NISTIR 4722

INVARIANTS IN VISUAL MOTION

DANIEL RAVIV

Florida Atlantic University
The Robotics Center and
The Electrical Engineering Department
Boca Raton, FL 33431
and
Sensory Intelligence Group
Robot Systems Division

U.S. DEPARTMENT OF COMMERCE
National Institute of Standards
and Technology
Robot Systems Division
Bldg. 220 Rm. B124
Gaithersburg, MD 20899

U.S. DEPARTMENT OF COMMERCE
Robert A. Mosbacher, Secretary
NATIONAL INSTITUTE OF STANDARDS
AND TECHNOLOGY
John W. Lyons, Director

NIST

NISTIR 4722

INVARIANTS IN VISUAL MOTION

DANIEL RAVIV

Florida Atlantic University
The Robotics Center and
The Electrical Engineering Department
Boca Raton, FL 33431
and
Sensory Intelligence Group
Robot Systems Division

U.S. DEPARTMENT OF COMMERCE
National Institute of Standards
and Technology
Robot Systems Division
Bldg. 220 Rm. B124
Gaithersburg, MD 20899

November 1991

U.S. DEPARTMENT OF COMMERCE
Robert A. Mosbacher, Secretary
NATIONAL INSTITUTE OF STANDARDS
AND TECHNOLOGY
John W. Lyons, Director

ABSTRACT

This paper deals with new and simple representations of 3-D points in a moving-observer coordinate system. Assuming rectilinear motion with no rotation of an observer where the optical axis coincides with the direction of motion, and a stationary scene, points in 3-D space that lie on a particular 3-D surface produce constant value of some nonlinear function of the measurable image optical flow. Five sets of different surfaces are introduced and there is one optical-flow based constant value for each surface. We called these values "invariants". It is shown how to extract these invariants and how to use them for representing 3-D space.

1. INTRODUCTION

During eye motion in a stationary environment, the projection of objects in the world is continuously changing, but we perceive the world as stationary. Are there properties of the image that under some transformation remain constant during the motion of the eye? In other words, are there *visual invariants?*

Gibson [11] capured this idea as follows:

"If invariants of the energy flux at the receptors of an organism exist and if these invariants correspond to the permanent properties of the environment, and if they are the basis of the organism's perception of the environment instead of the sensory data on which we have thought it based, then I think there is new support for realism in epistemology as well as for a new theory of perception in psychology."

This paper deals with new and simple representations of 3-D points in a moving-observer coordinate system. Assuming rectilinear motion with no rotation of an observer where the optical axis coincides with the direction of motion, and a stationary scene, points in 3-D space that lie on a particular 3-D surface produce constant value of some nonlinear function of the measurable image optical flow. Five sets of different surfaces are introduced and there is one optical-flow based constant value for each surface. We called these values "invariants". For example, one invariant corresponds to all points in 3-D that lie at the same distance from an observer. This means that a nonlinear function of optical flow will result in the same value for all points on a particular sphere that surrounds the observer.

The optical flow analysis takes place in a spherical $(R-\Theta-\Phi)$ coordinate system (rather than X-Y-Z). In this representation, since we deal with translation only along the optical axis, any point in this coordinate system moves along a constant ϕ radial line and can be processed independently of any other point.

There are several advantages to using these invariants:

1. They are measured in camera coordinates, i.e., there is no need to transform to object's coordinate system.
2. They can be very useful when local qualitative as well as quantitative vision-based motion related decisions need to be made e.g., how to avoid an obstacle.
3. Only one camera is needed to extract these invariants, and all measurements can be obtained from the visual data.
4. The magnitude of the camera velocity vector need not be known. The invariants are measured in *time* units rather than *distance* units.

5. Using the invariants it is possible to obtain simple representation of space, and define obstacles in a relatively simple way.
6. The measurement of these invariants become very simple when logarithmic retinas are used.
7. These calculated values of optical flow can be used to find relative locations of 3-D points without first reconstructing the 3-D world.

Several uses of the invariants such as understanding of fixation, analysis of looming, robust methods for 3-D reconstruction, etc. are mentioned in the text.

2. COORDINATE SYSTEM FOR 3-D SPACE

In a rectilinear motion with no rotation, points in the image plane move away from the Focus of Expansion (FOE) (Figure 1a) and towards the Focus of Contraction (FOC) (Figure 1b). Based on this observation we use an R-Θ-Φ spherical rather than an X-Y-Z cartesian representation of points in space, which reduces to a Θ-Φ representation in the image domain.

Figures 2 and 3 show the chosen coordinate system and the definitions of r and the angles θ and ϕ. (Note that the R-Θ-Φ coordinate system corresponds to the velocity egosphere defined in [7].) If the optical axis coincides with the translational vector then, in the image domain, constant ϕ corresponds to a radial line that emerges from the FOE and constant θ corresponds to a circle whose center is the FOE. Given a point in cartesian coordinate system, it can be transformed to a (r–θ–ϕ) point in the (R–Θ–Φ) domain and vice versa.

In this paper we assume that the camera (observer) undergoes translational motion along the optical axis. Using the R–Θ–Φ coordinate system, any point in the image domain moves radially i.e., along a constant ϕ line. In this coordinate system a point in the image can be processed independently of any other point, and a constant ϕ line may be processed as a 1-D image. For example, in a translational motion when the optical axis coincides with the translation vector, each point on an edge in the image moves radially away from the FOE. This point can be traced and both its optical flow and location in the image can be obtained independently of any other point in the image.

3. MOTION INVARIANTS

We describe five invariants. First, we derive the general relationship between the moving observer and a point in 3-D space. Based on this relationship the invariants are obtained. It is shown that points in space that lie on a specific 3-D surface share the same time-dependent invariant. Geometrical invariants are introduced, followed by a derivation of each invariant from the image optical flow.

The invariants are time-dependent only, i.e., points in space are described in terms of time (scaled space) [1,2,4] rather than their location in space.

Refer to Figure 4. For a rectilinear and continuous motion of the camera, any point in space obeys:

$$\Delta l = V \; \Delta t \; \sin\theta \tag{1}$$

$$\Delta l = (r - \Delta r) \tan\Delta\theta \tag{2}$$

Except for $\theta=0$ and $\theta=\pi$, then from (1) and (2), for $\Delta t \to 0$, $\tan\Delta\theta \to \Delta\theta$, and $\Delta\theta\Delta r \to 0$

$$V = r\frac{\dot{\theta}}{\sin\theta} \tag{3}$$

where the dot denotes differentiation with respect to time. $\dot{\theta}$ is the optical flow along a radial line, i.e., for constant ϕ. The value $r\frac{\dot{\theta}}{\sin\theta}$ is the same for *all points* (except those with $\theta=0$ and $\theta=\pi$) in 3-D space. If, in addition, the velocity V of the observer remains constant, then this relationship holds for all time instants. θ and $\dot{\theta}$ can be measured/computed at each instant of time, and hence the nonlinear function $\frac{\dot{\theta}}{\sin\theta}$ can be obtained.

The following is a description of the five invariants in terms of optical flow. We denote them by $\frac{1}{\tau_R}$, $\frac{1}{\tau_T}$, $\frac{1}{\tau_S}$, $\frac{1}{\tau_C}$, and $\frac{1}{\tau_P}$. They all have units of $\frac{1}{time}$.

1. *The equal range invariant* $\frac{1}{\tau_R}=\frac{\dot{\theta}}{\sin\theta}$ (Figure 5): All points in 3-D space (except those that lie on the motion axis, i.e., with $\theta=0$ and $\theta=\pi$) that lie on a sphere whose center is the pinhole point of the camera share this invariant, i.e., have the same τ_R. In this case the radius r is constant, and so, using Equation (3), the ratio $\frac{V}{r}$ is kept constant, and $\frac{\dot{\theta}}{\sin\theta}$ remains constant. The meaning of this invariant is that the modified optical flow $\frac{\dot{\theta}}{\sin\theta}$ is the *same* for all points on a sphere (except those which are on the axis motion of the camera). Points inside the sphere ("close" points) produce higher values of $\frac{\dot{\theta}}{\sin\theta}$ and points outside the sphere ("far" points) produce smaller values of $\frac{\dot{\theta}}{\sin\theta}$. Therefore it is possible to find the relative distance of a point simply by calculating or measuring this value. Points in 3-D space can be viewed as lying on shells, each of which has a different $\frac{1}{\tau_R}$ invariant.

2. *The fixation invariant* $\frac{1}{\tau_T}=\dot{\theta}$ (Figure 6): All points in 3-D space (except those that lie on the motion axis, i.e., with $\theta=0$ and $\theta=\pi$) that lie on a torus obtained by rotating a circle which is tangent to the direction of motion axis about this axis, share this invariant, i.e., have the same τ_T. This circle is an "Equal Flow Circle" (EFC) as described in [6]. In this case the diameter of the torus, i.e., $\frac{r}{\sin\theta}$ is constant, and so using Equation (3), $\dot{\theta}$ remains constant. (Note that r is the distance from the observer to the point and *not* the radius of the torus.) An extension of this invariant to a more general motion of a camera has led to a quantitative approach to camera fixation [5,6] and to road following [10].

3. *The looming invariant* $\frac{1}{\tau_S}=\frac{\dot{\theta}}{\tan\theta}$ (Figure 7): All points in 3-D space (except those that lie on the motion axis, i.e., with $\theta=0$ and $\theta=\pi$) that lie on a sphere which lies in front of the camera share this invariant, i.e., have the same τ_S. The center of the sphere lies on the optical axis of the camera, and the camera lies on the sphere's surface. In this case the diameter of the sphere $\frac{r}{\cos\theta}$ is constant, and so, using Equation (3), $\frac{\dot{\theta}}{\tan\theta}$ remains constant. It has been shown by Raviv in [9] that all points on a particular sphere result in the same visual looming.

4. *The clearance invariant* $\frac{1}{\tau_C}=\frac{\dot{\theta}}{\sin^2\theta}$ (Figure 8): All points in 3-D space (except those that lie on the motion axis, i.e., with $\theta=0$ and $\theta=\pi$) that lie on a cylindrical surface whose axis coincides with the camera translational motion vector share this invariant, i.e., have the same τ_C. In this case the radius of the cylinder $r\sin\theta$ is constant, and so, using Equation (3), $\frac{\dot{\theta}}{\sin^2\theta}$ remains constant. This invariant has been used by Raviv [3] to develop a robust, integration-based, and massively parallel method for reconstructing 3-D scenes. Albus [7] has suggested that this invariant can be used to measure clearance.

5. *The time to contact invariant* $\frac{1}{\tau_P}=\frac{2\dot{\theta}}{\sin 2\theta}$ (Figure 9): All points in 3-D space (except those that lie on the motion axis, i.e., with $\theta=0$ and $\theta=\pi$) that lie on a plane which is perpendicular to the direction of motion of the camera, share this invariant, i.e., have the same τ_P. In this case $r\cos\theta$ (the distance from the surface) is constant, and so, using Equation (3), $\frac{2\dot{\theta}}{\sin 2\theta}$ remains constant.

τ_P is the "time to contact" as described by Lee in [1], and by Lee and Reddish [2]. Lee observed that two macroscopic visual parameters are essential for animals' behavior, the time-to-contact τ and its derivative $\frac{d\tau}{dt}$. He also showed how to derive these parameters from the optical flow. These parameters "tell" the animal how to control its motion in order to avoid collision. Diving birds use it to fold wings for entry into water [2]. τ is specific to the immediacy of contact (i.e., when contact will be made), and $\frac{d\tau}{dt}$ is specific to the harshness of contact (i.e., the type of contact) with the environmental surface. In this paper we referred to the "time to contact" as τ_P and showed its formulation in an image centered spherical coordinate system. More about extraction of τ_P from visual data can be found in [3].

Figure 10 is a summary of the invariants. It shows the basic relationship between space, speed and optical flow (top equation), from which the five invariants are derived. Based on geometrical properties, the time-dependent invariants are shown as a function of optical flow. The geometrical interpretations of all invariants are summarized at the bottom of Figure 10.

4. REPRESENTATIONS USING MOTION INVARIANTS

Two different representations of 3-D space using the invariants are shown in Figure 11. Figure 11a shows a representation that uses the clearance invariant $\frac{1}{\tau_C}=\frac{\dot{\theta}}{\sin^2\theta}$ (Figure 8) and time to contact invariant $\frac{1}{\tau_P}=\frac{2\dot{\theta}}{\sin 2\theta}$ (Figure 9). Figure 11b shows a representation that uses the fixation invariant $\frac{1}{\tau_T}=\dot{\theta}$ (Figure 6) and the looming invariant $\frac{1}{\tau_S}=\frac{\dot{\theta}}{\tan\theta}$ (Figure 7). In both 11a and 11b Figures, the third dimension is Φ.

5. ACKNOWLEDGEMENT

The author would like to thank Marty Herman, Jim Albus, Don Orser and Ernie Kent for fruitful discussions regarding the invariants and some new related representations.

6. REFERENCES

[1] Lee, D.N., "A theory of visual control of braking based on information about time to contact", Perception, No. 5, pp. 437-459, 1976.

[2] Lee, D.N. and Reddish, D.E., "Plumming gannets: a paradigm of ecological optics", Nature 293, pp. 293-294, 1981.

[3] Raviv, D., "Parallel algorithm for 3-D surface reconstruction", In Proceedings of the SPIE conference on intelligent robots, Philadelphia, PA, Nov. 1989.

[4] Raviv, D., "Extracting the 'time to contact' parameter from real visual data", In Proceedings of the SPIE conference on intelligent robots, Philadelphia, PA, Nov. 1989.

[5] Raviv, D. and Herman, M., "Towards an understanding of camera fixation", In Proceedings of the IEEE Conference on Robotics and Automation, Cincinatti, Ohio, 1989.

[6] Raviv, D., "A quantitative Approach to Camera Fixation", In Proceedings of Computer Vision and Pattern Recognition (CVPR) Conference, Hawaii, June 1991.

[7] Albus, J.S., "Outline for a Theory of Intelligence", IEEE Transactions on systems, Man and Cybernetics, Vol. 21, No. 3, 1991.

[8] Cutting, J., *Perception with an Eye for Motion,* MIT Press, 1986.

[9] Raviv, D., "A Quantitative Approach to Looming", NISTIR, in preparation.

[10] Raviv, D., and M. Herman, "A New approach to Vision and Control for Road Following", NISTIR 4476-91, January 91.

[11] Gibson, J.J., "New Reasons for Realism", Synthesis 17, pp. 162-172, 1967.

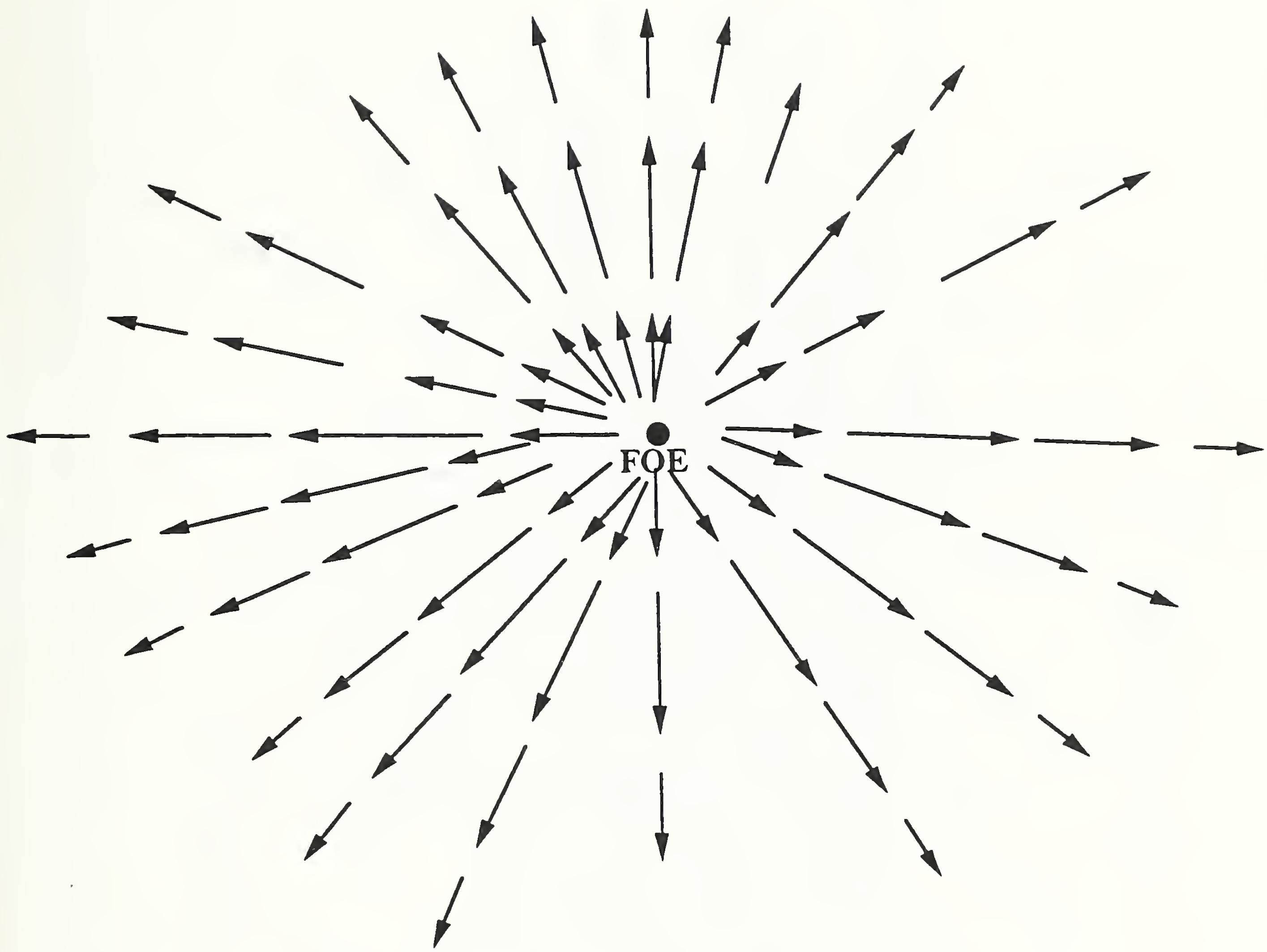


Figure 1a: Optical Flow Relative to the Focus of Expansion (FOE)

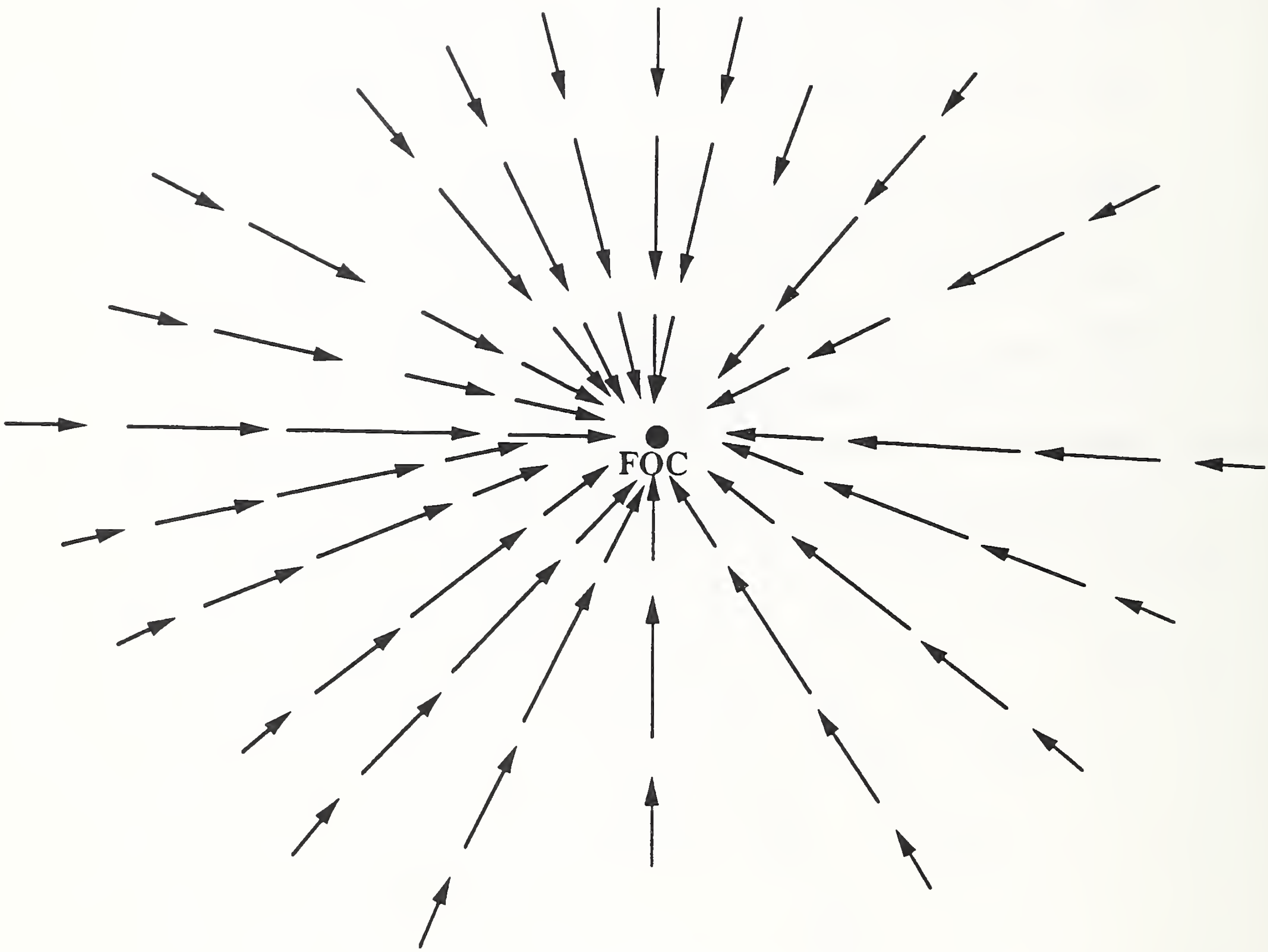


Figure 1b: Optical Flow Relative to the Focus of Contraction (FOC)

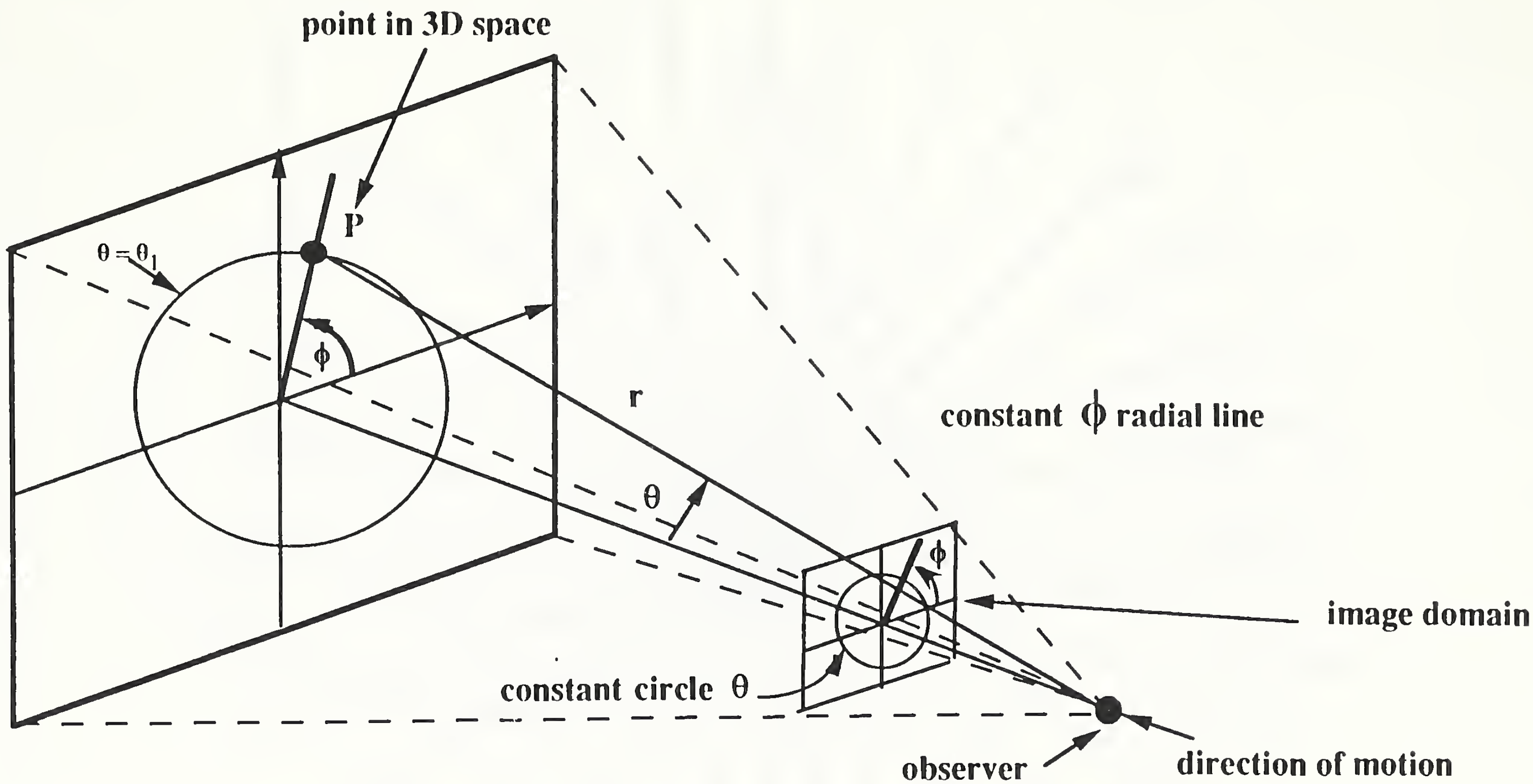


Figure 2: The R–Θ–Φ Coordinate System

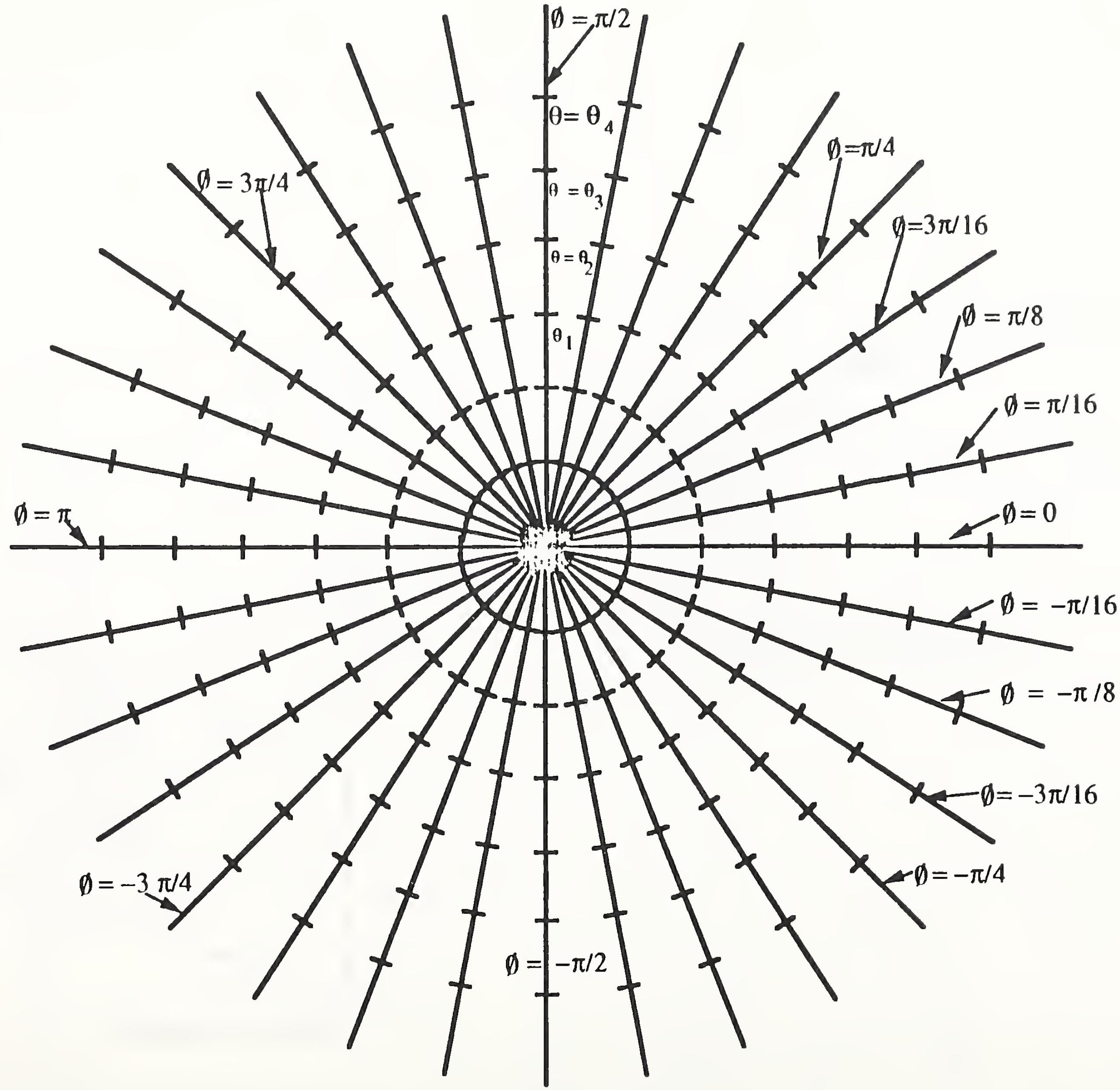


Figure 3: The Θ–Φ Image Domain

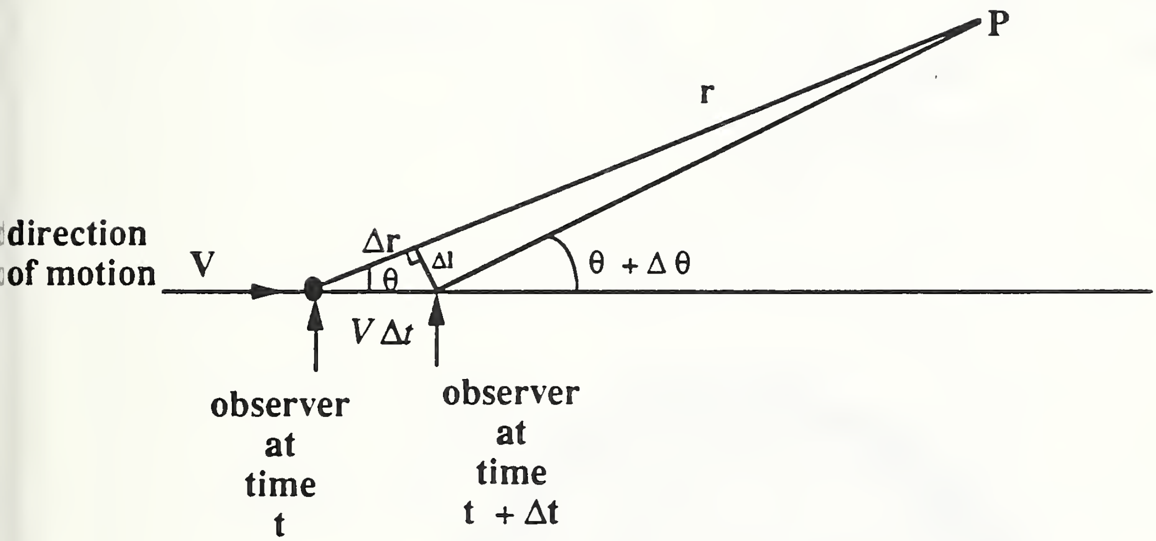


Figure 4: Geometry of Constant Φ Line

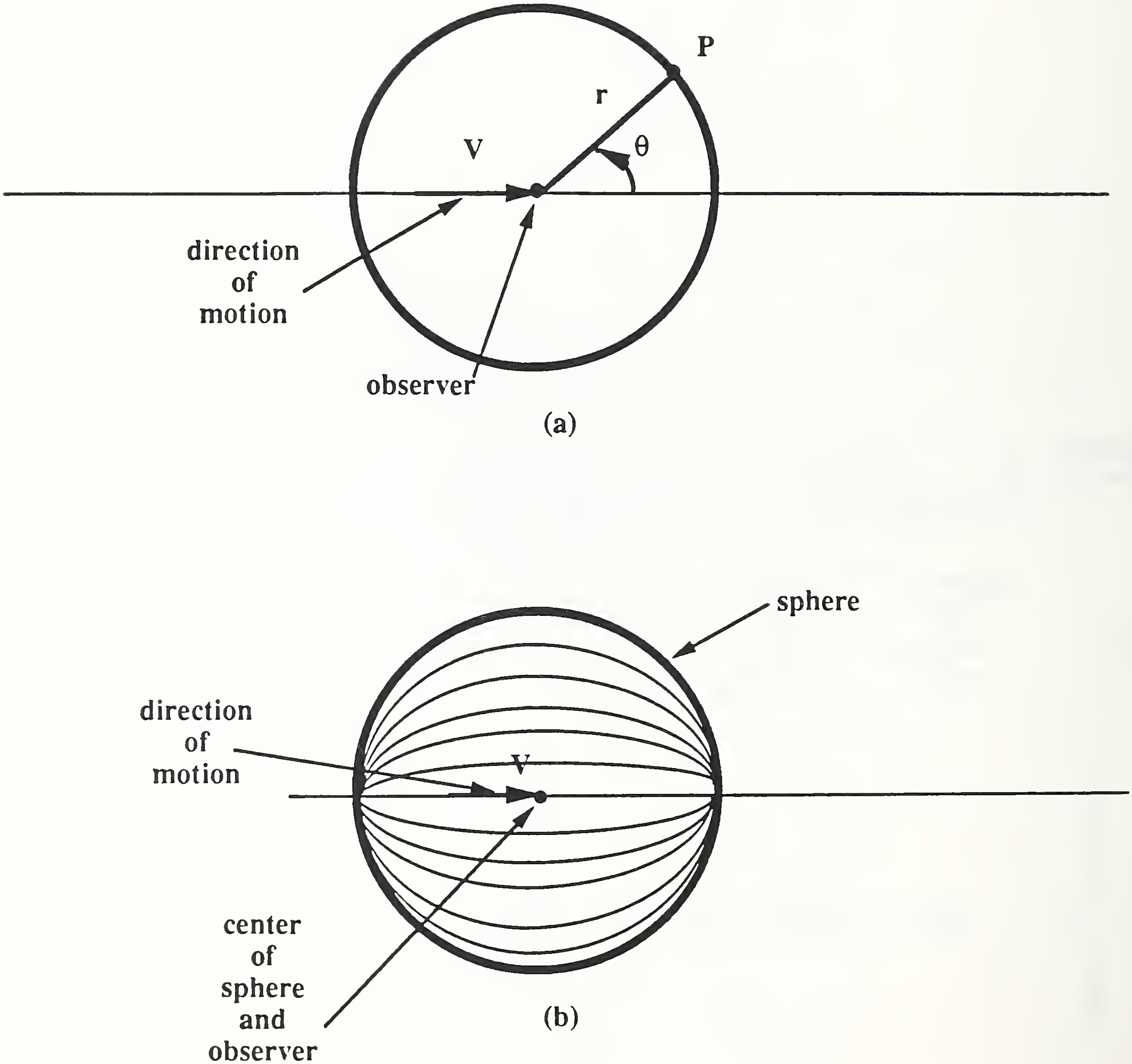
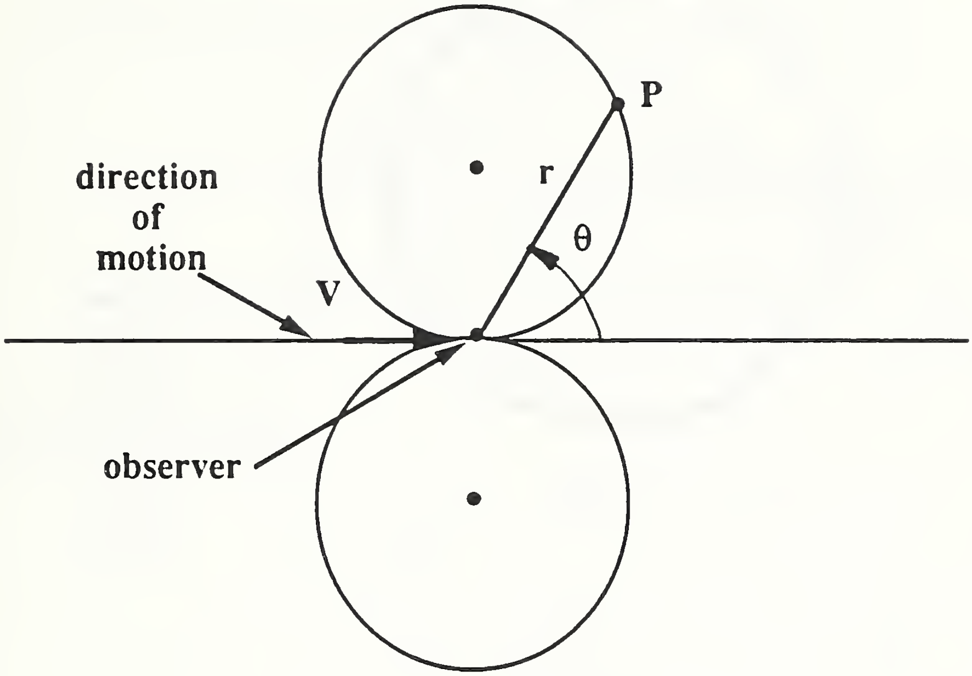


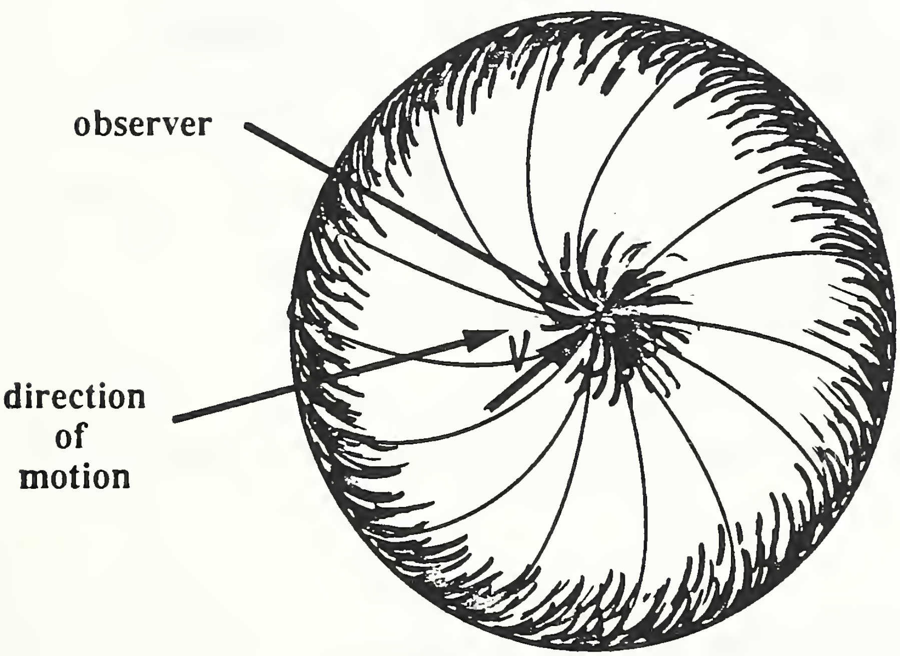
Figure 5: Meaning of the $\frac{1}{T_R} = \frac{\dot{\theta}}{\sin\theta}$ Invariant

(a) In 2-D

(b) In 3-D: Equal distance sphere



(a)

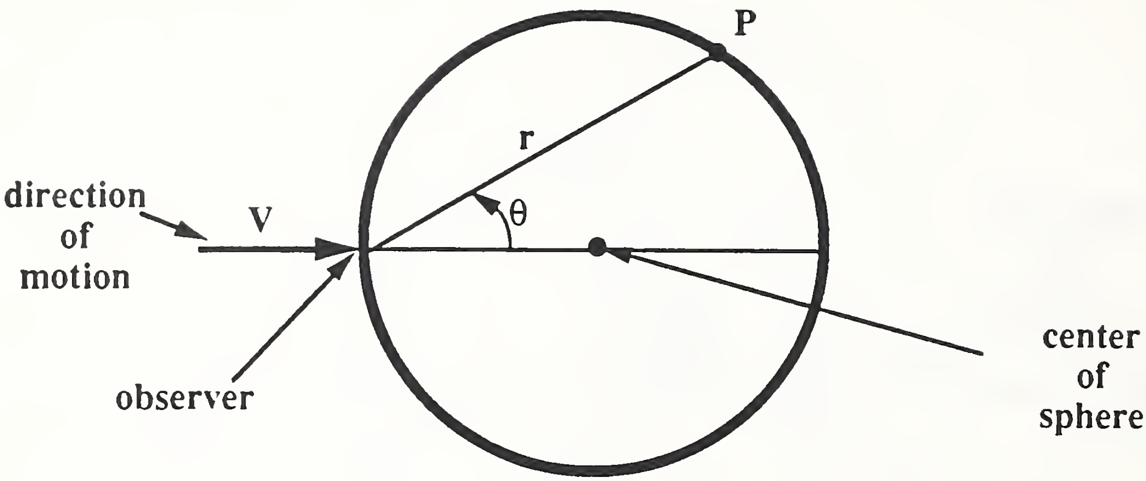


(b)

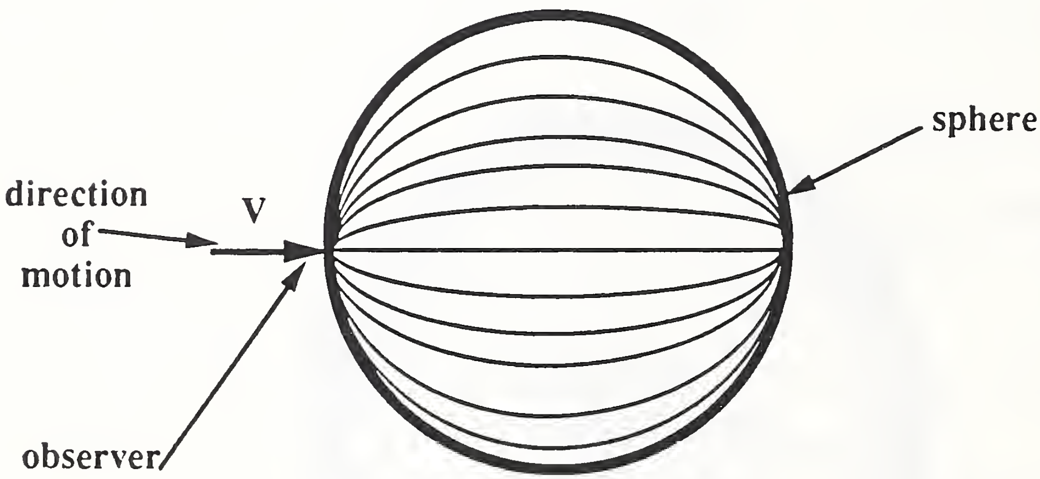
Figure 6: Meaning of the $\frac{1}{T_T} = \dot{\theta}$ Invariant

(a) In 2-D

(b) In 3-D: Equal flow torus



(a)



(b)

Figure 7: Meaning of the $\frac{1}{T_S} = \frac{\dot{\theta}}{\tan\theta}$ Invariant

(a) In 2-D

(b) In 3-D: Equal looming sphere

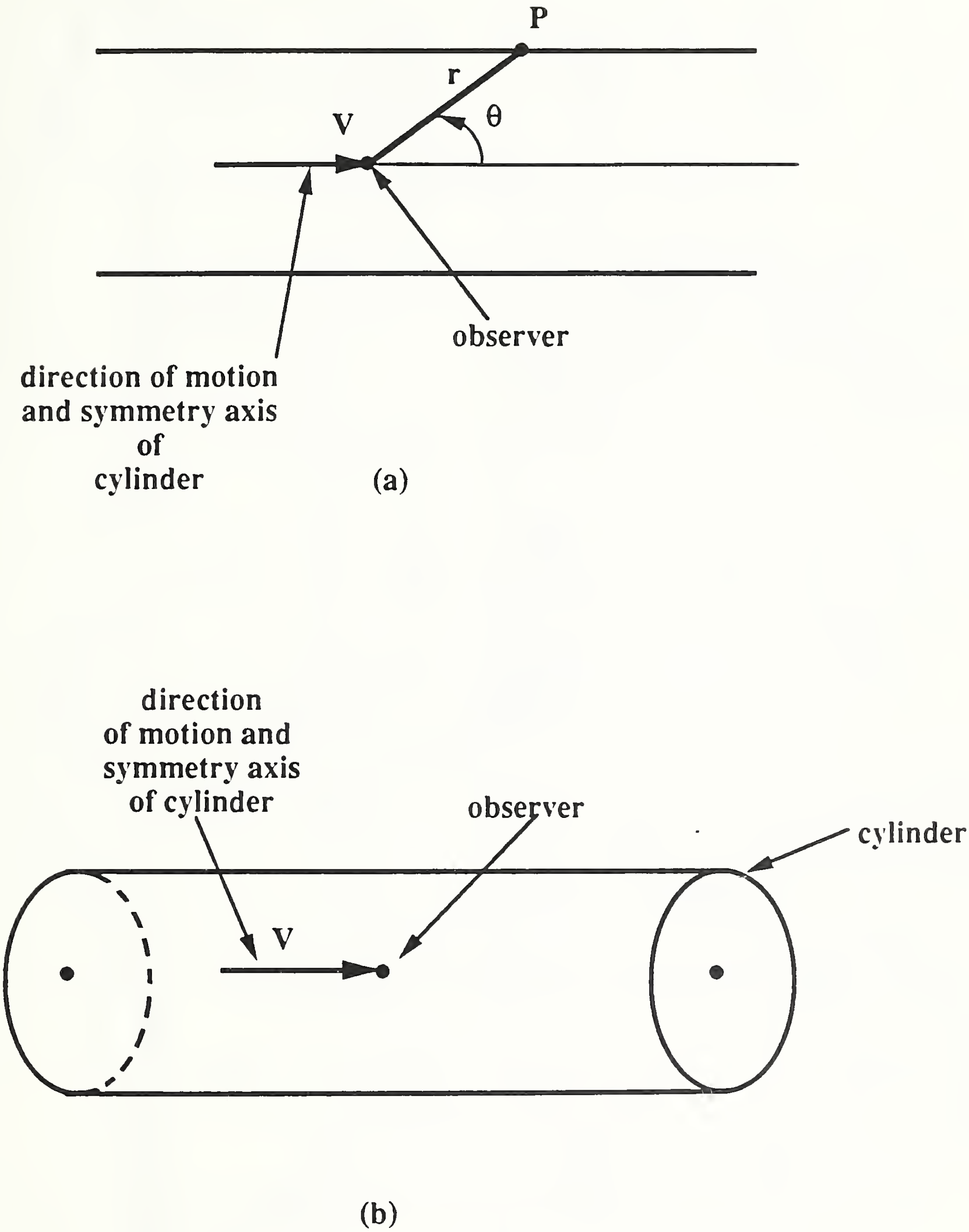


Figure 8: Meaning of the $\frac{1}{T_C} = \frac{\dot{\theta}}{\sin^2\theta}$ Invariant

(a) In 2-D

(b) In 3-D: Equal clearance cylinder

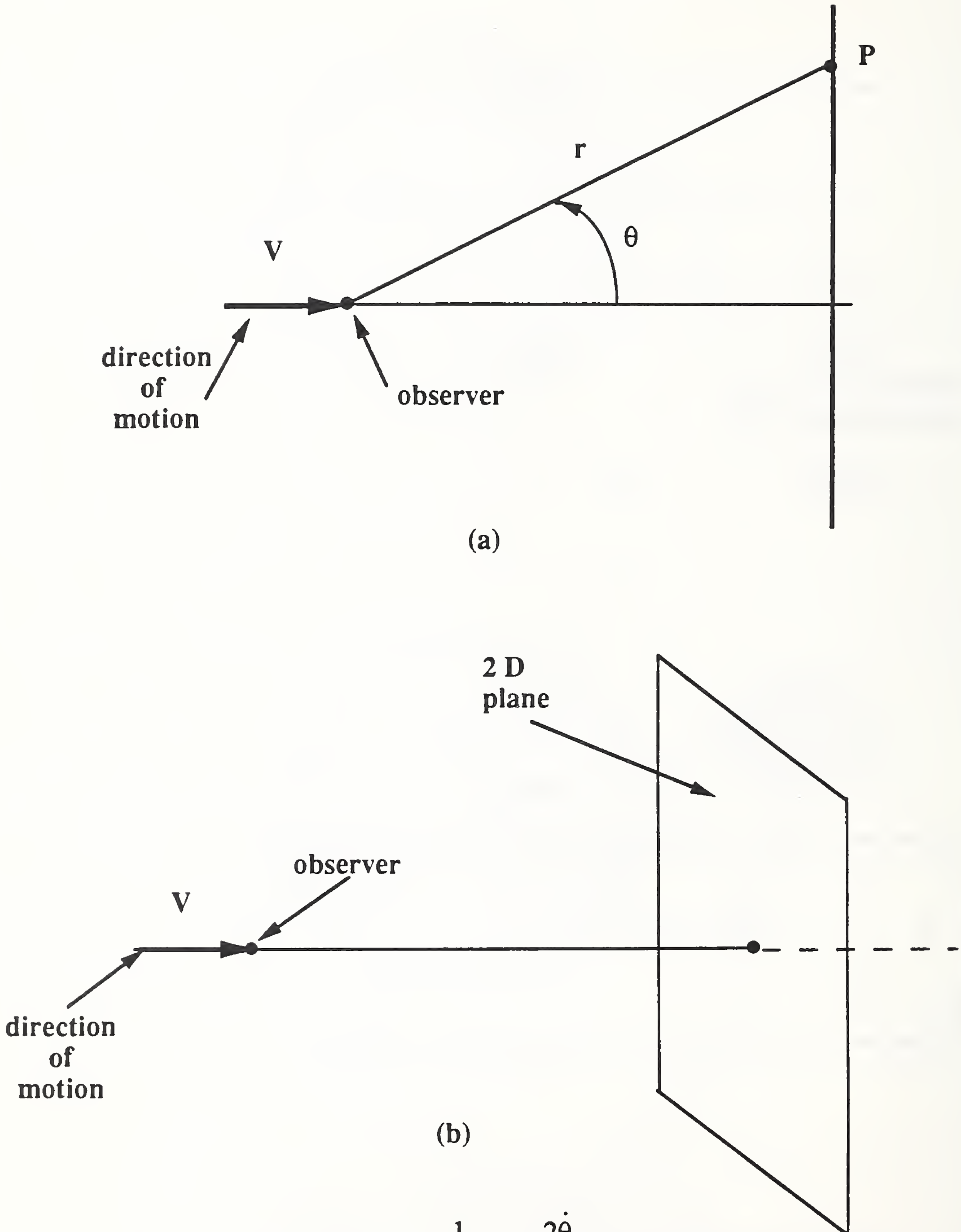


Figure 9: Meaning of the $\frac{1}{T_P} = \frac{2\dot{\theta}}{\sin 2\theta}$ Invariant

(a) In 2-D

(b) In 3-D: Equal depth plane

$V = r \dfrac{\dot{\theta}}{\sin \theta}$					Basic Relationship
$\dfrac{1}{\tau_P}$	$\dfrac{1}{\tau_S}$	$\dfrac{1}{\tau_R}$	$\dfrac{1}{\tau_T}$	$\dfrac{1}{\tau_C}$	notation of invariant
$r \cos \theta$	$\dfrac{r}{\cos \theta}$	r	$\dfrac{r}{\sin \theta}$	$r \sin \theta$	geometrical property of invariance
$\left(\dfrac{\dot{\theta}}{\sin \theta}\right) \dfrac{1}{\cos \theta} = \dfrac{2\dot{\theta}}{\sin 2\theta}$	$\left(\dfrac{\dot{\theta}}{\sin \theta}\right) \cos \theta = \dfrac{\dot{\theta}}{\tan \theta}$	$\left(\dfrac{\dot{\theta}}{\sin \theta}\right) 1 = \dfrac{\dot{\theta}}{\sin \theta}$	$\left(\dfrac{\dot{\theta}}{\sin \theta}\right) \sin \theta = \dot{\theta}$	$\left(\dfrac{\dot{\theta}}{\sin \theta}\right) \dfrac{1}{\sin \theta} = \dfrac{\dot{\theta}}{\sin^2 \theta}$	invariant in terms of optical flow

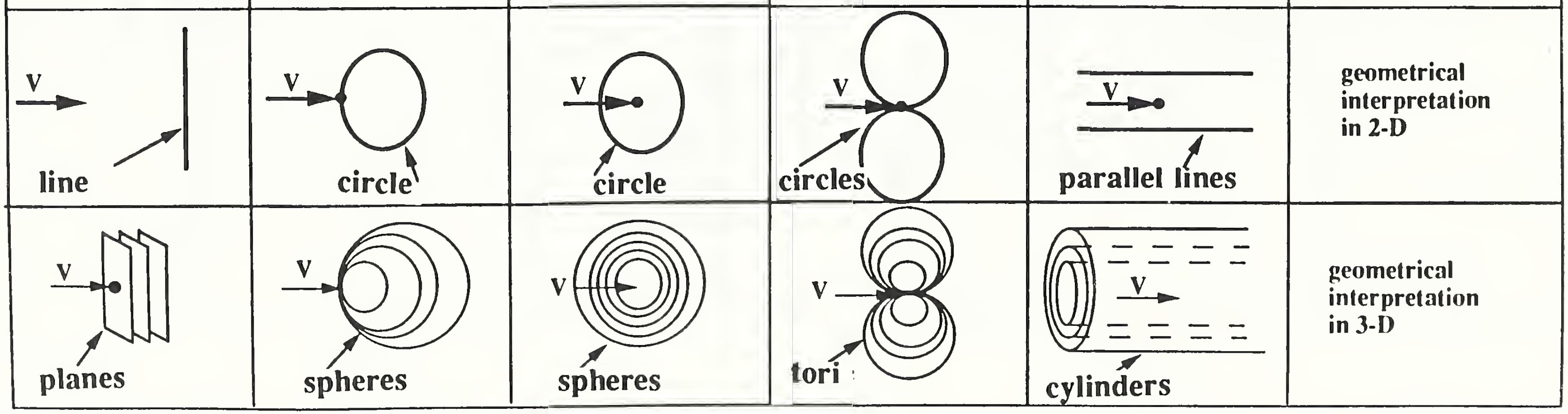


Figure 10: Summary of the Invariants

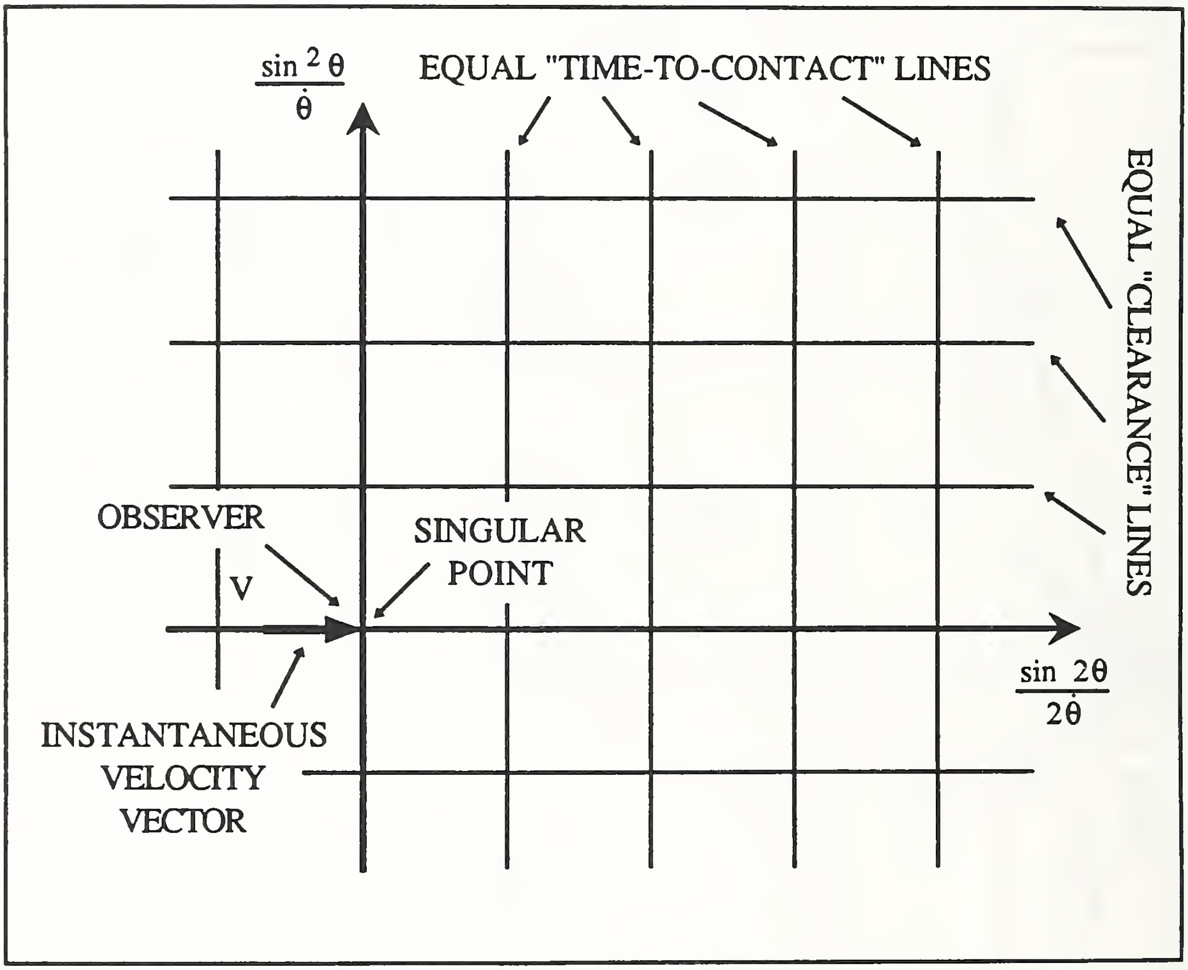


Figure 11a: Representation 1

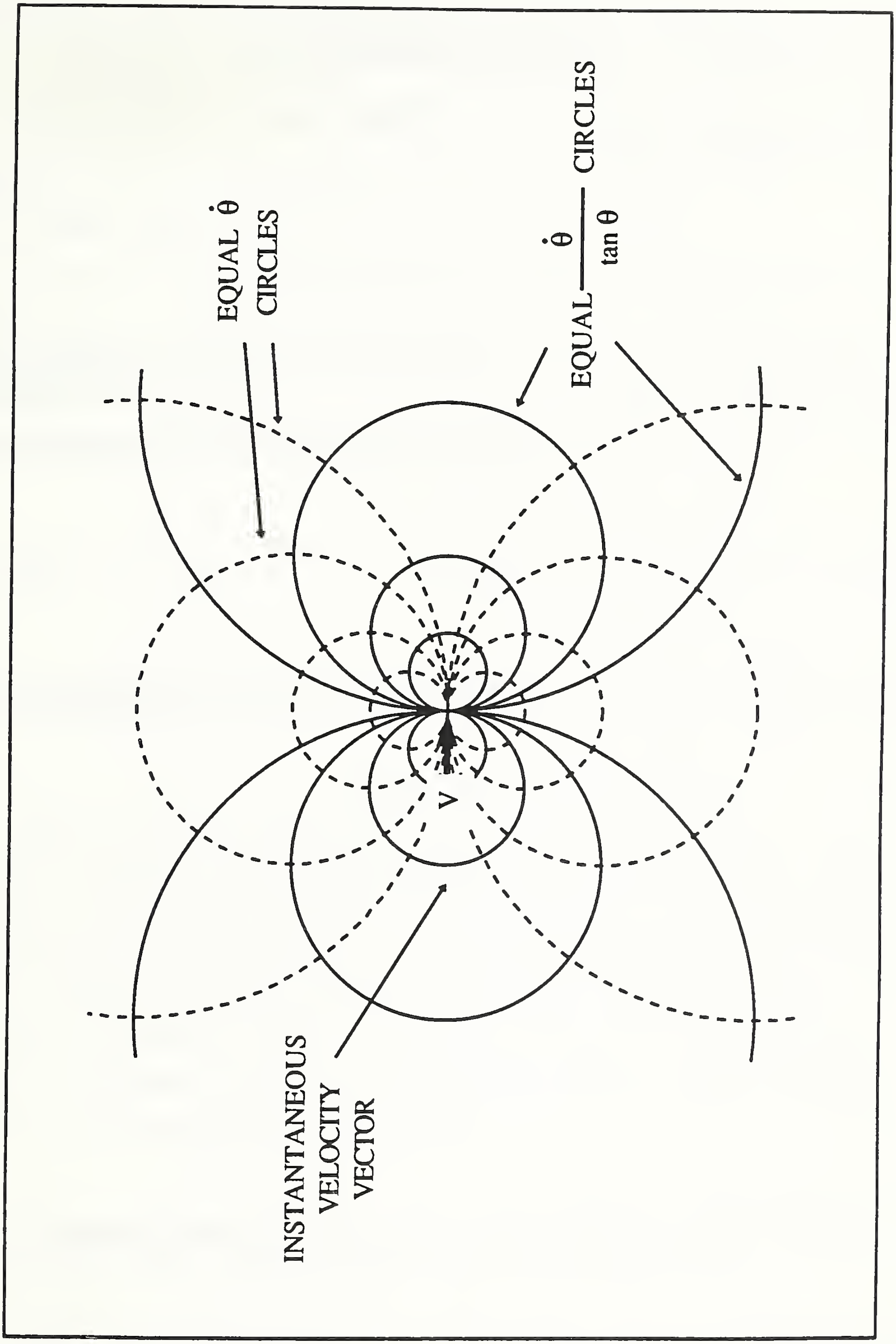


Figure 11b: Representation 2

NIST-114A
(REV. 3-90)

U.S. DEPARTMENT OF COMMERCE
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY

BIBLIOGRAPHIC DATA SHEET

1. PUBLICATION OR REPORT NUMBER
NISTIR 4722

2. PERFORMING ORGANIZATION REPORT NUMBER

3. PUBLICATION DATE
November 1991

4. TITLE AND SUBTITLE

Invariants in Visual Motion

5. AUTHOR(S)

Daniel Raviv

6. PERFORMING ORGANIZATION (IF JOINT OR OTHER THAN NIST, SEE INSTRUCTIONS)

U.S. DEPARTMENT OF COMMERCE
NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
GAITHERSBURG, MD 20899

7. CONTRACT/GRANT NUMBER

8. TYPE OF REPORT AND PERIOD COVERED

9. SPONSORING ORGANIZATION NAME AND COMPLETE ADDRESS (STREET, CITY, STATE, ZIP)

10. SUPPLEMENTARY NOTES

11. ABSTRACT (A 200-WORD OR LESS FACTUAL SUMMARY OF MOST SIGNIFICANT INFORMATION. IF DOCUMENT INCLUDES A SIGNIFICANT BIBLIOGRAPHY OR LITERATURE SURVEY, MENTION IT HERE.)

ABSTRACT

This paper deals with new and simple representations of 3-D points in a moving-observer coordinate system. Assuming rectilinear motion with no rotation of an observer where the optical axis coincides with the direction of motion, and a stationary scene, points in 3-D space that lie on a particular 3-D surface produce constant value of some nonlinear function of the measurable image optical flow. Five sets of different surfaces are introduced and there is one optical-flow based constant value for each surface. We called these values "invariants". It is shown how to extract these invariants and how to use them for representing 3-D space.

12. KEY WORDS (6 TO 12 ENTRIES; ALPHABETICAL ORDER; CAPITALIZE ONLY PROPER NAMES; AND SEPARATE KEY WORDS BY SEMICOLONS)
Visual Motion; Robot Vision; Visual Invariants; Visual Representations.

13. AVAILABILITY

- [X] UNLIMITED
- [ ] FOR OFFICIAL DISTRIBUTION. DO NOT RELEASE TO NATIONAL TECHNICAL INFORMATION SERVICE (NTIS).
- [ ] ORDER FROM SUPERINTENDENT OF DOCUMENTS, U.S. GOVERNMENT PRINTING OFFICE, WASHINGTON, DC 20402.
- [X] ORDER FROM NATIONAL TECHNICAL INFORMATION SERVICE (NTIS), SPRINGFIELD, VA 22161.

14. NUMBER OF PRINTED PAGES
23

15. PRICE
A02

ELECTRONIC FORM