther intervals. This would have led to intervals subtending equal visual angles, which was not found (Fig. 7). The farther of two matching intervals is typically smaller than the closer one, reflecting the basic fact that a fixed disparity corresponds to and is perceived as a larger depth difference as distance increases. This suggests that perhaps matches are based on constant disparity ratios with the ratio varying with perceived distance. But again this was not found empirically; disparity ratios tended to increase with the size of the interval, especially for intervals beyond 5 m (Fig. 7, subjects MF, JB, WA, GK in particular; only ZK’s ratios are fairly constant). Since it is not known how disparities are scaled with the large egocentric distances used in this experiment, it is not possible to decide whether or not matches were based on disparities.

A fundamental problem with using disparities as the primary cue to interval size is the lack of feedback in far visual space. There is no measure to evaluate the interpretation of disparities in terms of depth for distances that are beyond grasping. In contrast, the optic
flow metric is explicitly based on such a mechanism. Perhaps, then, optic flow is the primary cue for depth and is used to calibrate disparities. As mentioned in subsection: Collinear, sagittal intervals, perceived distance, \( d_p \), based on an optic flow metric is approximated closely by a power law: \( d_p = d^n \). On the ground plane \( d = h/\tan \theta \), so that \( d_p = (h/\tan \theta)^n \). Differentiating with respect to \( \theta \) yields:

\[
\frac{\partial d_p}{\partial \theta} = -nh^{-1}(\cos \theta)^{-2}d_p^{n-1} \approx -nh^{-1}d_p^{n-1}.
\]  

Combining this with the above equation for disparity and letting \( n = 1 \) yields:

\[
\delta \approx -ih^{-1}\partial \theta.
\]

Thus, on the ground plane disparity between two points is proportional to their difference in elevation. Substituting this back into equation (11) yields:

\[
\frac{\partial d_p}{\partial \theta} \approx \delta n^{-1}d_p^{n-1} = \delta n^{-1}d_p^{n-1}.
\]  

Perceived depth differences based on binocular disparity approximate the correct geometry as \( n \) approaches 1. Since \( n \) is less than 1, depth differences will be increasingly underestimated with distance, or, conversely, larger disparities are needed for the same perceived depth difference. For example, if \( n = 0.8 \) and distance doubles, \( d_p \) increases by a factor of 3.5 rather than 4; this means that to perceive the same depth difference disparity has to simultaneously increase by a factor of 1.15. Such scalings of disparity have indeed been found in stereopsis (Johnston, 1991).

Direct experimental evidence that image motion calibrates binocular disparity has been provided in a series of perceptual learning experiments by Wallach and his collaborators (Wallach & Karsh, 1963a,b; Wallach, Moore & Davidson, 1963). In the learning period, subjects looked at a small wireframe object, rotating about the vertical at 12 rpm, through a telestereoscope which increased their interpupillary distance by 7.6 cm. Initially, the wireframe appeared to be non-rigid, with its segments perceptually elongating as they approached the sagittal plane during rotation. This distortion disappeared after a few minutes, and after 10 min, judgments of depth with stationary wireframes were reduced by about 30%. Since judgments of depth in a rotating wireframe viewed monocularly did not change, it follows that optic flow stimuli inducing a depth percept (the kinetic-depth-effect) calibrate the mapping of binocular disparities to perceived depth.

**EXPERIMENT 2: EQUATING FRONTAL AND SAGITTAL INTERVALS**

**Materials and methods**

Subjects had to match the 3-D size of a frontal interval \( T_2T_3 \) to that of a given sagittal interval \( T_1T_2 \) (Fig. 9). A system of pulleys allowed subjects to adjust the frontal interval by moving the marker \( T_3 \) along a thin line extending from \( T_2 \) in the direction of \( T_3 \) for a distance of 12 m at approximately 14 cm above the ground. The targets were bright orange rectangles 10 cm high and 7 cm wide. Because of the line and the system of pulleys along \( T_2T_3 \), \( T_2 \) was in a fixed location in the field throughout the experiment and the subject and marker \( T_1 \) were moved back and forth to get the various sagittal intervals. The distance of the subject to marker \( T_1 \) was varied between 1.5 and 15 m and the sagittal interval \( (T_1T_2) \) was varied between 0.5 and 20 m. Table 2 lists all the combinations of distances between the observer and targets \( T_1 \) and \( T_2 \) that were used in the experiment.

The set of sagittal intervals was presented four times and each time the order of presentation was randomized. In half the trials, \( T_2 \) and \( T_3 \) were very close to begin with so that subjects had to increase the frontal interval size. In the other half, \( T_2 \) and \( T_3 \) were very far apart initially so that subjects had to decrease the frontal interval's size. Subjects viewed the markers binocularly and received no feedback about their matching performance. Subjects remained collinear with \( T_1T_2 \) during the experiment.

**Results**

The data from all subjects were expressed and analyzed in terms of the visual angles subtended by the frontal and sagittal intervals from the vantage point of the subject. As the eye height of the subjects ranged from
1.50 to 1.83 m, the sagittal and frontal intervals and the elevation of $T_1$ varies slightly among subjects in subsequent figures.

Figure 10 shows the matching data from a typical subject (ZK). The data from sagittal intervals with the same $T_1$ are connected in the figure, with the numbers indicating the distance of $T_1$ in meters. Note that the data points cluster around a line with a slope of 2, which means that a frontal interval is perceived to delineate the same 3-D interval size as a sagittal interval if it subtends a visual angle about twice as large as the sagittal interval. This follows the prediction derived in section 2.2 [equation (10)]. This relationship holds well if $T_1$ lies between 3 and 10 m in front of the observer. For distances of 10 m and beyond, the matching ratio tends to be larger than two; for closer distances, the ratio is less than two and approaches unity as $T_1$ gets very close to the observer.

To further analyze the dependence of the matching ratios on the distance of $T_1$ and the size of the sagittal interval, the data were plotted as frontal/sagittal matching ratios. The data from subject ZK and those of all other subjects are shown separately in Fig. 11. Clearly, matching ratios vary little with sagittal size for a given distance of $T_1$ and increase with distance to $T_1$. These patterns describe the data from all subjects, with the exception of subject JC, whose matching ratios are inversely proportional to sagittal interval size.

The dependency of the matching ratios on sagittal size was further quantified through linear regression of the matching ratios on sagittal interval size for each $T_1$ distance separately; i.e., $\text{ratio} = a + b\times\text{sagittal size}$. Table 3 shows the results of the regression analysis for each subject. The slope $b$ and its 95% confidence interval $[b_l, b_u]$ (Sokal & Rohlf, 1995) capture the influence of sagittal interval size on the matching ratio. (No confidence interval could be computed for $T_1 = 1.5$ m as there were only two data points—sagittal interval sizes—at that distance.) Slopes are generally small and negative, that is, matching ratios tend to decrease slightly with increasing sagittal size. However, this decrease is often not significantly different from no decrease, i.e., a slope of zero. Specifically, for subject SN none of the slopes differed significantly from 0. For JB only one did; for JA and ZK, only two did; for MF, JC and GK three did, and for WA four slopes differed significantly. Thus, only 18 out of a total of 64 slopes or 28% (or 72 and 25% if the slopes at $T_1 = 1.5$ m, all of which are very close to zero, are included) are significantly different from zero.

Matching ratios at different $T_1$ distances were then computed for a standard sagittal interval size, which is defined as the sagittal interval that minimizes the rms error between the matching ratios computed from the regression lines of Table 3 and the optic flow prediction [equation (8)]. The standard interval size ranged from 1 deg for subject SN to 5 deg for JC and GK (Table 3). The matching ratios increase with distance to $T_1$, as predicted by equation (9) (substitute $\sin \theta_2 \approx h/d_2$, where $d_2$ is the distance of $T_2$).

The standard matching ratios for each subject are plotted as a function of the elevation of $T_1$ in Fig. 12 (filled symbols). It is clear that the data are predicted quite well by optic flow (dotted lines). Small differences between subjects such as SN and GK are captured by the interval sizes over which optic flow is used to form the representation of the ground plane. Figure 13 illustrates how interval size affects the predictions from optic flow;
FIGURE 11. Matching ratios as a function of sagittal interval size (degrees of visual angle) and distance to $T_1$ for all subjects.
TABLE 3.
Matching ratios and their dependence on sagittal size

<table>
<thead>
<tr>
<th></th>
<th>SN (1 deg) 1.55 m</th>
<th></th>
<th>MF (2.5 deg) 1.60 m</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>sem</td>
</tr>
<tr>
<td>T1</td>
<td>b</td>
<td>b1</td>
<td>b2</td>
<td>f/s</td>
<td>sem</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------------------</td>
<td>---</td>
<td>------------------</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>-0.02</td>
<td>---</td>
<td>1.41</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>0.00</td>
<td>-0.11</td>
<td>0.10</td>
<td>1.34</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>3.0</td>
<td>0.00</td>
<td>-0.21</td>
<td>0.21</td>
<td>1.59</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>4.0</td>
<td>-0.02</td>
<td>-0.06</td>
<td>0.01</td>
<td>1.73</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>5.0</td>
<td>-0.03</td>
<td>-0.17</td>
<td>0.10</td>
<td>1.91</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td>8.0</td>
<td>-0.03</td>
<td>-0.16</td>
<td>0.19</td>
<td>2.22</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>10.0</td>
<td>0.01</td>
<td>-0.07</td>
<td>0.09</td>
<td>2.18</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>12.0</td>
<td>0.06</td>
<td>-0.17</td>
<td>0.29</td>
<td>2.20</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>15.0</td>
<td>0.16</td>
<td>-0.80</td>
<td>1.11</td>
<td>2.44</td>
<td>0.09</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>JA (3 deg) 1.50 m</th>
<th></th>
<th>ZK (3.75 deg) 1.52 m</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>b</td>
<td>b1</td>
<td>b2</td>
<td>f/s</td>
<td>sem</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------------------</td>
<td>---</td>
<td>------------------</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>-0.02</td>
<td>---</td>
<td>1.48</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>0.00</td>
<td>-0.22</td>
<td>0.21</td>
<td>1.51</td>
<td>0.13</td>
<td></td>
</tr>
<tr>
<td>3.0</td>
<td>-0.03</td>
<td>-0.08</td>
<td>0.03</td>
<td>1.79</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>4.0</td>
<td>-0.04</td>
<td>-0.10</td>
<td>0.02</td>
<td>2.02</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>5.0</td>
<td>-0.07</td>
<td>-0.08</td>
<td>-0.05</td>
<td>2.22</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>8.0</td>
<td>-0.14</td>
<td>-0.27</td>
<td>-0.01</td>
<td>2.45</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td>10.0</td>
<td>-0.29</td>
<td>-1.51</td>
<td>0.93</td>
<td>3.12</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>12.0</td>
<td>-0.16</td>
<td>-1.08</td>
<td>0.77</td>
<td>2.84</td>
<td>0.27</td>
<td></td>
</tr>
<tr>
<td>15.0</td>
<td>0.01</td>
<td>-2.11</td>
<td>2.13</td>
<td>3.19</td>
<td>0.23</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>JB (4 deg) 1.68 m</th>
<th></th>
<th>JC (5 deg) 1.80 m</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>b</td>
<td>b1</td>
<td>b2</td>
<td>f/s</td>
<td>sem</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------------------</td>
<td>---</td>
<td>------------------</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>-0.02</td>
<td>---</td>
<td>1.52</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>0.00</td>
<td>-0.06</td>
<td>0.04</td>
<td>1.56</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>3.0</td>
<td>0.00</td>
<td>-0.13</td>
<td>0.13</td>
<td>1.80</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>4.0</td>
<td>0.04</td>
<td>-0.06</td>
<td>-0.02</td>
<td>1.91</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>5.0</td>
<td>-0.03</td>
<td>-0.08</td>
<td>0.01</td>
<td>1.99</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>8.0</td>
<td>-0.05</td>
<td>-0.11</td>
<td>0.02</td>
<td>2.44</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>10.0</td>
<td>-0.02</td>
<td>-0.28</td>
<td>0.23</td>
<td>2.88</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>12.0</td>
<td>0.11</td>
<td>-0.09</td>
<td>0.31</td>
<td>3.32</td>
<td>0.08</td>
<td></td>
</tr>
<tr>
<td>15.0</td>
<td>-0.05</td>
<td>-0.58</td>
<td>0.48</td>
<td>3.73</td>
<td>0.09</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>WA (4.25 deg) 1.63 m</th>
<th></th>
<th>GK (5 deg) 1.83 m</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>b</td>
<td>b1</td>
<td>b2</td>
<td>f/s</td>
<td>sem</td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------------------</td>
<td>---</td>
<td>------------------</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>1.5</td>
<td>-0.01</td>
<td>---</td>
<td>1.48</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>-0.01</td>
<td>-0.02</td>
<td>0.02</td>
<td>1.54</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>3.0</td>
<td>-0.01</td>
<td>-0.29</td>
<td>0.31</td>
<td>1.87</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>4.0</td>
<td>-0.06</td>
<td>-0.08</td>
<td>-0.04</td>
<td>2.30</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>6.0</td>
<td>-0.11</td>
<td>-0.18</td>
<td>-0.05</td>
<td>2.55</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td>8.0</td>
<td>-0.20</td>
<td>-0.27</td>
<td>-0.13</td>
<td>2.99</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>10.0</td>
<td>-0.27</td>
<td>-0.44</td>
<td>-0.09</td>
<td>3.17</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>12.0</td>
<td>-0.29</td>
<td>-0.73</td>
<td>0.14</td>
<td>3.42</td>
<td>0.21</td>
<td></td>
</tr>
<tr>
<td>15.0</td>
<td>-0.28</td>
<td>-4.96</td>
<td>4.39</td>
<td>3.87</td>
<td>0.86</td>
<td></td>
</tr>
</tbody>
</table>

**T1** = distance to T1 in meters, **b** = slope, **b1** and **b2** are lower and upper bounds of 95% confidence intervals; **f/s** is the predicted matching ratio for the standard sagittal size shown in each table’s header (along with the subject's eye height), and **sem** is the standard error of the mean of the estimate.

For small intervals, the ratios are close to 2 regardless of elevation as predicted by equation (10); for larger intervals and small elevations, equation (9) has to be used, causing the ratios to increase as elevation gets smaller (i.e., distance to intervals increases). Finally, note in Fig. 12 that matches are not at all veridical (dashed curves), except for intervals that are very close (<3 m).

In summary, the matches of frontal and sagittal intervals from eight subjects were predicted well by a model which assumes that the comparison of the spatial intervals is based on relative image motion.

**Discussion**

Frontal–sagittal interval matches have been investigated previously and the main phenomenon, that sagittal intervals are considerably larger in terms of 3-D size than perceptually equal frontal intervals, has been previously noted (Haber, 1985; Levin & Haber, 1993; Loomis et al.,...
FIGURE 12. Standard matching ratios and their standard errors (filled symbols) as a function of the elevation of $t_1$. Open symbols are matching ratios from Fig. 11. Dotted curves are predicted ratios using the standard sagittal interval for each subject. Dashed curves are matching ratios for veridical matches.
1992; Wagner, 1985). The present experiment goes beyond these previous studies by using an apparatus that allowed subjects to perceptually match intervals quickly and reliably without interference from the experimenter. In contrast, Loomis et al. (1992) used experimenters to adjust the frontal interval manually under the direction of subjects. And Haber (1985), Levin and Haber (1993), and Wagner (1985) used magnitude estimation of sagittal and frontal intervals presented at different times. None of the previous studies proposed an explanation for why sagittal intervals appear so much shorter than frontal intervals.

There do not appear to be any other obvious cues for equating sagittal and frontal intervals. Binocular parallax and disparity do not vary along a frontal interval and can therefore not be used to compare the two intervals. Recently, the relationship between sagittal depth intervals and matching frontal intervals has been investigated in stereoscopic displays. Johnston (1991) had subjects adjust the disparity of a textured half-cylinder displayed on a computer screen until its half-height equaled its depth; that is, until its cross-section appeared to be circular. The ratio of the required disparity and the half-height was 0.082 at 53 cm, 0.066 at 107 cm and 0.052 at 214 cm distance. The optic flow metric predicts an asymptotic value of approx. 0.02 [according to equation (12), $\delta = l/2h\theta$, and since matching frontal angles are roughly twice as large as $\theta$, it follows that $\delta/\phi = l/2h \approx 0.02$.] It would be interesting to extend the stereoscopic matching experiment to distances of up to 10 m, to determine if the asymptotic value predicted by optic flow is obtained.

**REFERENCES**


Acknowledgements—I thank Monica Strauss and Susan Naylor for helping conduct the experiments, and Whitman Richards, Ellen Hildreth and the reviewers for their constructive comments and questions.