

Motion-based nearest vector metric for reference frame selection in the perception of motion

Department of Electrical and Computer Engineering and
Center for Neuro-Engineering and Cognitive Science,
University of Houston, Houston, TX, USA
Present address: School of Optometry,
University of California, Berkeley, Berkeley, CA, USA

Mehmet N. Agaoglu



Laboratory of Computational Vision, Psychology,
and Neuroscience Departments, Bilkent University,
Ankara, Turkey

Aaron M. Clarke



Laboratory of Psychophysics, Brain Mind Institute,
Ecole Polytechnique Fédérale de Lausanne,
Lausanne, Switzerland

Michael H. Herzog



Department of Electrical and Computer Engineering and
Center for Neuro-Engineering and Cognitive Science,
University of Houston, Houston, TX, USA

Haluk Öğmen



We investigated how the visual system selects a reference frame for the perception of motion. Two concentric arcs underwent circular motion around the center of the display, where observers fixated. The outer (target) arc's angular velocity profile was modulated by a sine wave midflight whereas the inner (reference) arc moved at a constant angular speed. The task was to report whether the target reversed its direction of motion at any point during its motion. We investigated the effects of spatial and figural factors by systematically varying the radial and angular distances between the arcs, and their relative sizes. We found that the effectiveness of the reference frame decreases with increasing radial- and angular-distance measures. Drastic changes in the relative sizes of the arcs did not influence motion reversal thresholds, suggesting no influence of stimulus form on perceived motion. We also investigated the effect of common velocity by introducing velocity fluctuations to the reference arc as well. We found no effect of whether or not a reference frame has a constant motion. We examined several form- and motion-based metrics, which could potentially unify our findings. We found that a motion-based nearest vector metric can fully account for all the data reported here. These findings suggest that the selection of reference frames for motion processing does not result from a winner-take-all process, but instead, can be explained by

a field whose strength decreases with the distance between the nearest motion vectors regardless of the form of the moving objects.

Introduction

Motion is defined as a change in position over time. Position is defined based on a reference frame (coordinate system) and, hence, the analysis of motion requires a reference frame. In physics, what makes a certain reference frame preferable or more convenient over others is its relevance to the context or its ability to represent phenomena in simpler terms. For instance, a major revolution in astronomy occurred with the shift from the geocentric to the heliocentric model of the solar system. This new reference frame simplified the expressions for planets' motions by eliminating complex epicycles that were introduced to account for irregularities that arose in the geocentric reference frame. As in physics, in order to make sense of the complex motion trajectories of multiple objects and their parts (i.e., to create representations of the environment), the perceptual system needs to choose an appropriate reference frame according to task de-

Citation: Agaoglu, M. N., Clarke, A. M., Herzog, M. H., & Öğmen, H. (2016). Motion-based nearest vector metric for reference frame selection in the perception of motion. *Journal of Vision*, 16(7):14, 1–16, doi:10.1167/16.7.14.

doi: 10.1167/16.7.14

Received February 12, 2016; published May 18, 2016

ISSN 1534-7362



mands. Through the optics of the eye, neighboring elements in the environment stimulate neighboring photoreceptors on the retina, and these retinotopic relations are preserved in early visual cortices (Serenio et al., 1995; Tootell, Silverman, Switkes, & de Valois, 1982). Most theories of vision are based on computations on a retinotopic reference frame and/or make use of features extracted by retinotopically organized receptive fields. However, under normal viewing conditions, retinotopic representations are highly dynamic and unstable due to object and observer movements, which render retinotopically based theories insufficient to explain the clarity and the stability of perception under dynamic conditions (Ögmen, 2007; Ögmen & Herzog, 2010). In fact, many visual processes, which have been previously thought to occur in retinotopic coordinates, have been shown to result from computations in nonretinotopic reference frames (e.g., form: Agaoglu, Herzog, & Ögmen, 2012; Ögmen, Otto, & Herzog, 2006; Otto, Ögmen, & Herzog, 2006; luminance: Shimozaki, Eckstein, & Thomas, 1999; color: Nishida, Watanabe, Kuriki, & Tokimoto, 2007; attention: Boi, Vergeer, Ögmen, & Herzog, 2011; size: Kawabe, 2008; and motion: Boi, Ögmen, Krummehner, Otto, & Herzog, 2009).

Nonretinotopic reference frames for motion perception

Nonretinotopic reference frames are, as the name suggests, those that are not based on retinotopic coordinates. There are many nonretinotopic reference frames available to the visual system (e.g., self-centered ones such as head-centered, etc., and world-centered or space-based ones such as spatiotopic, object-centered, or motion-based reference frames). The reference frames underlying motion perception have been extensively studied both under Gestalt psychology (Duncker, 1929) and ecological perception (Gibson, 1979; Johansson, 1950; Johansson, von Hofsten, & Jansson, 1980). In contrast to physics, the selection of a reference frame for perception cannot be done at will in most cases (for example, while you are moving your eyes try to perceive a stationary object in motion according to a retinotopic reference frame). In addition, the selection of a reference frame is not simply a result of a winner-take-all competition among available reference frames, but multiple reference frames can be used in combination by the perceptual system according to task demands and the relevance of the prevailing context. For instance, several studies showed that the effective reference frame for motion perception can be expressed as a weighted combination of multiple reference frames (Agaoglu, Herzog, & Ögmen, 2015a; Freeman, 2001; Freeman & Banks, 1998; Souman,

Hooge, & Wertheim, 2006; Swanston, Wade, & Day, 1987; Turano & Massof, 2001; Wade & Swanston, 1987; Wertheim, 1994).

As the well-known Gestalt principle of common fate suggests, motion signals influence form processing; objects or elements that share a common motion component are grouped into a single Gestalt. On the other hand, the perceived motion of a stimulus depends also on its own spatiotemporal properties. Motion trajectories and object properties such as elongation, symmetry axes, closest oriented element, spatial frequency, and presentation duration have been shown to affect the perceived direction of motion (e.g., Freeman & Banks, 1998; Löffler & Orbach, 1999, 2001; Magnussen, Orbach, & Löffler, 2013). Moreover, perceived motion of a stimulus is also influenced by the properties of spatiotemporally neighboring stimuli. Karl Duncker (1929) was one of the first scientists to investigate this issue systematically. He found, for instance, that a slowly moving large rectangle induces an illusory motion in the opposite direction for a stationary dot placed inside the rectangle. He explained this “induced motion illusion” by a Gestalt-like principle called the “stationarity tendency of large stimuli.” In other words, larger stimuli tend to be taken as a reference frame. Another interpretation of this finding is that the rectangle provides a natural frame because it is outside of the dot and surrounds it. However, later studies revealed that the inducer object need not be larger than, or surround, the target object to produce the illusion (e.g., Day, 1978; Wallach, 1959). Hence, a reference frame does not have to frame the stimulus in the sense of surrounding and enclosing the stimulus.

Johansson (1950, 1973) proposed the vector decomposition theory according to which the perceptual system decomposes the motion of each element in the display into common and relative components and the common component serves as the reference frame. Whereas several studies supported the general concept of vector decomposition, exceptions wherein the decomposition is imperfect have also been reported (e.g., Hochberg & Fallon, 1976; Shum & Wolford, 1983). More fundamentally, a major shortcoming of the vector decomposition approach is that it is an ill-posed problem in mathematical terms: There are infinitely many ways to decompose a given set of motion vectors into common and relative motion components. However, the percepts form only a small subset of possible solutions. For instance, when presented with Duncker’s (1929) wheel stimuli, where two point lights are attached to the hub and the rim of an otherwise invisible rotating and translating wheel, some observers reported rotational motion for the light at the rim whereas some others reported cycloidal motion (Johansson, 1974; Proffitt & Cutting, 1980;

Proffitt, Cutting, & Stier, 1979; Shum & Wolford, 1983). A “tumbling stick” percept has also been reported (Mori, 1984). Therefore, it is evident from a mathematical point of view that additional information (or constraints) is needed to reduce the number of solutions to the vector decomposition problem. Various constraints (such as minimum information load, minimum relative motion, and zero sum of residual motion vectors) have been proposed to explain how the visual system regularizes vector decomposition (Borjesson & von Hofsten, 1972; Cutting & Proffitt, 1982; Gershman, Jäkel, & Tenenbaum, 2013; Gogel & Koslow, 1972; Hochberg & McAlister, 1953; Proffitt et al., 1979; Restle, 1979). In other words, the constraints introduced in these studies provide heuristics to explain why the visual system selects a particular nonretinotopic reference frame from infinitely many possibilities. According to another approach, a specific vector decomposition may emerge from synergistic computations among neural networks computing, on the one hand, depth, and, on the other hand, figure-ground segmentation (Grossberg, Léveillé, & Versace, 2011).

The field approach to reference frame selection

Recently, taking an alternative approach to vector decomposition, we have proposed that nonretinotopic reference frames emerge from field-like interactions between motion vectors (Agaoglu, Herzog, & Ögmen, 2015b; Noory, Herzog, & Ögmen, 2015). In physics, field theories, such as gravitational and electromagnetic fields, are used to explain distance-dependent action through space without direct physical contact or connectivity. Similarly, in motion perception, the prevailing reference frame emerges in a distance-dependent manner without direct connectivity and physical contact between the inducing elements (e.g., Agaoglu et al., 2015a, 2015b; Gogel, 1974; Gogel & Koslow, 1972; Hochberg & Fallon, 1976; Mori, 1979; Noory et al., 2015; Shum & Wolford, 1983).

The metric for reference frame selection in motion perception

The motion of an object can be computed by tracking the spatiotemporal changes in its form, or independently of its form. The former requires processing of form in advance while the latter does not. In order to understand whether and how different reference frames for motion interact, and what type of metric is used (form-based vs. motion-based), well-controlled spatiotemporal manipulation of motion and form is needed. For instance, the distance between two moving objects can be varied to see the effect of

distance, but it does not help us to differentiate between form-based versus motion-based interactions. Therefore, the goal of this study was to investigate with which metric reference frames for motion interact.

Previous research has examined several factors influencing the perception of linear motion. However, pitting form-based and motion-based metrics against each other using linear motion comes with several confounding factors such as eccentricity and the relative distance between the stimuli (see General methods). For instance, when two objects move linearly with different speeds, the interaction of their motions will be confounded by the distance between them and retinal eccentricity. To eliminate (in some cases to minimize) these confounding factors, in all experiments reported here, two concentric arcs rotating around the center of the display at a fixed eccentricity were presented to the observers for one full cycle. The outer arc was defined as the target arc, and its angular velocity was modulated by a sine wave so as to keep the average distance between the two arcs constant. Depending on the modulation amplitude, the outer arc could slow down, and even briefly change its direction of rotation. The task was to report whether or not the outer arc changed its direction of rotation at any time throughout its motion on the display (see General methods). To illustrate the difference between these two broad categories of metrics, we have considered four potential candidates (Figure 1). The first two of these metrics, namely the object-centered and the object-nearest-contour, are form-based (Figure 1A, B), whereas the other two are motion-based (Figure 1C, D) metrics. The red arrows in each panel in Figure 1 represent the distance defined by the corresponding metric. In order to determine the best metric among these four, we varied the closest contour distance (Experiment 1), and the angular-contour distance (Experiment 2) between the arcs, and their relative sizes (Experiment 3). We have also tested a special scenario where the motions of both arcs were modulated in synchrony but by various amounts, to once again, pit form-based and motion-based metrics against each other. Finally, we present a quantitative model of reference frame selection for motion, which can account for all the findings reported here.

General methods

Participants

Five naive observers and one of the authors (MNA) participated in the study. The age of the participants ranged from 26 to 30 years and all participants had normal or corrected-to-normal vision. The experiments

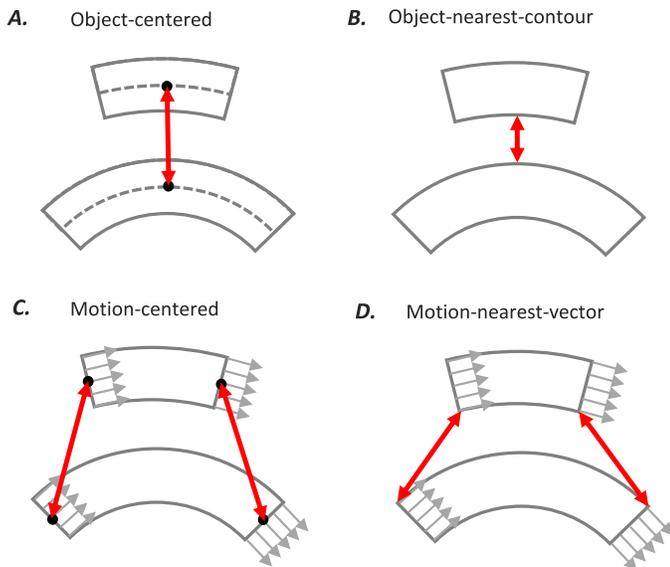


Figure 1. Illustration of various metrics that are considered in this study. (A) Object-centered metric. The center-to-center distance between two moving objects determine the strength of their interaction. (B) Object-nearest contour metric. The closest contour distance between the objects determines the strength of interaction. (C) Motion-centered metric. The average motion vector along a moving contour (or the central motion vector) controls motion interactions. (D) Motion-nearest-vector metric. The distance between the nearest motion vectors determine the way interactions occur. The red double-headed arrows in all panels represent the corresponding metric in each part. In (C) and (D), the gray arrows indicate the motion vectors when the direction of rotation of the arcs is clockwise.

followed a protocol approved by the University of Houston Committee for the Protection of Human Subjects and were in accordance with federal regulations, the ethical principles established by the Belmont Report, and the principles expressed in the Declaration of Helsinki. Each observer gave written informed consent before the experiments.

Apparatus

Visual stimuli were created via a visual stimulus generator card VSG2/5 (Cambridge Research Systems, Rochester, UK) and displayed at a resolution of 800×600 with a refresh rate of 100 Hz on a gamma-corrected Sony GDM-FW900 CRT monitor. Gaze position monitoring for both eyes was performed by means of an Eyelink-II eye-tracker (SR Research, Ottawa, Ontario, Canada) at 250 Hz sampling rate. The distance between observer's eyes and the display was 1 m and the dimensions of the display at this distance

were $22.7 \times 17.0 \text{ deg}^2$. A head/chin rest was used to help stabilize fixation. Observers performed the task via a joystick.

Stimuli and procedures

In all experiments, we investigated how perceived motion of an arc is influenced by another moving concentric arc. Figure 2 shows the spatial and temporal characteristics of the stimuli used. Two white (56 cd/m^2) arcs moving along a circular path around the center of display against a black ($<0.5 \text{ cd/m}^2$) background were used (Figure 2A). The task of the observers was to report if the outer arc (hereafter referred as the target arc) reversed its direction of motion (from clockwise to counter clockwise or vice versa) anytime during its presentation in the trial. The average angular speed of both arcs was $143^\circ/\text{s}$ so that one full cycle of circular motion was completed in 2.5 s (all speeds are expressed in this manuscript in terms of rotational angles per second; rotational angles and visual angles are denoted by $^\circ$ and deg , respectively). Since the distance between stimuli is known to affect reference frame selection (Gogel, 1974; Hochberg & McAlister, 1953; Mack & Herman, 1978; Mori, 1979), by using the same average speed for the two arcs, we kept the average distance between the arcs constant. The centers of the arcs were aligned at the beginning of each trial and the starting position of the arcs along the circular trajectory was selected randomly in each trial. The arcs were presented for only one full cycle of rotation. The direction of motion (i.e., clockwise or counter clockwise) was randomized across trials. An example of velocity profile for the target (outer) arc in a trial, where it rotated clockwise, is given in Figure 2C (the thin red curves). The angular velocity of the target arc was constant during the first and last 630 ms. From 630 to 1890 ms, the velocity of the target arc was modulated by a sine wave. As long as the amplitude of the sine wave is smaller than the average angular velocity of the target, the target might decelerate and accelerate but never moves backward according to a retinotopic/spatiotopic reference frame (e.g., the lightest thin red curve in Figure 2C). However, when the amplitude of the sine wave is chosen to be larger than the average speed, the target stops and reverses its direction for a short amount of time during its motion (e.g., the darkest thin red curve in Figure 2C). The inner arc (hereafter referred as the reference arc) had a constant angular velocity profile in all experiments except Experiment 4, where its angular velocity profile was also modulated with varying amounts in different blocks. We quantified the modulation amount, κ , by the ratio of the amplitude of sine modulation, A , to the mean velocity, w (i.e., $\kappa = A / w$). When $\kappa = 1$, the modulation amplitude equals the mean

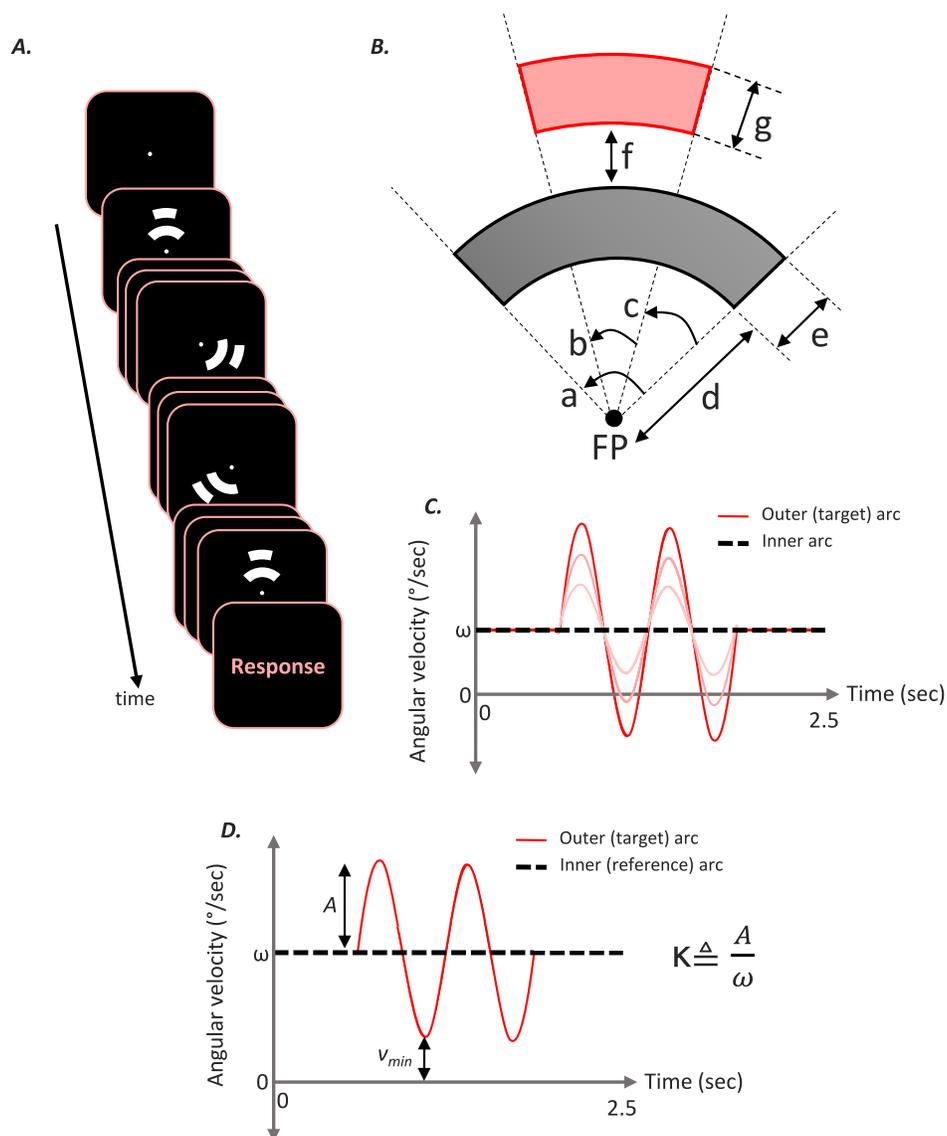


Figure 2. Spatial and temporal characteristics of the stimuli. (A) An example stimulus presentation over time. Observers fixated at a bright point at the center of the display while two concentric arcs rotated around it by one full cycle. The task of the observers was to report whether the outer arc was perceived to reverse its direction of rotation (from clockwise to counterclockwise or vice versa) at any time during its motion. (B) The spatial and figural parameters of the two arcs. (a) The angular size of the inner arc, (b) the angular size of the outer arc, (c) the angular-contour distance between the two arcs, (d) the inner radius of the inner arc, (e) the radial thickness of the inner arc, (f) the radial distance between the closest contours of the two arcs, and (g) the radial thickness of the outer arc. (C) The angular velocity profiles of the two arcs. The dotted black line represents the velocity profile of the inner arc whereas each thin red curve represents the velocity profile of the outer arc with a different velocity modulation factor (κ). (D) Given a specific modulation amplitude A , $v_{min} = \omega - A$ and $K = A / \omega$. See text for detailed explanations.

angular velocity, and there was no modulation when $\kappa = 0$. Experiments were conducted in a normally illuminated room. A small point at the center of the display was provided throughout each trial to maintain proper fixation during stimulus presentation. Trials during which gaze positions of observers deviated more than 2 deg from the fixation point were discarded and repeated immediately.

As soon as both arcs disappeared, observers were asked to report via a joystick whether the target reversed

its direction of motion (from clockwise to counterclockwise or vice versa) anytime during its presentation in the trial. The amplitude of the sine modulation in the target's velocity profile was varied across trials by an adaptive one-up/one-down staircase algorithm (see various thin red lines showing different modulation amplitudes in Figure 2C). For each reversal in observers' responses, the step size in the staircase was halved. Four independent staircases with randomly chosen initial amplitudes were interleaved in a block of trials. Each

	a (°)	b (°)	<c> (°)	d (deg)	f (deg)	(e, g) (deg, deg)	κ
Experiment 1	90	30	30	5.5	1, 4, 7, 10.5	(0.5, 0.5)	0
Experiment 2	15, 45, 90, 180, 270, 360	30	-7.5, 7.5, 30, 60, 120, ring	5.5	1	(0.5, 0.5)	0
Experiment 3	90	30	30	5.9, 5.5, 5.5, 5.5, 5.5, 2.75, 0	1	(0.1, 0.5), (0.5, 0.5), (0.5, 3), (0.5, 6), (0.5, 10), (3.25, 0.5), (6, 0.5)	0
Experiment 4	90	30	30	5.5	1	(0.5, 0.5)	0, 0.1, 0.2, 0.3, 0.5

Table 1. Summary of parameter values used in all experiments.

staircase was completed in 15–35 trials. A staircase was considered “converged” when it underwent 10 reversals and the last eight reversals were used to calculate the threshold for perceiving backward rotation. The minimum velocity of the target corresponding to this threshold amplitude (v_{min} in Figure 2D) was taken as the point of subjective stationarity (PSS). For instance, if the staircases converged to $A = \omega^\circ/s$ for the sinusoidal amplitude modulation, this would correspond to a minimum target velocity of $v_{min} = (\omega - \omega) = 0^\circ/s$. This would mean that backward rotation is perceived only when the target velocity goes below $0^\circ/s$ (“veridical”; i.e., spatiotopic percept), and hence, the PSS would be $0^\circ/sec$. On the other hand, if, for instance, the staircases converged to β°/sec , where $\beta < \omega$, corresponding to the minimum target velocity of $v_{min} = \omega - \beta^\circ/s$, it would mean that as soon as the target velocity fell below $\omega - \beta^\circ/s$, backward motion would be perceived (illusory percept), although it would never move backwards in spatiotopic coordinates. Therefore, the PSS in this case would be $v_{min} = \omega - \beta^\circ/s$. For each unique combination of stimulus parameters (a through g illustrated in Figure 2B and modulation factor for the inner arc in Experiment 4), each observer ran one block of trials (four staircases). Figure 2B illustrates the arcs along with the parameters manipulated in different experiments, and Table 1 summarizes the parameter sets used in all experiments.

In all experiments, we also ran a single block (four staircases) of baseline condition where the outer arc was presented alone. The PSS values obtained in the baseline conditions represent the response bias of observers. Consistent with our previous findings with linear motion stimuli (Agaoglu et al., 2015a), we found a small but statistically significant negative bias in observers’ responses in this study (average across observers $\pm SEM$: $-17.7 \pm 6.8^\circ/s$, one-sample t test: $t(5) = -2.611, p = 0.048$). All PSS values reported here are corrected for observer bias by subtracting the PSS values in the baseline conditions from corresponding effect sizes in each experiment.

Experiment 1: The effect of radial distance

In order to investigate and quantify how distance affects the effective reference frame, we varied the radial distance between the two arcs (f in Figure 2B, the radial distance between the closest contours of the two arcs).

Results and discussion

Table 1 summarizes the parameter values used in this experiment. Figure 3 shows that the baseline-subtracted

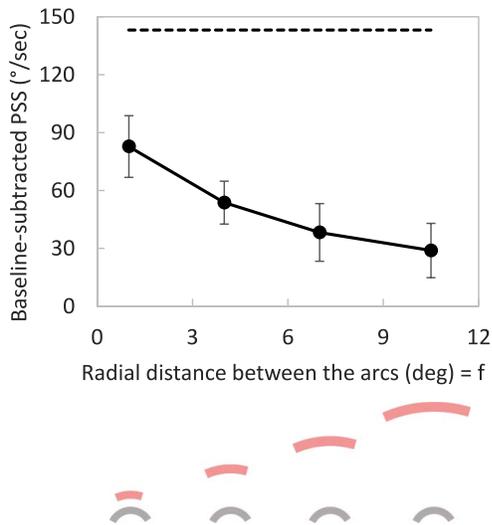


Figure 3. Baseline-subtracted PSS values as a function of the radial distance between the arcs in Experiment 1 (markers). The horizontal dashed line represents the prediction of perfect vector decomposition; if observers base their judgments solely on the relative angular velocity of the target arc with respect to the reference arc. Baseline-subtracted PSS equal to zero corresponds to the prediction of a purely retinotopic/spatiotopic reference frame. The arcs at different conditions are illustrated below the x-axis. The red arcs represent the target arc whereas the gray ones represent the reference arc. Error bars represent \pm SEM across observers ($n = 6$).

PSS values (see General methods) decrease as a function of the radial distance between the rotating arcs. A one-way repeated-measures analysis of variance (ANOVA) showed a significant effect of radial distance, $F(3, 13) = 9.969$, $p = 0.001$, $\eta_p^2 = 0.697$. Note that all data points are above zero, indicating that within the range of distances tested here, percepts were never “veridical” (i.e., they never followed a purely spatiotopic reference frame). Since observers’ eyes were stationary,¹ these results are also inconsistent with a purely retinotopic reference frame. Furthermore, these results are also inconsistent with complete extraction of common motion from the outer arc as predicted by the perceptual vector decomposition theory (Borjesson & von Hofsten, 1975; Johansson, 1950, 1973). According to the vector decomposition theory, the common angular velocity of the two arcs should be perceptually subtracted from the outer target arc, and hence, a slight deceleration in the angular velocity of the target arc (any velocity value below the average velocity, ω) should lead to backward motion percepts. In other words, all PSS values should lie on the horizontal dashed line in Figure 3. However, all data points are well below what is predicted from a perfect common motion extraction point of view.

These results are consistent with previous accounts of distance-dependent effects of moving reference frames (Agaoglu et al., 2015a, 2015b; Gogel, 1974;

Gogel & Koslow, 1972; Hochberg & Fallon, 1976; Mori, 1979; Shum & Wolford, 1983). Previously, we have shown similar distance-dependent effects with a variant of the stimuli used here (Agaoglu et al., 2015a, 2015b). Instead of rotating concentric arcs, we used two horizontally moving disks, one translating with a constant velocity profile whereas the other’s velocity profile was modulated by a sine wave as was the case for the outer arc in the present study. The PSS values showed a linear distance-dependent decrease with horizontally moving disks as well; however, the overall extent to which the extraction of common motion occurs (measured by the ratio of empirical PSS values and those predicted from perfect vector decomposition for the closest spatial separation) was significantly larger than what is reported here (~ 0.85 with translational motion vs. ~ 0.55 with rotational motion). Mori (1984) reported that speed also can influence the selection of the reference frame. However, the quantitative difference between the distance effects in the two studies cannot be explained by different speeds because the average linear speed of the moving elements were roughly the same in these studies. Bertamini and Proffitt (2000) assessed the degree to which different types of motion can serve as a reference frame, and found that translation and divergence are superior to rotation. Hence, a plausible explanation for the quantitative difference between the distance effects could be the ability of the perceptual system to establish reference frames based on translational versus rotational motion in the fronto-parallel plane.

In Experiment 1, we found that the effect of the reference arc’s motion on the perceived motion of the target arc decreases with increasing radial distance. Experiment 1 cannot distinguish between the metrics considered in Figure 1 since an increase in the radial distance between the two arcs results in an increased distance in all four metrics. However, Experiment 1 rules out the accounts based on perfect vector decomposition, and purely retinotopic and purely spatiotopic reference frames. The distance dependence of the reference frame can be viewed as an expression of the Gestalt principle of proximity. In grouping multiple motion vectors so as to extract local reference frames, it is reasonable to assume that Gestalt principles like common fate and proximity also apply to reference frame selection.

Experiment 2: The effect of angular-contour distance

While the contours of the arc with elongations that are parallel to the direction of motion cannot provide a reference frame for that motion, the contours perpen-

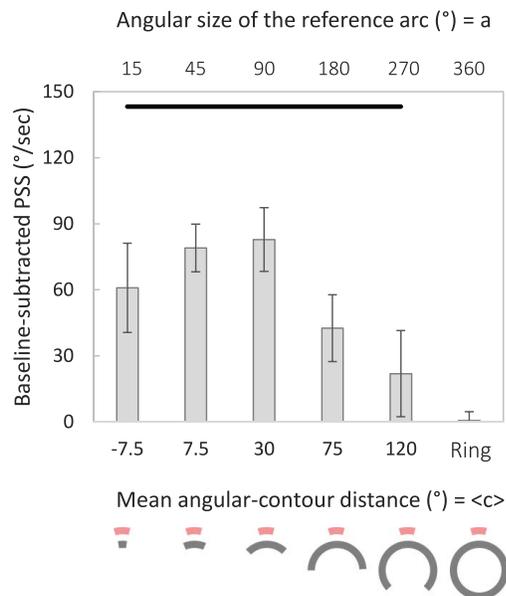


Figure 4. Baseline-subtracted PSS values in Experiment 2 are plotted as a function of mean angular-contour distance. On a secondary x-axis, the angular size of the reference arc is also shown. The horizontal line represents again the predictions of perfect vector decomposition. The arcs at different conditions are illustrated below the primary x-axis. The red arcs represent the target arc whereas the gray ones represent the reference arc. When the angular size of the reference arc is 360° , it becomes a ring and it no longer provides a motion signal. In this case, a zero baseline-subtracted PSS is predicted. Error bars represent \pm SEM across observers ($n = 6$).

pendicular to the direction of motion can. In fact, since the surface of the arc is uniform, the rotational motion information is generated at the leading and trailing contours that are perpendicular to the motion direction. Hence, another way to manipulate the distance between the motion vectors of the reference and the target arcs is to vary the angular distance between the edges of the arcs. In this experiment, the radial distance between the two arcs was kept fixed at 1 deg and the angular-contour distance between the edges (see Figure 2A) was varied systematically. The angular size of the target arc was always 30° , whereas the angular size of the reference arc took one of the following values in a block of trials: 15° , 45° , 90° , 180° , 270° , and 360° . The corresponding angular-contour distances are -7.5 , 7.5 , 30 , 75 , and 120 . When the inner arc's angular span was 360° , it became a ring, therefore there was no angular contour in this case. The parameter values used are summarized in Table 1.

Results and discussion

Figure 4 shows the baseline-subtracted PSS values averaged across observers as a function of mean

angular-contour distance (denoted by $\langle c \rangle$). When the inner arc's angular span is 360° , it becomes a ring. In this case, the rotation of the ring cannot be perceived since its surface is homogenous. In addition, since the contours of the ring are parallel to the direction of motion, they cannot serve as reference and hence the results are, as expected, identical to the baseline condition, yielding a zero baseline-subtracted PSS. Whenever the angular extent of the inner arc was less than 360° , the reference arc appeared to rotate and illusory percepts of direction reversals were perceived, as indicated by positive PSS values in Figure 4.

A one-way repeated-measures ANOVA showed a significant effect of angular-contour distance of the PSS values, $F(5, 25) = 11.718$, $p < 0.001$, $\eta_p^2 = 0.701$, indicating that the effectiveness of the inner arc as a reference frame for the motion of the outer (target) arc was strongly modulated by the changes in angular-contour distance between the two. The condition with $\langle c \rangle = 30^{\circ}$ is identical to the condition in Experiment 1 with $f = 1$ deg, and we found similar PSS values in the two experiments. Increasing $\langle c \rangle$ beyond 30° caused a decrease in the reference frame effect, consistent with the distance-dependent decreases observed in Experiment 1. However, bringing the contours of the reference disk closer to those of the target disk ($\langle c \rangle < 30^{\circ}$) did not cause a further increase in the effect size. Moreover, particularly interesting is the comparison of the two cases when the angular size of the reference arc is smaller than the target arc, and when it is larger (i.e., with radial-contour distances of -7.5° and 7.5° , respectively). Although there is an apparent drop in the effect size when the inner arc is smaller than the target, this difference did not reach significance (paired t test: $t[5] = -1.349$, $p = 0.235$).

If reference frames for motion perception are object-based, changes in the center-to-center, or the radial distance between the arcs should account for the changes in the perceived motion of the target arc. Although these two metrics can explain the results in Experiment 1, they fall short in explaining the data of Experiment 2 in which the angular size of the reference arc is varied while distance according to these two metrics is kept constant. Both the object-centered and the object-nearest contour metrics predict no change in the effect size in this case. However, as our results in Experiment 2 show, that is not the case. On the other hand, as the average angular contour distance increases, the distances defined by motion-based metrics (see Figure 1C, D) also increase, thereby accounting for the results in Experiment 2. However, these results still cannot distinguish between the two motion-based metrics.

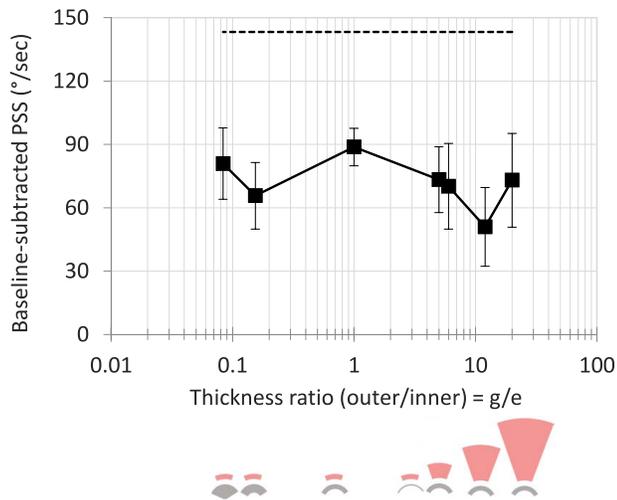


Figure 5. Baseline-subtracted PSS values in Experiment 3 are plotted as a function of the thickness ratio of the two arcs. The horizontal dashed line represents the prediction of perfect vector decomposition. The arcs at different conditions are illustrated below the x-axis. The red arcs represent the target arc whereas the gray ones represent the reference arc. Error bars represent \pm SEM across observers ($n = 6$).

Experiment 3: The effect of radial size

Duncker (1929) proposed a principle called “the stationarity tendency of large stimuli” and suggested that large stimuli tend to serve as a reference. In Experiment 2, we looked at the effect of radial-contour distance between the two arcs on how the target arc’s motion is perceived. A change in radial-contour distance was accompanied by a change in the radial size of the reference arc. In order to investigate the effect of size more directly, we kept the closest contour radial distance and the average angular-contour distances between the arcs constant, and we varied the sizes of the arcs by changing their thickness. This manipulation causes changes in the distance between motion centers (Figure 1C) of the arcs but does not affect the distance between the nearest motion vectors (Figure 1D), allowing us to pit these metrics against each other. Parameters used in this experiment are also summarized in Table 1.

Results and discussion

Figure 5 shows the average baseline-subtracted PSSs as a function of thickness ratio of the two arcs. There is no discernable pattern in the results suggesting that relative size of the arcs does not influence the perceived motion. A one-way repeated-measures ANOVA re-

vealed no significant effect of relative thickness, $F(6, 30) = 1.038$, $p = 0.421$, $\eta_p^2 = 0.172$.

Several studies looked at the effect of size on motion perception during smooth pursuit eye movements. Mateeff, Ehrenstein, and Hohnsbein (1987) showed that when the target object is either small or fast, its retinotopic motion (i.e., motion relative to the pursuit target) is perceived. Increasing the size or reducing the speed of the target object resulted in spatiotopic motion percepts (i.e., motion with respect to the stationary display). They concluded that when the ratio of size and velocity of the target object exceeds 300 ms, spatiotopic motion will be perceived. It is not clear how we can relate these findings to the stimuli in the present study for several reasons. First, in their study, perceived motion judgments were measured during smooth pursuit eye movements. Second, they used dots and disks, which can be described with only one parameter (radius or diameter); however, we used arcs whose shape can be described by at least two parameters (e.g., radius and thickness). Third, whether and how the relative motion between the two arcs can be taken account is not clear. Turano and Heidenreich (1999) also investigated how perceived speed of a distal stimulus changes during smooth pursuit eye movements. Although they did not directly examine the effect of stimulus size, they identified an interaction between the eye movements relative to distal motion and the size of the distal stimulus: When the eyes move in the same direction as the distal stimulus, retinal motion mostly determines the percepts. However, when they are in opposite directions, perceived speed depends on stimulus size. For sizes smaller than 12 deg, perceived speed is overestimated, whereas for larger sizes, perceived speed is underestimated. In short, the size of the stimulus has been shown to affect perceived motion during smooth pursuit eye movements; however, these findings cannot be linked to motion perception during fixation as there are other processes (e.g., efference copy signaling) involved in the former.

Since the arcs underwent rotational motion and since they have homogeneous surface areas, motion vectors are only generated at the leading and trailing contours (Figure 1C, D). The motion-centered distance metric is ruled out by the results of Experiment 3 because the distance between the midpoints of the leading or trailing edges of the arcs changes while the motion reversal thresholds in Experiment 3 do not. The last metric we considered was the motion-nearest-vector metric (Figure 1D). This metric is defined as the distance between the nearest motion vectors of the two rotating arcs (denoted by the red double-headed arrows in Figure 1D). Since the distance defined by this metric does not change with changing relative thicknesses in Experiment 3, this metric predicts no change in effect size here. Therefore, the best metric (among those

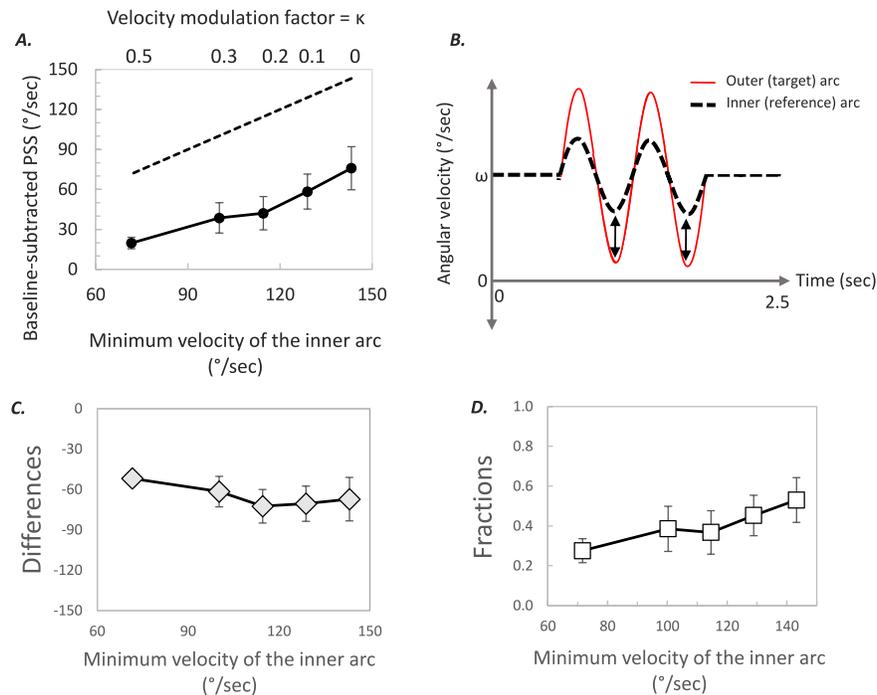


Figure 6. (A) Baseline-subtracted PSS values as a function of the minimum angular velocity of the reference arc. On a secondary x-axis, the corresponding velocity modulation factors (κ) are given. The dashed line represents the prediction of perfect vector decomposition. (B) An example of velocity profiles of the target and reference arcs in Experiment 4. The thin red curve and the dashed black curve represent the velocity profiles of the target and the reference arcs, respectively. Note that within a block of trials, the modulation of the reference arc's motion was fixed. The results in (A) suggest that, as long as the difference between the minimum angular velocities of the two arcs (double-headed arrows) is kept constant, the effectiveness of the inner arc as a reference frame does not change. PSS values in (A) are also presented as a difference from (C), and a fraction of (D) the predictions of perfect vector decomposition. Error bars represent $\pm SEM$ across observers ($n = 6$).

considered here) that can account for all the results presented so far is the motion-nearest-vector metric.

equal so that they decelerated and accelerated with the same time course.

Experiment 4: Constant motion?

The results of Experiment 3 suggest that the figural aspects of stimuli do not play a systematic role in determining the selection of reference frames and the results of Experiments 1 and 2 suggest that the distance with respect to motion vectors has an important influence. The goal of the fourth experiment was to examine further how the motion of the stimuli influences reference frame selection. In Experiment 4, the velocity profile of the inner arc was also modulated by a sine wave (Figure 6B) to determine whether a reference frame is required to have a constant motion. In fact, previous studies suggested that constant motion is more likely to serve as a reference frame (Cutting & Proffitt, 1982; Rubin & Richards, 1988). In different blocks, the amplitude of modulation and correspondingly the minimum velocity was different. Phases of sine wave modulations of the velocity of both arcs were

Results and discussion

Results are given in Figure 6. If common motion of the arcs is extracted perfectly, the target arc should be perceived as reversing its direction of rotation only when its angular velocity goes below that of the inner arc. Therefore, PSS values should be equal to the minimum velocity of the inner arc, as depicted by the dashed line in Figure 6A. However, all PSS values fall well below this line indicating, once again, that common motion extraction is incomplete. A one-way repeated-measures ANOVA showed a significant effect of level of velocity modulation of the inner arc (i.e., its minimum velocity; $F[4, 20] = 9.246$, $p < 0.001$, $\eta_p^2 = 0.649$). These results suggest that in order for the inner arc to serve as a reference frame for the motion of the target arc, it need not have a constant velocity profile: In all levels of velocity modulation used here, PSS values were significantly different from zero, which indicates illusory percepts of rotation reversals. More-

over, the fact that PSS values increase with increasing minimum velocity of the inner arc (i.e., decreasing modulation in its velocity) does show that the reference arc's time-modulated velocity profile is a better determinant of its strength as a reference frame than its average velocity. In fact, a one-way repeated-measures ANOVA on the difference between the empirical PSS values and the predictions of perfect vector decomposition (Figure 6C) revealed no effect of level of velocity modulation of the reference arc, $F(4, 20) = 1.413$, $p = 0.266$, $\eta_p^2 = 0.220$. On the other hand, a one-way repeated-measures ANOVA on the fractions (computed by dividing the PSS values in Figure 6A by corresponding prediction of perfect vector decomposition) revealed a significant effect of velocity modulation, $F(4, 20) = 3.095$, $p = 0.039$, $\eta_p^2 = 0.382$. These results suggest that as long as the difference between the minimum angular velocities of the two arcs (double-headed arrows in Figure 6B) is kept constant, the amount of deviation from perfect vector decomposition remains constant.

After the completion of all experiments, observers were asked to verbally report whether they were aware of the fact that in Experiment 4, the inner arc's velocity profile was also modulated by various amounts. Surprisingly, all five naive observers reported that they were not, which suggests that the inner arc was perceived as rotating at a constant angular velocity. When attention was allocated to the target arc, the reference arc appeared to move with a constant angular velocity while its actual time-varying velocity profile determined how the motion of the target arc was perceived. This observation suggests that the variations in the velocity profile of the reference arc served as a reference for both the target and the reference arc itself. The average angular velocities of the two arcs were equal to each other and remained constant. This common motion was attributed to both arcs, as they were perceived to be rotating with this average velocity. The velocity variations were judged with respect to variations of the reference arc in a distance-dependent manner. Since the distance of the reference arc to itself was zero, the variations of the reference frame matched perfectly the variations of its own velocity and hence it appeared to move at a constant velocity. On the other hand, the target arc being distant from the reference frame, the effect of the reference frame was only partial, resulting in perceived variations in its velocity profile. For the velocity modulation factor values used here, it was difficult to perceive the modulations in the velocity of the reference arc. In a previous study, we have shown that attention modulates the strength of reference frames (Noory et al., 2015); hence, we would predict that with focused attention and large modulation factors, one can also observe the modulations of the reference arc (Demos 3 and 4). In this study we did

not give any specific instructions to the subjects in terms of where they should focus their attention. Our goal was to minimize any a priori bias in the way observers made their judgments. However, it would be also interesting to investigate whether and how attentional effects depend on distance by systematically controlling the focus of attention.

A unifying metric for nonretinotopic reference frames for motion perception

Our results show that the effective reference frame deviates from the prediction of perfect vector decomposition in a distance-dependent manner and the motion-based nearest vector metric provides the best account of this distance dependence. In order to quantify this finding, we computed the deviation from perfect vector decomposition by subtracting the empirical PSS values from those that are predicted from perfect vector decomposition. This allowed us to plot in the same graph the results from Experiments 1, 2, and 3, in which the modulation factor was constant, along with the results of Experiment 4, in which the modulation factor varied. A deviation equal to zero indicates perfect vector decomposition, as shown by the dashed horizontal line at the bottom of Figure 7. This deviation can reach a maximum when motion is perceived according to a retinotopic/spatiotopic reference frame as shown by the dotted horizontal line at the top of Figure 7. In order to assess quantitatively how well the motion-based nearest vector metric accounts for all the data reported here, we plotted the data from all experiments against this metric in Figure 7. A regression analysis revealed a simple linear relationship between the motion-nearest-vector metric and the systematic distance-dependent deviations of the reference-frame from the predictions of perfect vector decomposition.

Modeling

Here, we provide a quantitative account for the reference field theory with motion-nearest-vector metric. We used the same modeling approach as in our previous study (Clarke, Ögmen, & Herzog, in press). In order to model the experimental data, we first created movies of the stimulus for each and every stimulus condition. The first stage of the model consists of filters that extract local motion vectors. In general, any number of motion-extraction algorithms suffice to find the appropriate motion vectors (e.g., Adelson & Bergen, 1985; Simoncelli & Heeger, 1998; Watson & Ahumada, 1985) and the choice of motion vector

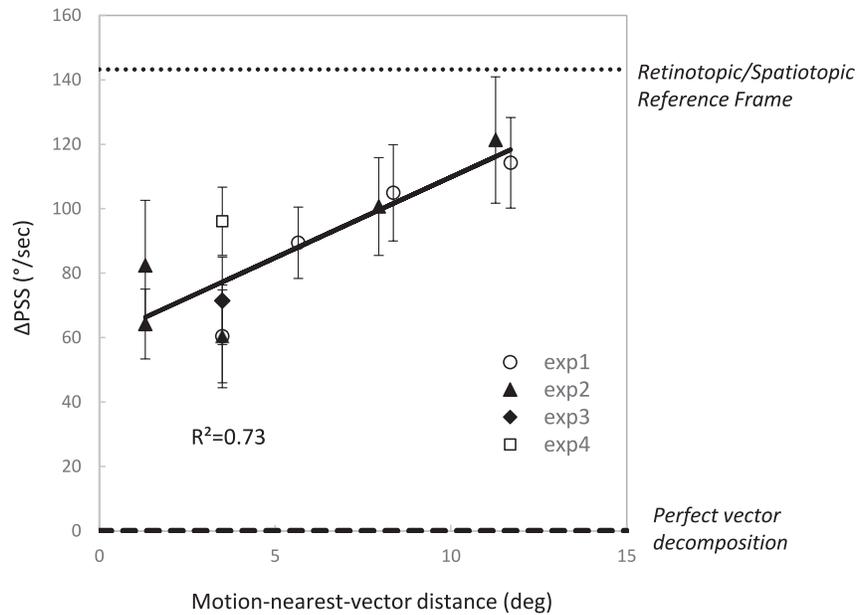


Figure 7. Deviation from perfect vector decomposition, defined as the difference between the prediction of perfect vector decomposition and the baseline-subtracted PSS values, in each experiment are plotted as a function of the motion-based nearest vector metric. Each symbol represents a different experiment. All data points in Experiments 3 and 4 have the same motion-nearest-vector distance and the average of all conditions is plotted as a single data point for each of these two experiments. The solid line represents the linear fit to the data with a coefficient of determination of 0.71. The dotted line represents percepts based on a retinotopic/spatiotopic reference frame, whereas the dashed line indicates perfect vector decomposition. Error bars represent SEM across subjects ($n = 6$).

extraction algorithm is not crucial to model performance. Here, since we generated the stimuli artificially, we know the motion vectors exactly, and to save computational time, we replaced the outputs of a filtering stage with the exact, known motion vectors. Since all of our stimuli involve circular motion, the motion vectors were coded in terms of their angular velocity around the fixation point. Following this stage, a reference field is established around each motion vector for each object. The field follows a Gaussian weighting function of the form:

$$w_{ij} = G \exp(-d_{ij}^2 / 2\sigma^2) + C,$$

where d_{ij} is the Euclidian distance between a pair of motion vectors' spatial locations (i.e.,

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2},$$

where x and y represent the horizontal and vertical coordinates of a motion vector), σ is a constant specifying the spatial extent of each vector's influence, G is a gain factor which accounts for imperfect interactions, and C is a small constant representing the default long-range interactions between motion vectors. Since our results suggest that the motion-nearest-vector distance metric is the best among all four considered, the model computes the distance between each pair of motion vectors at every point in time. The

weight fields of the motion vectors that yield the smallest distance determine the strength of interaction. In quantitative terms, this is equivalent to $r(t) = m_i - w_{ij}m_j$, where $r(t)$ is the perceived motion of the target arc at time t , m_i and m_j are instantaneous motion vectors of the closest points on the target and reference arcs, respectively, and w_{ij} is the instantaneous weight of the velocity field between the two closest motion vectors. The model finds the modulation amplitude that is necessary to get an instantaneous sign change in $r(t)$. This amplitude corresponds to the empirically measured baseline-subtracted PSS. Model fitting was carried out by varying three parameters, G , C , and σ . The best-fitting values were 0.4 for G , 85 pixels (2.41 deg) for σ , and 0.15 for C . The same values of these parameters are used to fit the data from all experimental data in this study. Simulation results are plotted in Figure 8. In short, the reference field theory with a nearest-motion-vector metric provides a good account for all of our results. It should be noted, however, that the model slightly overestimates the effect size for large velocity modulation factors in Experiment 4. Note also that alternatively, the combined influence (through weighted averaging) of all neighboring motion vectors could also be subtracted from the motion of the target arc. However, that would indirectly use the form information, and hence, cannot account for all of our results.

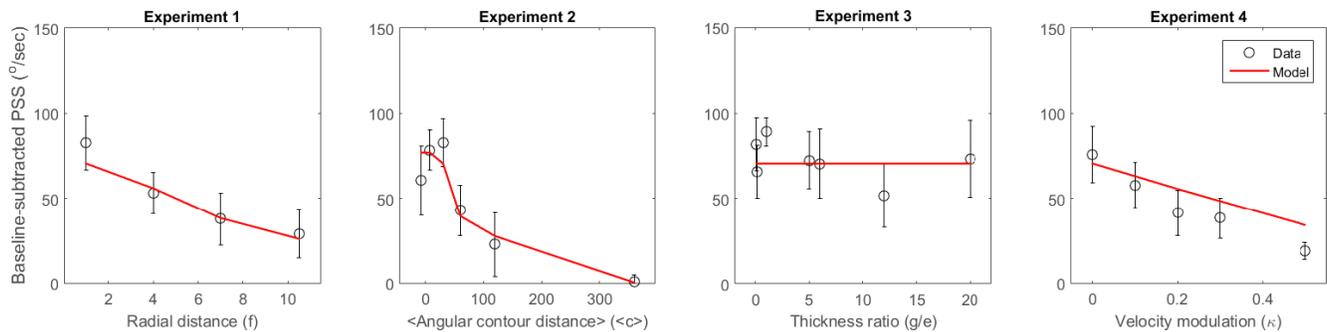


Figure 8. Comparison of the quantitative predictions of the reference-field theory with motion-nearest-vector metric (solid lines) with experimental data (markers) from the four experiments. Note that data from Experiment 4 are plotted here against velocity modulation.

General discussion

There are two broad types of reference frames for perception. Endogenous reference frames are internal to the organism (e.g., retinotopic, head-based, body-based, etc.), whereas exogenous reference frames are external to the organism. In general, our perception is anchored to exogenous reference frames since, in most cases, the perception of our environment remains stable despite the movements of our eyes, head, and body. The early visual system is organized retinotopically and hence the initial coding of visual information is in retinotopic coordinates. Thus, the early retinotopic representations need to be transformed to representations based on exogenous reference frames. In the case of self-generated movements, these necessary transformations can be carried out by means of efference-copy signaling (see reviews: Bridgeman, Van der Heijden, & Velichkovsky, 1994; Wurtz, 2008). The neural motor-planning signals provide the brain with a means to predict retinotopic motions before they occur. Furthermore, the observer's motion generates global and stereotypical retinotopic motion patterns such as translating, expanding optic flow, which can be used to carry out reference frame transformations (e.g., Gibson, 1979; Morrone et al., 2000; Rushton, Bradshaw, & Warren, 2007; Rushton & Warren, 2005; Warren & Rushton, 2009).

A more challenging situation arises in the case of motions of the objects in the environment since the retinotopic changes that occur due to the movements originating from objects external to the observer are neither predictable nor global. The brain has absolutely no information in advance about the changes in retinal motions as a result of motions of external objects. This needs to be computed online in real time. Since each object in the environment may move in a different direction and since motion trajectories can be arbitrarily complex, there must be reference frame selection mechanisms that process visual motion. Here, we

investigated how the visual system selects reference frames for motion perception. In our experimental design, we used rotational motion to eliminate or minimize the confounding effects of eccentricity and distance in the determination of reference frames. With this paradigm, we examined the effects of spatial and figural factors on reference frame selection. The vector decomposition theory has been successful in explaining the selection of exogenous reference frames for motion stimuli, including complex motion configurations as in biological motion displays. However, a shortcoming of this theory is its inability to take into account distance dependency of reference frames. However, if one assumes that observers use linear velocity, instead of angular velocity, vector decomposition approach can account for the results of our Experiments 1, 3, and 4.² Nevertheless, the vector decomposition approach fails to explain the results of Experiment 2, even when linear velocity is used. On the other hand, a motion-based nearest vector metric is able to fully account for all the data reported here. By using the same modeling approach as in our previous study (Clarke et al., in press), we have also provided a computational account for reference-frame selection.

Keywords: reference frames, motion perception, form perception, nonretinotopic, computational modeling

Acknowledgments

Michael Herzog is supported by the Swiss National Science Foundation (SNF) project (320030-153001/1): “Basics of visual processing: from retinotopic encoding to non-retinotopic representations.”

Commercial relationships: None.

Corresponding author: Mehmet Naci Agaoglu.

Email: mnagaoglu@gmail.com.

Address: School of Optometry, University of California, Berkeley, Berkeley, CA, USA.

Footnotes

¹ With the exception of miniature eye movements such as ocular drifts and microsaccades as well as small torsional eye movements.

² We thank an anonymous reviewer for pointing this out.

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*(2), 284–299.
- Agaoglu, M. N., Herzog, M. H., & Ögmen, H. (2012). Non-retinotopic feature processing in the absence of retinotopic spatial layout and the construction of perceptual space from motion. *Vision Research*, *71*, 10–17.
- Agaoglu, M. N., Herzog, M. H., & Ögmen, H. (2015a). The effective reference frame in perceptual judgments of motion direction. *Vision Research*, *107*, 101–112.
- Agaoglu, M. N., Herzog, M. H., & Ögmen, H. (2015b). Field-like interactions between motion-based reference frames. *Attention, Perception & Psychophysics*, *77*(6), 2082–2097.
- Bertamini, M., & Proffitt, D. R. (2000). Hierarchical motion organization in random dot configurations. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(4), 1371–1386.
- Boi, M., Ögmen, H., Krummenacher, J., Otto, T. U., & Herzog, M. H. (2009). A (fascinating) litmus test for human retino- vs. non-retinotopic processing. *Journal of Vision*, *9*(13):5, 1–11, doi:10.1167/9.13.5. [PubMed] [Article]
- Boi, M., Vergeer, M., Ögmen, H., & Herzog, M. H. (2011). Nonretinotopic exogenous attention. *Current Biology: CB*, *21*(20), 1732–1737.
- Borjesson, E., & von Hofsten, C. (1972). Spatial determinants of depth perception in two-dot motion patterns. *Perception & Psychophysics*, *11*(4), 263–268.
- Borjesson, E., & von Hofsten, C. (1975). A vector model for perceived object rotation and translation in space. *Psychological Research*, *38*, 209–230.
- Bridgeman, B., Van der Heijden, A. H. C., & Velichkovsky, B. M. (1994). A theory of visual stability across saccadic eye movements. *Behavioral and Brain Sciences*, *17*(2), 247–293.
- Clarke, A. M., Ögmen, H., & Herzog, M. H. (in press). A computational model for reference-frame synthesis with applications to motion perception. *Vision Research*, in press, doi:10.1016/j.visres.2015.08.018.
- Cutting, J. E., & Proffitt, D. R. (1982). The minimum principle and the perception of absolute, common, and relative motions. *Cognitive Psychology*, *14*(2), 211–246.
- Day, R. (1978). Induced visual movement as non-veridical resolution of displacement ambiguity. *Perception & Psychophysics*, *23*(3), 205–209.
- Duncker, K. (1929). Über induzierte Bewegung. *Psychologische Forschung*, *12*(1), 180–259.
- Freeman, T. C. A. (2001). Transducer models of head-centred motion perception. *Vision Research*, *41*(21), 2741–2755.
- Freeman, T. C. A., & Banks, M. S. (1998). Perceived head-centric speed is affected by both extra-retinal and retinal errors. *Vision Research*, *38*(7), 941–945.
- Gershman, S., Jäkel, F., & Tenenbaum, J. (2013). Bayesian vector analysis and the perception of hierarchical motion. In *Cooperative minds: Social interaction and group dynamics. Proceedings of the 35th Annual conference of the Cognitive Science Society* (pp. 489–494). Berlin, Germany.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- Gogel, W. C. (1974). Relative motion and the adjacency principle. *Quarterly Journal of Experimental Psychology*, *26*(3), 425–437.
- Gogel, W. C., & Koslow, M. (1972). The adjacency principle and induced movement. *Perception & Psychophysics*, *11*(4), 309–314.
- Grossberg, S., Léveillé, J., & Versace, M. (2011). How do object reference frames and motion vector decomposition emerge in laminar cortical circuits? *Attention, Perception & Psychophysics*, *73*(4), 1147–1170.
- Hochberg, J. E., & Fallon, P. (1976). Perceptual analysis of moving patterns. *Science*, *194*(4269), 1081–1083.
- Hochberg, J. E., & McAlister, E. (1953). A quantitative approach, to figural “goodness.” *Journal of Experimental Psychology*, *46*(5), 361.
- Johansson, G. (1950). *Configurations in event perception: an experimental study*. Stockholm: Almqvist & Wiksell.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, *14*(2), 201–211.
- Johansson, G. (1974). Vector analysis in visual

- perception of rolling motion. *Psychologische Forschung*, 36(4), 311–319.
- Johansson, G., von Hofsten, C., & Jansson, G. (1980). Event perception. *Annual Review of Psychology*, 31, 27–63.
- Kawabe, T. (2008). Spatiotemporal feature attribution for the perception of visual size. *Journal of Vision*, 8(8):7, 1–9, doi:10.1167/8.8.7. [PubMed] [Article]
- Löffler, G., & Orbach, H. S. (1999). Computing feature motion without feature detectors: A model for terminator motion without end-stopped cells. *Vision Research*, 39(4), 859–871.
- Löffler, G., & Orbach, H. S. (2001). Anisotropy in judging the absolute direction of motion. *Vision Research*, 41(27), 3677–3692.
- Mack, A., & Herman, E. (1978). The loss of position constancy during pursuit eye movements. *Vision Research*, 18(1), 55–62.
- Magnussen, C. M., Orbach, H. S., & Löffler, G. (2013). Motion trajectories and object properties influence perceived direction of motion. *Vision Research*, 91, 21–35.
- Mateeff, S., Ehrenstein, W., & Hohsbein, J. (1987). Constancy of visual direction requires time to develop. *Perception*, 16(2), 253–253.
- Mori, T. (1979). Relative locations among moving spots and visual vector analysis. *Perceptual and Motor Skills*, 48(2), 587–592.
- Mori, T. (1984). Change of a frame of reference with velocity in visual motion perception. *Perception & Psychophysics*, 35(6), 515–518.
- Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., & Burr, D. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nature Neuroscience*, 3(12), 1322–1328.
- Nishida, S., Watanabe, J., Kuriki, I., & Tokimoto, T. (2007). Human visual system integrates color signals along a motion trajectory. *Current Biology*, 17(4), 366–372.
- Noory, B., Herzog, M. H., & Ögmen, H. (2015). Spatial properties of non-retinotopic reference frames in human vision. *Vision Research*, 113, 44–54.
- Ögmen, H. (2007). A theory of moving form perception: Synergy between masking, perceptual grouping, and motion computation in retinotopic and non-retinotopic representations. *Advances in Cognitive Psychology*, 3(1–2), 67–84.
- Ögmen, H., & Herzog, M. H. (2010). The geometry of visual perception: Retinotopic and non-retinotopic representations in the human visual system. *Proceedings of the Institute of Electrical and Electronics Engineers*, 98(3), 479–492.
- Ögmen, H., Otto, T. U., & Herzog, M. H. (2006). Perceptual grouping induces non-retinotopic feature attribution in human vision. *Vision Research*, 46(19), 3234–3242.
- Otto, T. U., Ögmen, H., & Herzog, M. H. (2006). The flight path of the phoenix—The visible trace of invisible elements in human vision. *Journal of Vision*, 6(10):7, 1079–1086, doi:10.1167/6.10.7. [PubMed] [Article]
- Proffitt, D. R., & Cutting, J. E. (1980). Perceiving the centroid of curvilinearly bounded rolling shapes. *Perception & Psychophysics*, 28(5), 484–487.
- Proffitt, D. R., Cutting, J. E., & Stier, D. M. (1979). Perception of wheel-generated motions. *Journal of Experimental Psychology: Human Perception and Performance*, 5(2), 289–302.
- Restle, F. (1979). Coding theory of the perception of motion configurations. *Psychological Review*, 86(1), 1–24.
- Rubin, J., & Richards, W. A. (1988). Visual perception of moving parts. *Journal of the Optical Society of America A*, 5(12), 2045–2049.
- Rushton, S. K., Bradshaw, M. F., & Warren, P. A. (2007). The pop out of scene-relative object movement against retinal motion due to self-movement. *Cognition*, 105(1), 237–245.
- Rushton, S. K., & Warren, P. A. (2005). Moving observers, relative retinal motion and the detection of object movement. *Current Biology: CB*, 15(14), R542–R543.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., & Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, 268(5212), 889–93.
- Shimozaki, S. S., Eckstein, M., & Thomas, J. P. (1999). The maintenance of apparent luminance of an object. *Journal of Experimental Psychology: Human Perception and Performance*, 25(5), 1433–1453.
- Shum, K., & Wolford, G. (1983). A quantitative study of perceptual vector analysis. *Perception & Psychophysics*, 34(1), 17–24.
- Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research*, 38, 743–761.
- Souman, J. L., Hooge, I. T. C., & Wertheim, A. H. (2006). Frame of reference transformations in motion perception during smooth pursuit eye movements. *Journal of Computational Neuroscience*, 20, 61–76.

- Swanston, M., Wade, N., & Day, R. (1987). The representation of uniform motion in vision. *Perception, 16*(2), 143–159.
- Tootell, R. B., Silverman, M., Switkes, E., & de Valois, R. (1982). Deoxyglucose analysis of retinotopic organization in primate striate cortex. *Science, 218*(4575), 902–904.
- Turano, K. A., & Heidenreich, S. M. (1999). Eye movements affect the perceived speed of visual motion. *Vision Research, 39*(6), 1177–1187.
- Turano, K. A., & Massof, R. W. (2001). Nonlinear contribution of eye velocity to motion perception. *Vision Research, 41*(3), 385–395.
- Wade, N., & Swanston, M. (1987). The representation of nonuniform motion in vision. *Perception, 16*(5), 555–571.
- Wallach, H. (1959). The perception of motion. *Scientific American, 201*(1), 56–60.
- Warren, P. A., & Rushton, S. K. (2009). Optic flow processing for the assessment of object movement during ego movement. *Current Biology: CB, 19*(18), 1555–1560.
- Watson, A. B., & Ahumada, A. J. J. (1985). Model of human visual motion sensing. *Journal of the Optical Society of America A: Optics and Image Science, 2*(2), 322–341.
- Wertheim, A. (1994). Motion perception during self-motion: The direct versus inferential controversy revisited. *Behavioral and Brain Sciences, 17*(2), 293–355.
- Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research, 48*(20), 2070–2089.