

N 9 2 - 2 1 4 7 2

## OPTICAL FLOW VERSUS RETINAL FLOW AS SOURCES OF INFORMATION FOR FLIGHT GUIDANCE

James E. Cutting  
Cornell University  
Ithaca, New York

A recurring issue in the workshop concerned the appropriate description of visual information for flight guidance—optical flow vs. retinal flow. Most descriptions in the psychological literature are based on the optical flow. However, human eyes move and this movement complicates the issues at stake, particularly when movement of the observer is involved.

The basic question addressed here is: Can an observer, whose eyes register only retinal flow, use information in optical flow? The answer, I suggest, is that he/she cannot and does not reconstruct optical flow; instead, he/she uses retinal flow. To clarify what is meant, some definition of terms is needed.

### GLOSSARY

**Optical array.** The projections of a three-space environment to a point within that space. All measurements to this point should be made in steradians, solid degrees of subtended angle; typically, however, most descriptions are in degrees. The relations among these projections provide an important beginning to an understanding of information as used in visual perception. The optical array is best and most conveniently represented as a spherical projection surface, and centered on an observer's eye, which is at the nodal point of the projection.

**Retinal array.** The projections of a three-space onto a point and beyond to a movable, nearly hemispheric sensing device, like the retina. (1) Its movability, (2) its differential ability to register detail (acuity differences in the fovea, parafovea, and periphery), (3) its boundedness (edges at the orbit and nose), and (4) its slight deformations (due to the difference between center of rotation of the eye and its nodal point) distinguish it from the optical array. Movability is the critical factor in separation of optical and retinal flow; movability is evolutionarily designed to counter problems of acuity differences.

**Flow.** Global motion represented as a field of vectors, best placed on a spherical projection surface, as shown in figures 1 and 2. (These figures are Figures 11.2 and 11.3 from Cutting, 1986.) Specifically, flow is the mapping of the field of changes in position of corresponding points on objects in three-space onto a point, where that point has moved in position. Conventionally, the field of vectors is registered from two different but nearly adjacent points in three space (such as locations along the path of a moving observer). I will call these two points along the path registration points. Example of the differences between optical and retinal flow are given in figures 1 and 3 (Figure 11.4 from Cutting).

**Focus of expansion.** The point projected out to the horizon along the linear path one is taking. It is the (putative) point in the optical array from which all mapping vectors (flow) appear to be oriented.

**Curl.** The curvature of local vectors in a flow field. Some examples are given in figure 3.

**Wayfinding.** Determining the instantaneous direction one is traveling in.

## GROUND RULES

One must be able wayfind (to determine one's heading to guide flight, particularly helicopter flight) from perceptual information. This information could be purely visual, or combined with other modalities (vestibular activity). Analyses of the wayfinding requirements during running, skiing, and landing fixed-wing aircraft (Cutting, 1986, p. 152, 277-278) suggest we need an accuracy of 1 degree of visual angle at any point in time.

**Caveat:** These comments do not consider overt control of an aircraft, only on the perceptual needs of the pilot to initiate control adjustments.

## CAN WE DERIVE OPTICAL FLOW FROM SUCCESSIVE OPTICAL ARRAYS?

Participants at the workshop differed in their understandings about how optical flow and retinal flow information might be useful to a pilot, or any other moving observer. The issues are based on deeper conceptions of how flow in the optical array is determined. In essence, all workshop participants agree on the optical array, its generation, and its importance; there is disagreement, however, about optical flow.

In considering optical flow, should we be concerned with three degrees of freedom or six? Again, the optical array is all the projections of a three-dimensional space to a point. Since the spatial position of a point can be specified by three coordinates [ $x$  (lateral),  $y$  (vertical), and  $z$  (depth)], the optical array is concerned with only these three degrees of freedom.

However, to move the focal point of the optical array through three space, more than the change in  $x$ ,  $y$ , and  $z$  may be entailed. Specifically, rotations around  $x$  (pitch), around  $y$  (yaw), and around  $z$  (roll) might have to be considered, as seen in figure 3. Here's why:

## GENERAL PROBLEMS FOR CONSIDERATIONS OF OPTIC FLOW

(1) The contents of the optical array cannot be registered at a point, but only on a projection surface (hemisphere or plane) behind that point. Changes in positions of the contents of the optical array

need a surface on which to be represented. This means the surface of registration must be oriented to the points in the three-space from which they project.

(2) Optical flow vectors are typically extremely small within 15 degrees of the focus of expansion (see fig. 1). Threshold measures for registering motion become very important.

(3) To plot optical flow, one must map (at minimum) corresponding projected points of the environment across at least two successive optical arrays. For convenience's sake, call these registration points  $t_1$  and  $t_2$ , measured at two successive time intervals. The central question is: Can these optical flow mappings be constrained on reasonable (1) mathematical, (2) optical, or (3) extraocular grounds, and/or (4) can optical flow be bypassed and wayfinding be done purely on retinal grounds?

## **I. MATHEMATICAL METHODS FOR CONSTRAINING OPTICAL FLOW**

A purely mathematical approach can consider only projections from points in three space or from surfaces with unknown orientation and Gaussian curvature in three space, onto a spherical projection surface. In particular, the mathematics does not, a priori, allow horizons or other reference points.

1. Random choice. Pick a point anywhere in three-space and use it as the origin of the mapping system. That is, the projection of this point and only this point will necessarily map onto itself across  $t_1$  and  $t_2$  on the projection surface. It is an identity element; it will be a vector of zero length and hence no orientation. This fact is the essence of Brouwer's theorem in topology, which states that any field of mappings must have at least one point that maps onto itself. Brouwer's theorem is silent on the location of this identity element.

Result: A vector (flow) field will be generated.

Problem: Unless one has, accidentally or a priori, picked a point at (functionally) infinite distance, the flow field will have curl as shown in figure 3. Hence, this mapping will hide the location of the focus of expansion, hide the direction one is going, and give misinformation about most optic flow variables.

[Functional infinity, here, means any distance which is, say, at least three orders of magnitude larger than the distance between registration points,  $t_1$  and  $t_2$ .]

2. Selection of many points and comparison of curl. Simultaneously consider many mappings using a large number of points as origins, and compare curl across mappings.

Result: Any solution without curl reveals the focus of expansion.

Problems: (1) If there are no solutions without curl, the center of the projection (the pilot's eye) is in a three-space environment with no point at (functional) infinity. In practical terms, this environment has no horizon. Nap of the earth (NOE) flight may entail such environments. (2) The procedure is either iterative (repeated stepping through comparisons of flow fields), or requires multiple coordinated registration systems (e.g. many visual systems). This is computationally expensive and

psychologically impossible. (3) If the observer pilot is on a curvilinear path, all representations of flow (optical and retinal) will have curl, and no solution can be obtained.

3. Flow decomposition. In artificial intelligence approaches to the problem, it is assumed that since contributions to flow from pitch, yaw, and roll on the one hand, and translations in x, y, and z, on the other hand are additive, the flow can be decomposed by subtractive methods. That is, any curl in the flow field occurs because of rotations and can be subtracted out.

Result and Problem. Mathematically this is problematic. It reduces to  $a + b = c$ , if a is flow due to rotations, b is flow due to translations, and c is the resultant flow. However, although knowing a and b specifies c, knowing c does not allow one to determine the values of a and b. This formulation reduces then to the iterative method above and has the same problems.

Mathematical Conclusion. Since, in mathematical terms no point in three-space can have privilege in the mapping across  $t_1$  and  $t_2$ , and hence the formation of the vector field, there is no a priori, noniterative mathematical way to generate an optical flow field. The problem reduces to the fact one cannot guarantee the coordinates of the registration of flow will not have undergone pitch, yaw, or roll (rotations around x, y, and z axes), while simultaneously undergoing pure translation in x, y, and/or z.

## II. OPTICAL METHODS FOR CONSTRAINING OPTICAL FLOW

Optical methods presuppose an environment. For a pilot at any substantial altitude, this environment is essentially planar and has a horizon. Any information from or about eye movements, however, is excluded in this approach.

1. Horizon. The horizon may provide one explicit method for determining the traditional representation of optical flow (e.g., figure 1), preventing unruly mappings involving "spurious" pitch (x rotation) or roll (z rotation). Because the horizon is a series of points in three space at infinite distance, any and all of these points provide anchor for flow registration across  $t_1$  and  $t_2$ .

Result. Since the horizon holds a constant position in the mapping from one optical array to the next, flow vectors will not undergo curl due to pitch and roll (see the top and bottom panels of figure 3). This would facilitate location of a focus of expansion, and orientation of the optical array.

Problem. Unless texture along the horizon is used, there is no guarantee that yaw (y rotation) will not occur in flow mapping, and yaw rotations induce substantial curl, as shown in the middle panel of figure 3.

2. Horizon plus distant texture. Use the horizon to anchor the registration of flow against pitch and roll, and use any available texture near the horizon to anchor it against yaw.

Result: The result will be the "standard" optical flow pattern.

Problems: The horizon isn't usually available in some environments (inside buildings), and more particularly to the goals of the workshop the horizon is not generally available in NOE flight. It cannot be guaranteed that there will be any objects or texture at infinite distance.

3. Extrapolation from most rapid flow velocities. A look at the panels in figure 3 shows substantial curl in the vector fields near the direction of heading, but also shows large vectors (corresponding to rapid optical velocities) generally without curl beneath the individual, nearest the ground. Since these are relatively immune to curl, they could be used to determine heading.

Result: In principle this method could work and one could find one's heading perhaps within several degrees of visual angle even if the horizon were occluded.

Problems: (1) A pilot often cannot, or cannot afford to, look directly below to observe such flow, particularly in NOE flight. (2) Most rapid flow is contingent on the environment one is in and one's relation to it. In NOE flight, most rapid flow need not be beneath the aircraft and hence could be difficult to find. (3) When one "looks" at anything, one fixates on it and will use pursuit movements, creating retinal flow radically different from optical flow. To use rapid optical flow one must anchor fixation. There are three methods of doing this. Two are looking at the horizon or any other point at functional infinity. The problem here is that most rapid flow is usually at 90 degrees to this anchor, and difficult to register by eye. The third is by looking at the reference point at the edge of or on the windscreen of the craft. Such edge rates are known to be useful in open environments (e.g. over planar fields in simulations) but they would not be in NOE flight. (4) If a pilot is on a curvilinear path, the most rapid flow will have curl. Vectors will point to various locations in the distance generally in the direction of the curved path, but not along it (see Cutting, 1986, p. 209).

4. Marks on the windscreen. The best method for negating eye movements is to look at a fixed point on a windscreen and observe flow. Flow will always be in the opposite direction to one's line of movement.

Results and Problems: If one had enough marks on the windscreen one could generally pick out the focus of expansion by fixating various marks. This method is important for the practical task of piloting an aircraft, but is not available to a pedestrian or runner. Moreover, too many marks on a windscreen will impair visibility.

Optical conclusion. When one is in an environment without a guaranteed horizon, as in NOE flight, there is no optical method (other than many marks on a windscreen) that can guarantee finding the focus of expansion and anchoring the vectors in the optical flow field. Hence the standard optical flow variables are (or may be) indeterminate.

### **III. EXTRAOCULAR METHODS FOR CONSTRAINING OPTICAL FLOW**

These methods use information from extraretinal sources, such as feedback from eye muscles and/or from the semicircular canals of the vestibular system, to anchor the registration of flow in the optical array.

1. Registering sensor rotations. Since flow due to rotations are (thought exclusively to be) due eye movements, and flow due to translation are due to observer movement, one could decompose these contributions in retinal flow by registering eye movement extent from muscle activity.

Result. In principle this would work.

Problems. In practice it does not. We are not sufficiently kinesthetically sensitive to eye movements to drive guidance within 1 degree of visual angle.

2. Preventing sensor rotations. In principle the vestibulo-ocular system is gyroscopic. The vestibular system can be used (particularly in VOR) to hold eye position, preventing rotations.

Result. Again, in principle this would work.

Problem. The normal activity of the eye is driven not by VOR but by a field holding response which serves to direct the eye to an object and hold it (or some part of it) in position on the fovea. This field holding response overrides VOR. It is extremely difficult to stare off into the distance at a fixed angle and observe flow. One is constantly captured by objects, which one pursues, then one saccades back in the opposite direction. This phenomenon is optokinetic nystagmus (OKN).

Extraocular conclusion. In principle optical flow could be determined from extraretinal sources of information. In practice, however, the information is probably too coarse, at least in human observers.

#### **IV. A RETINAL METHOD FOR BYPASSING OPTICAL FLOW: Differential motion parallax (DMP)**

A retinal method of determining flow tries to bypass problems of reconstructing optical flow. That is, rather than trying to nullify rotations of the optical array, this approach embraces the rotations in the retinal array and see what regularities fall out. In essence, the claim is that while the optical array is relevant to perception, optical flow is not; only retinal flow is relevant.

Cutting (1986, Chapters 10-13) describes a retinal invariant that serves, in most environments, to indicate the direction of forward movement, whether linear or curvilinear, with respect to gaze. Its crux is that, when one is fixating an object while moving through an environment, the retinal velocity of near objects will generally be faster and in the opposite direction to far objects. Most rapid flow is in the opposite direction from heading. Thus, regardless of one's path, so long as one is fixated on an object in mid-distance, if the most rapid flow is leftward, one heading is to the right of gaze. If most rapid flow is right, heading is left-of gaze. Thus, more simply:

$$N > -F, \tag{1}$$

where N stands for the retinal velocities of near objects (and given positive sign) and F stands for the retinal velocities of far object. Zero retinal velocity, of course, is where one is looking.

Results of four experiments were presented in the workshop supporting DMP: three (Cutting, 1986, Chapters 12 and 13) indicate the efficacy of DMP along linear and curvilinear paths, the fourth (unpublished) indicated its efficacy in situations mimicking the bounce and sway of normal gait. Moreover, an analysis of errors indicated that DMP was used throughout, not just on correct trials. That is, given certain situations, DMP can fail due to the relative positions of objects in an environment (see Cutting, 1986, p. 197).

**Retinal conclusion.** Information is available in the retinal array for guidance. This information is generally trustworthy, but fails in certain environments with particular distributions of objects in it. A consideration of the failures is important to testing DMP in experimental situations.

**Constraints.** For DMP to operate, the minimum requirements are that objects be laid out in depth around a fixated object such that there are both nearer and farther objects near the line of sight. When there are no objects farther than the fixated object there is no problem; the motion of farther possible objects is zero and does not effect the inequality in equation 1 above. When there are no objects nearer than the fixated object, however, DMP fails completely.

What kinds of visual information DMP ignores. Differential motion parallax is measurement about pure motion. It is unconcerned with occlusions of near objects by far objects. It knows nothing about the sizes of objects, their identity, or their location in three space. It is also a measurement in the retinal array unconcerned with any ability to resolve motion. That is, it assumes retinal velocities can be measured with (roughly) equal efficacy everywhere, particularly above and below the line of gaze.

## **RECURSIVE RULES FOR WAYFINDING**

- (1) Fixate an object of potential interest in your environment.
- (2) If there is no flow across the line of gaze (or vertical plane passing through the line of gaze), you are looking in the direction you are going.
- (3) If there is flow across the line of gaze, follow the fixated object with pursuit eye movements,
- (4) During this pursuit, register the relative motions of objects near and far around the fixated object,
- (5) Assess if the information is adequate as an update of your current heading.
- (6) If it is, go back to Step 1.
- (7) If it is not, shift eyes in the direction opposite from the most rapid flow,
- (8) Go back to Step 1.

Caveat: Despite the consideration of fixations, pursuit movements, and saccades, DMP is not an eye movement theory of wayfinding; it is a partial and corollary explanation of the efficacy of eye movements and fixations as part of visual explorations of the environment during self movement.

In particular, notice one need not implement Steps 7 and 8. Once one has the general idea of one's heading, fixations, pursuit movements, and saccades need only reinforce the perception or knowledge of heading. In other words, the pilot can implement Steps 1 through 6, never (or perhaps only occasionally) looping through the whole set.

## **FIVE EXAMPLES OF THE RETINAL OPTICS RELEVANT TO DIFFERENTIAL MOTION PARALLAX IN CLUTTERED ENVIRONMENTS**

1. Looking at or near the focus of expansion. There is no DMP when looking at the focus of expansion. Moreover, DMP will often fail when looking near it. That is, because one wants to avoid objects in one's path, a pilot has already changed course to remove them from the path. This means there are few objects available to create the foreground motion required in DMP.

Implication: This fact, the failure of DMP near the focus of expansion, may be why pilots and car drivers spend so little time looking in the direction of their path of movement. Retinal motion information there is either (1) nil, or (2) DMP information is contradictory.

2. Looking at an object in the very far distance off the path of movement: All retinal will opposite in direction from heading. That is, leftward retinal flow indicates a gaze angle to the left of one's direction of movement; rightward flow indicates a gaze to the right. Again, technically there is no differential motion parallax, only motion perspective.

3. Looking at an object in the mid-distance: DMP reigns. Any object or texture along the line of gaze (or along a vertical plane through the line of gaze) that is half the distance (or less) to the fixated object will move faster than, and in the opposite direction to, any object or texture at infinite distance. If there are no objects nearer than half the distance, DMP will fail.

4. Looking at a close object: DMP will always fail. But looking at nearby things is generally not what one does when one is interested in where one is going.

Implication: Never look at close objects when wayfinding. This may be part of the problem with in cockpit instrumentation. Wayfinding is not about looking nearby.

5. Looking at the marks on the windscreen. Although marks on the windscreen are closer to a pilot than any external object in the environment, the fact that they do not move with respect to a pilot (sitting still) makes these marks at a distance of functional infinity. This situation becomes exactly like situation 2 above, provided the craft is not undergoing any rotations.

Meta-rules for visual guidance by DMP: Both rules have practical reasons for implementation known for along time; both are reinforced by the optics of DMP.



- (1) Spend little time looking exactly in the direction one is going;
  - (a) there is either no suprathreshold motion there, or
  - (b) what suprathreshold motion there is may yield contradictory DMP information.
  
- (2) Spend little time looking at objects in the near foreground;
  - (a) one may have to rotate one's eye too rapidly to maintain fixation.
  - (b) there can be contradictory DMP information there.

## **EXPLORATIONS OF DMP IN SIMULATIONS RELEVANT TO HELICOPTER FLIGHT**

During the course of the workshop several of us (Tom Bennett, John Flach, Dean Owen, Lawrence Wolpert, Greg Zacharias, and I) designed an experimental situation to explore the use of DMP and optical flow. The situation uses a head-mounted display responsive to head rotations (pitch, yaw, and roll). Thus, the simulator pilot moves his/her head to obtain new vistas of the environment he/she is flying through. In this environment, one can fly a dog-leg path in NOE flight through and around a series of poles, towards a goal. Computer software is being developed to record the head movements and derive the objects the pilot is looking at. The objective, from my perspective, is to determine if pilots use DMP to guide their flight.

We assume (1) that participants can generally fly an aircraft and (2) that they can learn to use headmovements instead of eye movements. In this manner changes in the position of the center of the display should indicate the locus of the two dimensional array they find interesting. This second assumption may be prove wrong, but the participants of the workshop agreed that positive evidence of the use of DMP would also serve as positive evidence for this assumption.

## **REFERENCES**

Cutting, James E., Perception with an Eye for Motion. MIT Press, 1986.

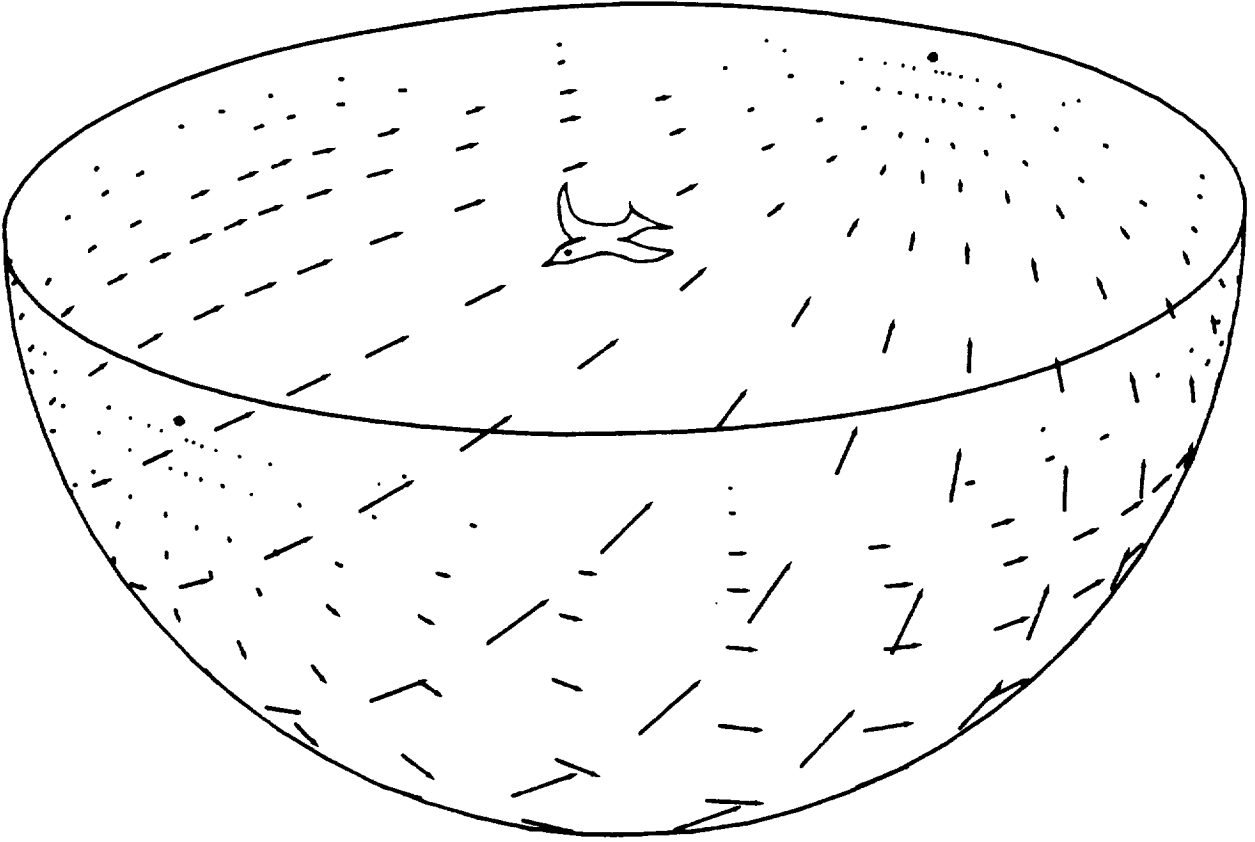


Figure 1. A computer-generated plot of the projection shown in figure 11.1. Notice that the vectors near the foci of expansion and contraction are smaller than in gibson's figure.

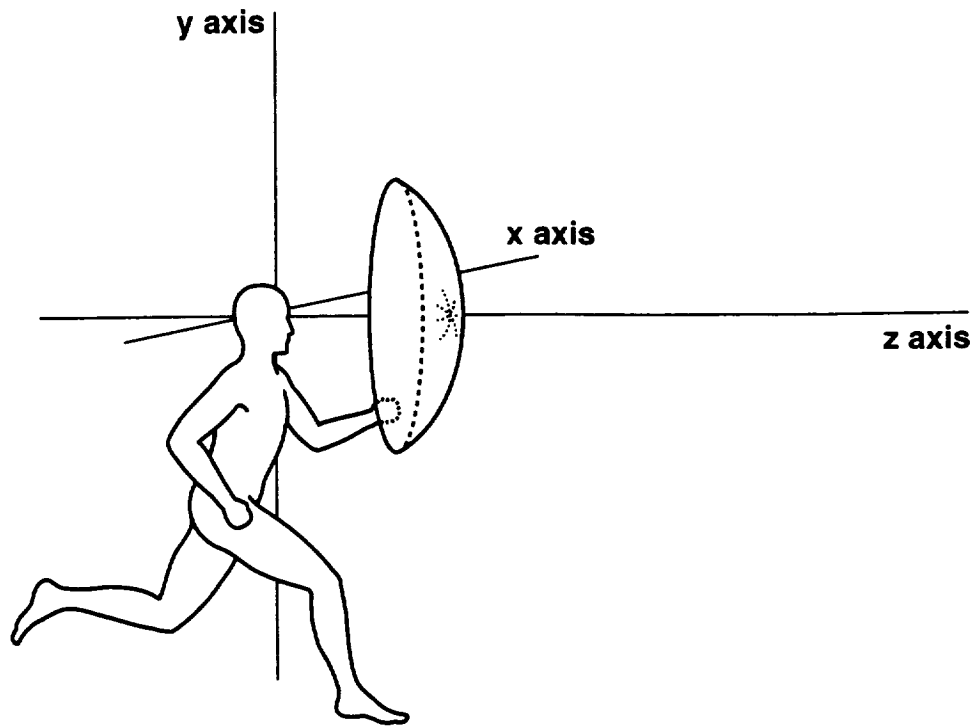


Figure 2. A representation of leonardo's window, in which the moving observer carries the projection surface through the environment. In accordance with the lower panel of figure 10.1, I call this leonardo's windshield. It is a section of the spherical projections shown in figures 11.1, 11.2, 11.4, and 11.5. Three axes of potential rotation are noted: x (which extends side to side across the observer, y (which runs vertically), and z (which extends along the linear path of movement. These axes of rotation are used in figure 11.4.

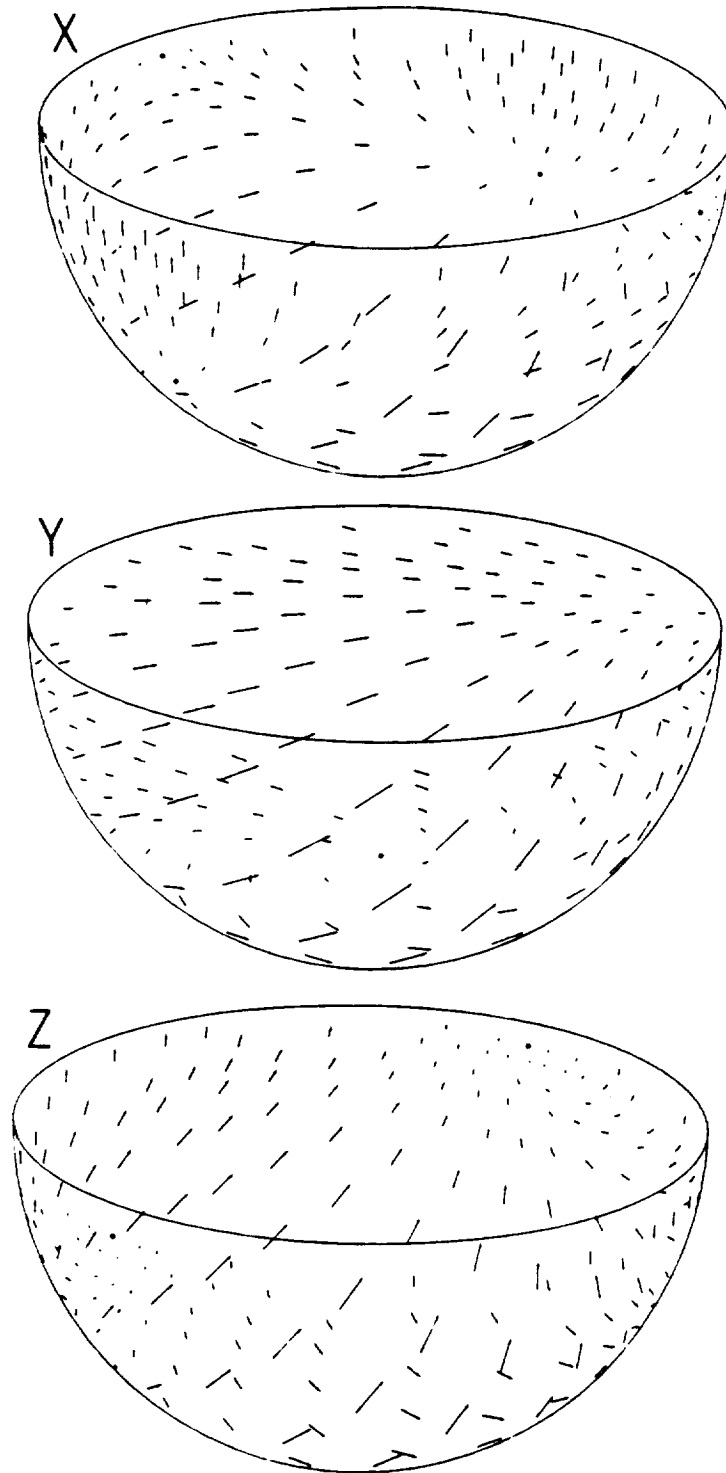


Figure 3. Alternative mappings of flow during forward linear locomotion. (Top) The spherical projection surface has been rotated around the x axis. Singularities are created where the x axis meets the horizon and beneath the observer. (Middle) Rotation around the y axis. Two new singularities are created to the left of the observer (one hidden at the edge of the drawing). (Bottom) Rotation around the z axis. The displays are valid representations of optic flow, as much so as those in figures 11.1 and 11.2.