Motion: Representation and perception,
SIGGRAPH/SIGART Interdisciplinary Workshop.
New York: Association for Computing Machinery,
1983

Perceiving and recovering structure from events

James E. Cutting
Department of Psychology, Uris Hall
Cornell University
Ithaca, NY 14853 USA

How do perceivers identify a moving object as seen against a changing background? How do figure and ground separate? Such questions have engaged psychologists for at least seventy years. In particular, the Gestalt psychologists were deeply concerned with the latter, but had only the ill-defined notion of common fate, or uniform density, for dealing with the former. The coherent flow of a moving object is seen, somehow, by extracting those aspects of the whole that segregate it from the ground; the uniform destiny of all parts of the object was thought both to make the whole cohere and to separate the whole from all else. Two pairs of ideas, from two researchers who came out of the Gestalt tradition, helped elucidate the notion of common fate as applied to motion perception.

The first pair of ideas is due to Johansson (1950): The motion of an object can be parsed into common motion and various relative motions. The common motion of elements in a visual display is the vector path shared by all elements; the relative motions of elements are the residuals, moving with respect to the whole. Johansson was quick to see, as others have subsequently (Cutting & Proffitt, 1982), that some decoding principles were needed for a human being or a machine to be able to recover structure from the motion of parts of a visual display. The problem is seen most clearly when we introduce an equation of three terms:

Absolute motion = common motion + relative motion

The new term, absolute motion, is the vector path over time and through space for any given part of a dynamic display without regard to other parts. This motion, of course, is generated by the motion common to the whole and the motion of the particular part in relation to the whole. To be more concrete, let me choose an example used by Johansson (1973) and by Duncker (1929/1938) before him, as shown in Figure 1a. Two lights are mounted on a rolling wheel and all else darkened. One light is mounted at the rim of the wheel and the other at the axle. The absolute motion paths are shown in the top panel of the figure, and two possible interpretations at the bottom. Which of these interpretations is seen depends somewhat on the viewing conditions (Cutting & Proffitt, 1982; Börjesson & von Hofsten, 1975), but the problem is general: The observer could see a smoothly rolling wheel if the motion of the axle-mounted light is seen as the center of the configuration, or she could see a stick with lights mounted on its ends rotating about an unseen center and the stick tumbling through space following the path of a prolate cycloid. The general problem, pointed out by the multiplicity of possible percepts, is one of induction: How does one obtain two component sets of motions when only the absolute motion is given? A priori, there is an indefinitely large set of common motion/relative motion pairs that could, when vectorially combined, yield the same absolute motion. The visual system either needs some rules of thumb to guide choice among possible alternatives in this situation and in others, or it needs some design feature inherent in its construction that constrains possibilities.

The second set of ideas is due to Wallach (1965/1976), and they express essentially the same idea in a different way: The motion of objects can be thought of in terms of two types of displacements, those relative to the object itself and those relative to the observer. It seems most efficacious to discuss these in terms of coordinate systems, and many will immediately recognize these ideas as close to those of Marr (1982). Object-relative displacements are motions or changes of a part of an object with respect to the whole of the object. The origin of this coordinate system is the "center" of the object, and is the origin of Marr's object-centered coordinate system. Of course, the critical issue is to find this center of this coordinate system, and most of this paper is focussed precisely on that point. Observer-relative displacements in Wallach's terms are measured in Marr's viewer-centered coordinate system. In a sketchy way, these ideas can now be applied to the lower two panels of Figure 1. If the observer perceives a rolling wheel from this two-light configuration, the origin of the object-centered system is at Light B. In turn, this coordinate system rotates relative to the observer and follows the translatory vector path of common motion through viewer-centered coordinates. If, on the other hand, the observer perceives the configuration as a tumbling stick then the origin of the system is midway between Lights A and B and the object-coordinate system rotates as it follows the prolate cycloidal path of translation through viewer coordinates.

With this set of interrelated ideas in mind--common and relative motions and observer-relative and object-relative displacements--we can now consider six different types of events where underlying structure is perceived and recovered through motions or displacements of elements in the display.

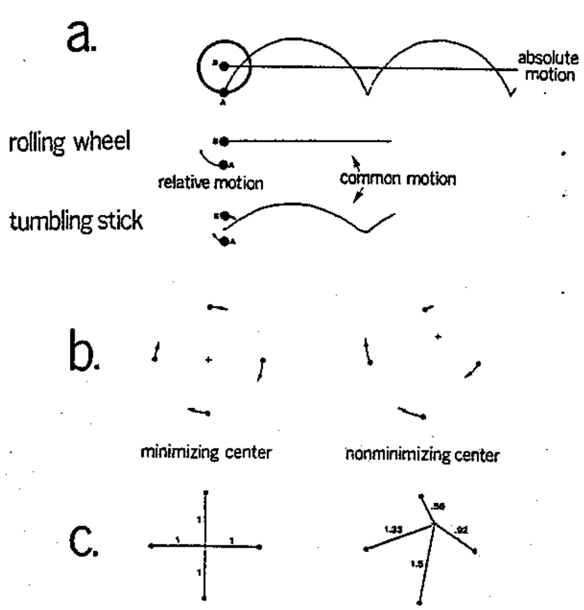

Figure 1. Panel a shows the absolute motion paths of two lights mounted on a rolling wheel, one at the perimeter (Light A) and one at the hub (Light B). Two interpretations of this stimulus are also given, one as a rolling wheel and the other as a tumbling stick. Relative and common motions are indicated for each of these interpretations. What is taken as the center of moment determines perception. Panels b and c show four points in rotation for demonstration of how a minimum principle might apply in extracting relative motions. In Panel b, if these lights are seen to rotate around their centroid, then all momentary vectors sum to zero. If, on the other hand, they are seen to rotate around some other point they will not sum to zero. Instead they will sum to a clockwise rotational vector whose radius is the distance from the centroid. In order to perceive such a configuration, a common vector of the same magnitude but rotating counterclockwise would be needed. The configuration on the left is invariably seen. Panel c demonstrates a second, interrelated minimizng procedure. Rotation around a centroid minimizes the squared lengths of moment arms from the radius to each light. On the left, the summed squares of moment arms is 4.0, and on the right it is 5.18. (After Cutting & Proffitt, 1982).

Event 1: Rolling wheels

The oldest phenomenon of real motion (as opposed to apparent motion) that has come under systematic scrutiny by psychologists is that of rolling motion. What is shown to observers are the movements of a few scattered lights mounted on an unseen wheel. Rubin (1927), following the lead of Galileo and some interests of 17th century mathematicians, noted that the absolute motion paths of such lights (which were always members of the family of cycloids) were rarely seen as such. Instead, viewers had a strong tendency to see objects undergoing rotation (relative motion) and translation (common motion). This finding had strong influence on Gestalt psychologists as demonstrating that perceivers organized noncontiguous elements into coherent wholes. But Rubin and subsequent researchers in the Gestalt tradition (Duncker, 1929/1938; Johansson, 1950, 1973; Wallach, 1965/1976) were always interested in particular stimuli as they elicited particular percepts. In contrast, my colleagues and I have been interested in the whole population of stimuli and in what people generally saw. Such effects cannot be measured through demonstration, but only through experimentation (Proffitt, Cutting & Stier, 1979; Proffitt & Cutting, 1979, 1980; Cutting & Proffitt, 1982).

The results have been remarkably clear-cut and promote a process-orientation as to how the visual system extracts information from these dynamic displays. Stimuli have generally consisted of two to four lights mounted in various places on the wheel--at the rim, within the interior and including the axle, and even exterior to the rim. Although such factors as the number of lights and the symmetry of their arrangement play some role in perception, these are swamped by a separate factor: What is seen by most observers at most times is dictated by the distance of the centroid of the configuration of lights from the wheel's axle. That is, if the center of the configuration is at the axle, then that configuration will look very much like a rolling wheel. If, on the other hand, the center of the configuration is far from the axle (say, three lights mounted closely together near the rim) then it will generally not look like a rolling wheel, but instead will look more like a hobbling or hopping object. Moreover, the degree to which the configuration looks like a rolling wheel is determined by the relative distance between the centroid of the light pattern and the axle. Such a simple notion accounts for about 90% of the variance in all observers' responses across a dozen experiments with over two dozen stimuli.

The process of perceiving these stimuli seems to be as follows: The viewer first minimizes the relative motions of the lights with one another. This minimization process can occur by either of two methods, as shown in Figures 1b and 1c. Imagine these four lights in rotation. What unseen axis will they appear to rotate around? In object-centered terms the rotation will be seen to take place around the centroid because either (1) this is the only point at which the momentary vectors of all points will add to zero, or (2) this is the point at which the length of the moment arms generating the movements are minimal. No other point satisfies these two interlocked minimization procedures. Once the relative motions have been extracted, then the common motion falls out as residual.

Such a procedure--first extract relative motions according to a minimum principle, then observe the residual common motion--is an adequate way to solve the equation given earlier, at least in many situations. It appears that for such wheel-generated motions, this procedure is used most of the time. There are, however, some exceptions. I will consider several later in the discussion of other events, but there is also one here, the rolling-wheel interpretation of the configuration shown in Figure 1. But this is one of the few failures in our object-parsing system, and its perception may be due to an alternative minimization procedure (Cutting & Proffitt, 1982). Moreover, this rolling-wheel percept disappears when external reference frames are removed (Börjession & von Hofsten, 1975).

Let me now introduce a neologism. We have called the point around which these motions appear to take place, the point that also nullifies relative motions in these displays and which minimizes moment arms, the center of moment. Borrowing from Dürer (1528/1972) and capitalizing on the dual meanings of motion and importance, we claim that the center of moment allows the perceiver to recover the structure of the stimulus and to perceive the event. This point is the origin of the object-centered coordinate system. Its own motion through observer-relative coordinates dictates how wheel-like the event will be perceived. If the center of moment (in this case the centroid) moves smoothly and linearly then the observer will see a wheel-like event; if if moves up and down in its translation (following a prolate cycloid) then it will be seen as less wheel-like.

This first event is one in which all elements remain in rigid relation to one another. The second, although it is also basically a rigid stimulus, has many nonrigid relations among its various parts.

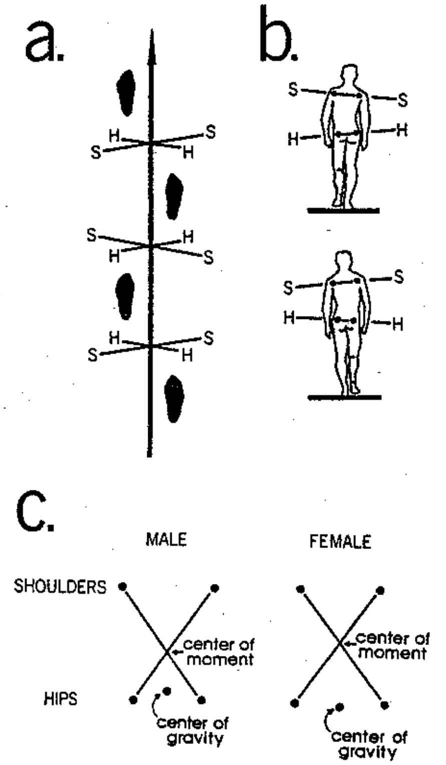

Figure 2. Panel a shows four footfalls of a walker and the relative torsion in the horizontal plane of shoulders against hips in each of the three double support phases. Panel b shows the same torsion in the frontal plane. Panel c draws stress lines across the diagonals of the torso to demonstrate the locales of the center of moment for a male and a female walker. The center of moment is not the center of gravity, also shown. The location of the center of moment in dynamic, computer-generated displays determine the perceived gender of the walker. (After Cutting, Proffitt, & Kozlowski, 1978).

Event 2: Walking People

Johansson (1973, 1976), using a variant of a technique from 19th century photography (Marey, 1895/1972), mounted lights on the joints of an individual, darkened the surround, and asked observers what they saw. Invariably, observers untutored in looking at such displays immediately recognized the presence of a human being. Moreover, they recognized the activities of the individual--walking, running, doing jumping jacks--with as little as 100 msec exposure to the dynamic displays. It is difficult to overestimate the impact and perceptual salience of such configurations; they yield remarkably convincing percepts of people and do so only with an arrangement of a dozen lights or less moving against a dark ground. One is never confused about which lights are dynamically connected to which others: a light at the wrist is seen as connected to the light at the shoulder, and not the light at the hip, even though it is closer to the hip than the shoulder. An adjacency principle (Gogel, 1978) cannot hold in such situations.

My colleagues and I became interested in such displays. We first demonstrated that friends could identify themselves and one another from such information when they walked laterally across the field of view (Cutting & Kozlowski, 1977). We then found that new observers unfamiliar with the particular displays could identify the gender of the walker (Kozlowski & Cutting, 1977; Barclay, Cutting & Kozlowski, 1978). The former result, although probably the most interesting that we have, seemed difficult to pursue: Person identity is wrapped up in too many vaguely known factors about personality and mood. The latter result, then, is the one that engaged us (Cutting, Proffitt & Kozlowski, 1978). The question is: On what informational basis do perceivers make their judgments of gender? After many false starts--finding that such factors as walking speed, step size, and arm swing were all sufficient for gender recognition but not necessary--we began a biomechanical analysis. Ultimately, this is only a little more sophisticated than the idea of "the knee bone is connected to the ankle bone": It is that the human form is a hierarchy of related elements moving through space in an economical way. The economy lies in the pendular motions of the arms and legs and in the dynamically crossed symmetry of arms and legs--left arm moves in synchrony with right leg, right arm with left leg. Everything is either in phase or 180° out of phase. As arms and legs move, so move shoulders and hips. The general arrangement can be seen in Figure 2a.

Since the shoulders move against the hips, the torso, in faithfulness to its etymology, undergoes torsion. This torsion can be likened to that of a flat spring. If stress lines are drawn across the diagonals, as shown in Figure 2c, one has in the intersection a fairly good approximation to the center of moment of a walker. All points in the torso and all points in the arms and legs can be thought of as moving around this point. It is the origin of the object-centered coordinates for the perceiver.

What was interesting to us is that the locus of this point is generally different for males and females. That is, because males have slightly wider shoulders than hips, and because in females these dimensions are roughly the same, male and female torsos differ in their style of torsion. In normal gait, the most ergonomically efficient mode of locomotion, the amounts of torsion and pendular motions of the limbs are constrained (Beckett & Chang, 1969; Murray, 1967). Because of the difference in origin of movements, males and females have many systematic differences in their walks: females swing their arms more, males their shoulders more (see Figure 2b), both in accomplishing the same end of countering the forces generated in the legs and maintaining balance. Because of their generally wider hip girdle, women rotate their hips more and walk more smoothly across a surface. Males, in contrast, tend to bob up and down more when they walk.

What excited us most, however, was the fact that the difference in locus of the centers of movement for males and females accounted for roughly 75% of the variance in gender judgments. That is, as determined by anthropometric procedures adapted to this study (Cutting et al., 1978), males generally had lower centers of moment than females, and the individuals who were systematically misidentified as either female or male had locations of center of moment that more nearly approximated those of the opposite sex than of their same sex.

With such results in hand I set out to computer-synthesize human gait. My purpose was to hold all else constant in the displays except the location of the center of moment and those differences that it generated. The outcome was quite successful (Cutting, 1978a, 1978c). In fact, I was able to generate hypernormal synthetic males and females; they were more often identified as male or female than any of the real people that I had previously videotaped.

Our account of how human observers accomplish the perceptual feat of extracting the information from these dynamic displays is, in its essence, the same as the account for the perception of rolling wheels. Observers appear to extract the relative motions of the lights from one another by some minimum principle that assumes that

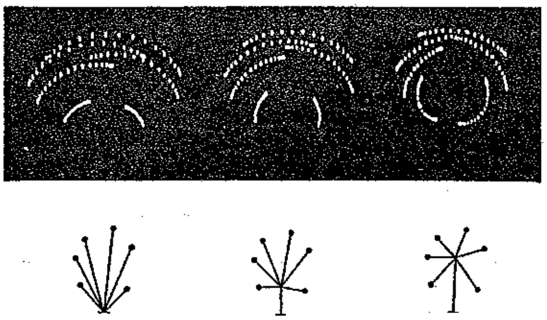

Figure 3. Static representations of three tree- and bush-like stimuli used by Cutting (1982a). At the top are superimpositions of 24 frames in each dynamic stimuli; and at the bottom are stick-figure versions of the same three stimuli. First-order centers of moment are located where the trunks meet the ground; second-order centers are where the limbs meet the trunk. Perceivers can determine the arborization pattern of a bush or tree quite accurately, appearing to use information about centers of moment.

rigid pendular motions are simpler than the various nonrigid comparisons that could be made. Once isolated, the common motion of the whole--an undulating motion with linear translation--falls out as the residual. Evidence for this view is not as solid as in the case of rolling wheels; about all we know is that distorting the common motion of a walker has little effect on perception (Johansson, 1973), but that distorting the relative motions has great effect (Barclay *et al.*, 1978; Cutting, 1981). If extraction of relative *motion* were logically dependent extraction of common motions, then a different pattern of results ought to have accrued.

Event 1 was a mechanical event (rolling wheels) and Event 2 biomechanical. The former involved rigid and circular relative motions and the latter rigid and nested pendular relative motions, with nonrigid torsion. For purposes of generality I looked at a third mechanical event with somewhat more complex relative motion, one that is completely nonrigid.

Event 3: Swaying trees and bushes

Most of us have seen the following event at a suburban shopping mall. At year's end before the holidays, small lights are placed among the limbs of deciduous trees. When the wind blows, the trees are set in motion, and the pattern of motion of the lights is fairly rich in revealing the limb structure of the tree. Since this so nearly approximated the experimental situations used in the events discussed previously, I decided to extend the technique to study the perception of underlying arborization patterns in the motions of trees and bushes (Cutting, 1982a).

Locations of lights were varied, and the pattern in which the limbs intersected the trunk was also varied. Over a sequence of frames, as before, the displays oscillated back and forth as if the tree or bush were blown by the wind. Here, all motions were not strictly inverse pendular. Instead, all branches and limbs were flexible, and motions slightly phase-staggered. Thus, unlike the displays of the previous two events, there were no rigid relations among the moving lights. Motion depended on the length of the unseen limb (from trunk to light) and on the unseen angle with which the unseen limb intersected the unseen trunk. Static renditions of such displays are shown in Figure 3.

Results of several experiments suggested that viewers are highly attuned to structural information of this kind. For highly stylized trees with six limbs branching from the same locale, observers could not only make systematic comparisons among members of the set of stimuli, but they could also discern the location of the unseen node of all limbs within about a half a degree of visual angle (about 5% of the stimulus height). For less stylized trees, those with six limbs branching from several different locations on the trunk, observers could still make systematic judgments about the relative similarity of the trees or bushes based on the arborization patterns implied by the moving lights on the ends of the limbs.

In this experimental situation there is no common motion; trees and bushes do not uproot themselves and move laterally across the field of view. Thus, no statement can be made for this type of display about priority of relative and common motions or parsing of object and observer coordinate systems. The observer, of course, could move relative to the display (the tree or bush) in the real world. I suspect our same scheme would apply and the motion-segregation problem would be invoked. But the displays here had no such variation. Instead, the displays present differentially nested patterns of relative motions, and the results suggest that observers can parse this structure in order to perceive a coherent whole. Since all motion in these displays is referred to the point at which the trunk meets the ground, but only through the motions of the limbs and the manner and height at which they branch from the trunk, the stimuli represent a complex hierarchy of *nonrigid* motions. Yet the visual system seems unperturbed by such complexity, and seems to handle such information as well as it handled the complex hierarchy of *rigid* motions in the walker displays. Any account that assumes that the visual system filters out rigidity in order to perceive objects (e.g. Ullman, 1979) must eventually try to handle such data.

The three types of events considered thus far are all events that occur in real time at a relatively rapid rate. The focus of all of these is a center of moment, the nonarbitrary origin of an object-centered coordinate system around which relative motion takes place for the object as a whole. In the fourth event I will not consider motion at all, but nonrigid change as it takes place over the span of a human lifetime.

Event 4: Aging faces

As we mature our heads and faces change. Orthodontists and surgeons trying to correct cleft palate realize that they must deal with these plastic changes over time: Growth and change must be built into the surgical correction (Todd & Mark, 1981). Much of the change that the human head undergoes can be captured by a relatively simple algorithm--a cardiodial transformation (Todd, Mark, Shaw & Pittenger, 1980). This transformation captures head growth and change, in part, because it mimics the gravitational tugs on tissues.

What interests me about the similarity between growth and this algorithm is that growth changes (Enlow, 1975) and cardiodial transformations (Pittenger & Shaw, 1975) occur around a point. This point is a head-centered origin of a coordinate system that is analogous to the centers of moment found in other events. The experimental question then arose: Can observers judge the goodness of growth changes (represented as cardiodial changes) in profiles of a human head as a function of the origin of the head-centered (cardiodial) origin of those changes?

The procedure and results (Cutting, 1978b) are as follows, as shown in Figure 4. Panel *a* shows the profile selected, representing an early adolescent, and the matrix of nine points used as cardiodial centers. Panel *c* shows nine youthened profiles generated from the nine different centers in Panel *a*, and Panel *d* shows the nine aged

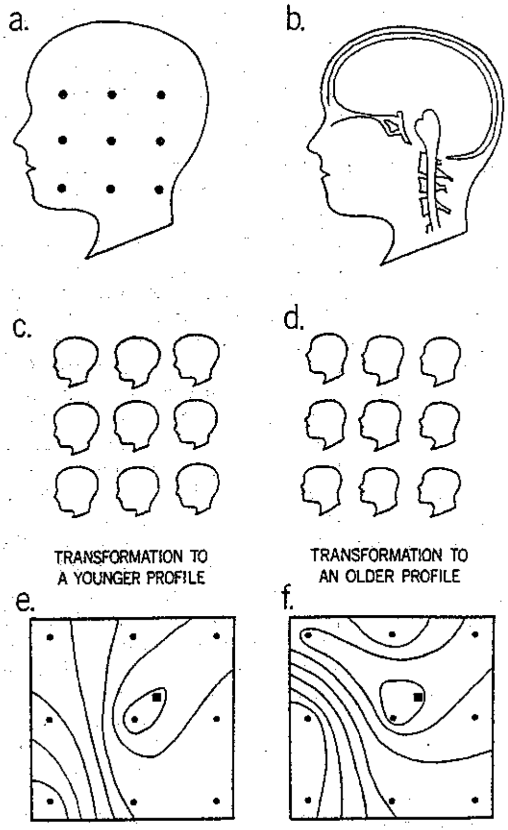

Figure 4. Panel a shows an intermediate-aged (adolescent) profile and the nine centers of moment chosen for cardiodial transformation. Panel b shows the internal anatomy of the same person. Panels c and d show the nine transformed profiles, each one corresponding to the locations given in the 3 x 3 matrix in Panel a. Panel c shows the youth-transformed profiles, and Panel d the aged profiles. Panels e and f show the judgment data of all subjects plotted as contour maps with respect to the 3 x 3 matrix. Squares indicate the maxima of best fitting paraboloids to the data, and should be mapped onto the matrix shown in Panel a and then onto the internal anatomy shown in Panel b. (After Cutting, 1978b).

profiles that correspond to them. Viewers were presented the standard profile (that shown in Panel a without the nine dots) and then the youthened profile and asked to make a judgment on a scale from 1 to 7 as to how well the second profile represented the younger version of the person represented in the first profile. The same procedure was then followed for the standard and the aged profiles. The data for the two tasks are shown in Panels e and f, plotted as contour maps for the averaged judgments of all observers. Because those data approximated elliptic paraboloids, I then calculated the maxima of the best fitting functions, and those are shown as the square dots in the bottom panels. Notice that the locations of these maxima are near the central point in the 3 x 3 matrix, but that they are slightly to the upper right. One can find this locale in the profile of Panel a and then transpose that point to the internal anatomy of a head shown in Panel b. Notice that the maxima of the viewers' judgments correspond to a point very near the foramen magnum, the point where the spinal column disappears into the skull and the brain stem. Interestingly, Enlow (1975) used this point as the origin of the framework for describing the developmental anatomy of the head. What I conclude is that observers are fairly good at discerning proper age and growth changes, and that these are generated by a mathematical transformation that closely mimics growth. In essence, people know the geometry of growth to the extent that they can recognize good and poor exemplars of the process.

As with the blown movements of trees and bushes, age changes in human profiles have no common motion. At least in the manner that the stimuli were generated, only object-centered coordinates and relative change are pertinent. This system and these changes are a structural description which best captures the aging process as it could be used in face perception. As before, a center of moment (here, the origin of the cardiodial transform) proves perceptually useful.

Thus far I have considered relatively fast, rigid (rolling wheels and walking humans) and nonrigid (swaying trees) events, and a relatively slow nonrigid event (aging faces). The fifth event type is a relatively slow rigid event, revealing important information to infrahumans.

Event 5: Rotating night sky

What information do migratory songbirds use to guide their long-distance flights? Among the many sources of information appear the locations of stars in the night sky. In particular, certain birds appear to sit in the upper branches of trees at night, hopping around limb to limb, observing the patterns of stars over time in a relatively cloudless sky. Over the period of several hours, as the earth rotates under the celestial sphere, the vectors paths of the stars begin to make concentric circles around the celestial poles--the North Star (Polaris) in the Northern Hemisphere, as shown in Figure 5. These paths, because they are all referred to the polar direction, could be used by the birds in selection of their migratory direction--south in the Fall and north in the Spring. Emlen (1975) demonstrated that naive birds (those that had neither migrated before nor had a chance to observe the patterns of stars to "memorize" the constellations) appeared to use the information, in Ptolemaic coordinates, about the rotation of the night sky. Placing the birds in small cages in a planetarium, Emlen noticed that they oriented in the direction of Polaris in the Spring and away from it in the Fall. Most convincingly, however, Emlen had the night sky rotate around a completely different point, Betelgeuse, and the birds oriented with respect to it.

Here, the center of moment of the night sky (Polaris for the Northern Hemisphere) is used as information for directing migratory behavior. It is a source that could be invaluable for long flights over dark water or terrain. But the perception of this center of moment must follow a different logic than those generally used in the previous events. Since the night sky is unbounded and is chock full of stars, no location in the sky will nullify relative rotational vectors nor minimize moment arms. All locations should be equally good. All computed relative

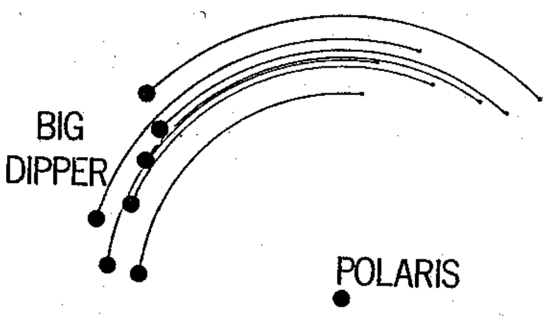

Figure 5. A six-hour rotation of the night sky would reveal circular vector paths around the North star, Polaris. Polaris is the center of moment for the night sky and is used for navigational purposes by migratory songbirds.

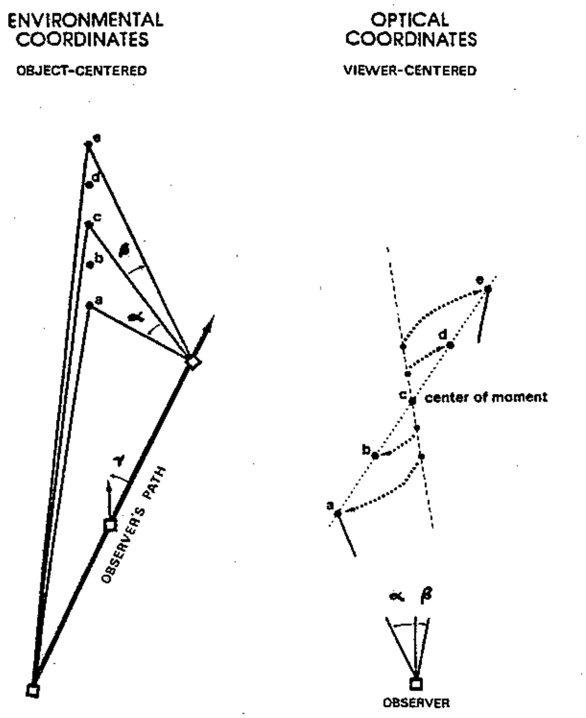

Figure 6. Representations of egomotion through space while observing an object off to the left side. Panel a shows these relations in environmental coordinates, Panel b in viewer-centered coordinates. The latter seems to be used for determining direction of locomotion. Object c is fixated, for example, nullifying common motion in the visual field and the object fixated becomes the center of moment for all rotational (and expansional) vectors. Relative velocities of these vectors in polar projection determine the direction headed: The growth of angle α is greater than that of β for the line of objects, a and e rotating around c. Angle γ, deviation of gaze from direction, is not considered.

motions would be the same. However, for all interpretations of this event other than rotation around the celestial poles, the stars would have a compensatory common counterrotation. But since no such counterrotation is seen, the common motion must be null. Thus, the equation given in the introduction reduces to absolute motion equals relative motion. In this event, the center of moment, or object-centered origin, is determined by superimposing the viewer- and object-centered coordinates, and letting the object coordinates (the night sky) rotate against the viewer coordinates creating a singularity in the sky that is the directional reference, North.

In the five events listed and discussed above, there is a center of moment within them that appears useful to the observer--dictating the wheel-likeness of a rolling object, gender of a human walker, relative arborization of a tree or bush, goodness of growth pattern in a human face, and celestial poles of the night sky. For completeness sake, I must now point out a sixth event that has a global center of moment that observers appear not to be able to use, but which has others that they can use.

Event 6: Expanding flowfields

Gibson (1950) claimed that when moving over a terrain one could tell one direction of motion optically by determining the location of the focus of expansion of visual flow. In my terms, the focus of expansion, the point dead ahead and directly behind one when locomoting, would serve as a center of moment of the event. More recently, however, Koenderink and van Doorn (1976) and Regan and Beverly (1982) have demonstrated that this is not true: The focus of expansion does not unambiguously specify direction because it is confounded with direction of gaze. A focus of expansion will always correspond to where one is looking, provided one is looking at a planar surface within 45° of where one is headed. Thus, this type of center of moment only tells the observer where he or she is looking, not where he or she is going.

One source of salient information for judging direction of motion, however, appears to be motion parallax. If one looks generally, but not directly ahead at an object in the middle distance, that object will lay at the origin of a viewer-centered coordinate system. Consider a concrete example. If one looks 5° off to the left at a tree while walking through a park, the angle between the line of gaze and line of direction will increase. But because registration of ocular rotation may not be very acute, and because researchers in this area are generally interested in optic information rather than kinesthetic information, this increase in gaze direction/motion direction angle is irrelevant: The tree always stays registered on the fovea, and what we are concerned with is the change in relative motion in viewer-centered coordinates of other objects in the field of view. Thus, as one walks through the park staring at the tree, the grass and other objects between the walker and the tree will shear across the line of sight to the left; grass and other objects farther away will shear across the line of sight to the right. In essence, the optic field undergoes rotation around the object of scrutiny, and this object becomes the new center of moment of import. This situation is shown in Figure 6. Direction of gaze, left or right, as it deviates from direction of motion is dictated by the differential rates of shear of textures across the gaze line, or alternatively the rotary motion, around the scrutinized object. Because of polar projection, the more rapidly moving textures always shear in the same direction that gaze deviates from motion. Thus, when looking left the more rapid shears are leftward; when looking right they are rightward. Such information seems adequate for telling gaze/motion deviations as small as a half a degree of visual angle (Cutting, 1982b).

Thus in this event general centers of moment fore and aft, the foci of expansion and contraction, are perceptually useless (Llewellyn, 1971; Johnston, White & Cumming, 1973). What is important are viewer-centered descriptions where the relative motions of objects with respect to the viewer and to an object under scrutiny can be registered. Fixating an object nullifies the common motions in the visual field with respect to the observer, and the perception and recovery of rigid structures can proceed from there. The object under scrutiny serves as the local center of moment for the rotation of the visual world. Motions with respect to that object dictate the direction of motion as it deviates from the object of gaze.

Summary

What I have presented are six types of situations that we might want to call events. In the first pair, rolling wheels and walking humans, I argued that for the viewer to perceive these structures she must generally extract the relative motions of parts from one another, accomplished perhaps through some minimization algorithm, and observe the common motion as the residual. These objects, wheels and walkers, have their own object-centered coordinates and an origin, the center of moment, around which their parts move. As wholes they move through viewer-centered coordinates following the vector of common motion. By definition, the center of moment for configurations on a wheel is the geometric center, or centroid, of that configuration. The center of moment for a walker need not be exactly at the center of the body and is not generally at the center of mass (Figure 2c): The center of moment, instead, is roughly at the intersection of diagonals drawn across the torso from shoulders to opposite hips.

In the second pair of events, swaying trees and aging faces, there is no common motion and thus viewer and object coordinates can be superimposed. Again, structural change is most properly discerned with respect to object-centered coordinates with an origin at the center of moment. The center of moment for a swaying tree is definitely not at the center of the configuration: it is at the point where the trunk meets the ground. The center of moment for a profile is generally near the center, located near the brainstem.

In the third pair of events, rotating night skies and expanding flowfields, the objects to be discerned encompass either most or all of the field of view: In the former event they are the stars in the rigid sky and in the latter they are the objects as laid out in rigid terrain. In both cases, in order to extract perceptually useful information, any common motion of the whole field must be nulled first. In the former case, for celestial navigation, nulled common motion reveals the relative vectors of the stars around the celestial pole. The location of the celestial pole, logically speaking, need not be at the center of the sky, nor does it even have to be visible. It need only be discernible from the vector paths of a few stars. In the latter case, for terrestrial navigation, null common motion of flow in the field of view is accomplished by looking at any object and maintaining fixation on it. That object then becomes the center of moment of the optic array and optic information rotates about that object, clockwise for leftward looking and counterclockwise for rightward viewing. If one happens to be looking directly in the direction one is going, no rotation occurs, only expansion. The center of moment for both events is important in viewer-centered coordinates and directional information available only after that has been determined.

In conclusion, the first four events are about object information. It comes as no surprise, then, that object-centered coordinates are most important and that motions and changes relative to the object's center of moment appear to be central to their perception. The latter two events are about viewer location. That is, the viewer wants to know where she is with respect to the environment. It comes as no surprise, then, that viewer-centered coordinates are most important and that motion changes relative to the viewer, the common motions of the flow, must be dealt with first. It would appear, then, that the task of perception, identifying a moving object or identifying where one is in a moving environment, dictates the use of the various coordinate structures that one might use--object-centered or viewer-centered. It seems likely that this would be true for both human and machine vision.

Acknowledgement

Supported by NIMH Grant MH37467.

References

Barclay, C.B., Cutting, J.E., & Kozlowski, L.T. Temporal and spatial factors in gait perception that influence gender recognition. Perception & Psychophysics, 1978, 23, 145-152.

Beckett, R. & Chang, K. An evaluation of the kinematics of gait by minimum energy. In D. Bootzin & H.C. Muffley (Eds.), Biomechanics. New York: Plenum Press, 1969.

Börjesson, E. and von Hofsten, C. A vector model for perceived object rotation and translation in space. Psychological Research, 1975, 38, 209-230.

Cutting, J.E. Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. Perception, 1978, 7, 393-405. (a)

Cutting, J.E. Perceiving the geometry of age in a human face. Perception & Psychophysics, 1978, 24, 566-568. (b)

Cutting, J.E. A program to generate synthetic walkers as dynamic point-light displays. Behavior Research Methods & Instrumentation, 1978, 10, 91-94. (c)

Cutting, J.E. Coding theory for gait perception. Journal of Experimental Psychology: Human Perception and Performance, 1981, 7, 71-87.

Cutting, J.E. Blowing in the wind: Perceiving structure in trees and bushes. Cognition, 1982, 12, 25-44. (a)

Cutting, J.E. Motion parallax and visual flow: How to determine direction of locomotion. Paper read at the 4th meeting of the International Society for Ecological Psychology, Hartford, CT, October, 1982. (b)

Cutting, J.E. & Kozlowski, L.T. Recognizing friends by their walk: Gait perception without familiarity cues. Bulletin of the Psychonomic Society, 1977, 9, 353-356.

Cutting, J.E. & Proffitt, D.R. The minimum principle and the perception of absolute, common, and relative motions. Cognitive Psychology, 1982, 14, 211-246.

Cutting, J.E., Proffitt, D.R., Kozlowski, L.T. A biomechanical invariant for gait perception. Journal of Experimental Psychology: Human Perception and Performance, 1978, 4, 357-372.

Duncker, K. Induced motion. In W.D. Ellis (Ed.), A sourcebook of Gestalt psychology. London: Routledge & Kegan Paul, 1937.

Dürer, A. The human figure: The Dresden sketchbooks. New York: Dover, 1972. (Originally published in 1528).

Emlen, S.T. The stellar-orientation system of a migratory bird. Scientific American, 1975, 233 (2), 102-111.

Enlow, D. Handbook of facial growth. Philadelphia: W.B. Saunders, 1975.

Gibson, J.J. The perception of the visual world. Boston: Houghton Mifflin, 1950.

Gogel, W. The adjacency principle in visual perception. Scientific American, 1978, 238 (5), 126-139.

Johansson, G. Configurations in event perception. Uppsala, Sweden: Almqvist & Wiksell, 1950.

Johansson, G. Visual perception of biological motion and a model for its analysis. Perception & Psychophysics, 1973, 14, 201-211.

Johansson, G. Spatio-temporal differentiation and integration in visual motion perception. Psychological Research, 1976, 38, 379-393.

Johnston, I.R., White, G.R. & Cumming, R.W. The role of optical expansion patterns in locomotor control. American Journal of Psychology, 1973, 86, 311-324.

Koenderink, J.J. & van Doorn, A.J. Local structure of movement parallax of the plane. Journal of the Optical Society of America, 1976, 66, 717-723.

Kozlowski, L.T. & Cutting, J.E. Recognizing the sex of a walker from a dynamic point-light display. Perception & Psychophysics, 1977, 21, 575-580.

Llewellyn, K.R. Visual guidance of locomotion. Journal of Experimental Psychology, 1971, 91, 245-261.

Marey, E.J. Movement. New York: Arno Press & New York Times, 1972. (Originally published in 1895).

Marr, D. Vision. San Francisco: Freeman, 1982.

Murray, M.P. Gait as a total pattern of movement. American Journal of Physical Medicine, 1967, 46, 290-333.

Pittenger, J.B. & Shaw, R.E. Aging faces as viscal-elastic events: Implications for a theory of nonrigid shape perception. Journal of Experimental Psychology: Human Perception and Performance, 1975, 1, 374-382.

Proffitt, D.R. & Cutting, J.E. Perceiving the centroid of configuraton on a rolling wheel. Perception & Psychophysics, 1979, 25, 389-398.

Proffitt, D.R. & Cutting, J.E. An invariant for wheel-generated motions and the logic of its determination. Perception, 1980, 9, 435-559. (a)

Proffitt, D.R., Cutting, J.E., & Stier, D.M. Perception of wheel-generated motions. Journal of Experimental Psychology: Human Perception and Performance, 1979, 5, 289-302.

Regan, D. & Beverly, K.I. How do we avoid confounding the direction we are looking and the direction we are going. Science, 1982, 215, 194-196.

Rubin, E. Visuell wahrgenommene wirkliche Bewegungen. Zeitshrift für Psychologie, 1927, 103, 384-392.

Todd, J.T. & Mark, L.S. Issues related to the prediction of craniofacial growth. American Journal of Orthodontics, 1981, 79, 63-80.

Todd, J.T., Mark, L.S., Shaw, R.E. & Pittenger, J.B. The perception of human growth. Scientific American, 1980, 242 (2), 132-144.

Ullman, S. The interpretation of visual motion. Cambridge, MA: MIT Press, 1979.

Wallach, H. Visual perception of motion. In G. Kepes (Ed.) The nature and the art of motion. New York: George Braziller, 1965. Revised in H. Wallach, On perception. New York: Quadrangle/New York Times, 1976.