N92-21471

THE PERCEPTION OF SURFACE LAYOUT DURING LOW LEVEL FLIGHT

John A. Perrone
NASA Ames Research Center
Moffett Field, California and
Stanford University
Stanford, California

INTRODUCTION

Low level flight in helicopters presents a particularly challenging visual situation to the pilot, especially if his normal field of view is degraded or restricted by the use of night-vision aids. The pilot must somehow extract information about his heading direction and the layout of the environment in front of him from the two-dimensional pattern of light reaching his retina. Humans can perform this task well in most situations, but low-level flight taxes this ability to the limit, because the time to respond to obstacles is greatly shortened and there is often a large rotational component to the observer platform that is not normally present in normal ambulatory motion.

There are many sources of information that a pilot can use to infer the three dimensional environmental layout from the two-dimensional images that feed into his visual system. This paper will concentrate on one of these sources, namely the 2-D motion flow-field that is generated during observer motion. We are interested in how the pilot can infer his heading and the surface layout from these 2-D velocity flow fields. This is part of the more general "structure from motion problem" and the special case of observer motion has been labelled the structure from ego-motion problem.

There now exists many theoretical treatments of this problem, which demonstrate how 3-D information can be extracted from the two-dimensional motion field. (e.g. [1, 2, 3, 4, 5]). However, explicit theories as to how human observers actually extract this information are not so common. Many of the above theories also assume that the 2-D flow field is available, but this stage can represent a major stumbling block in trying to solve the structure from motion problem [6].

This problem is obviously very relevant to the visually guided control of movement. It is a very difficult problem, however, and one which may not be fully understood for some time. A complete understanding of the process would help detect potentially dangerous situations (illusions) and help in the design of display systems such as FLIR. The designers could economize on features that were known to be unimportant to the control problem and make more informed decisions about parameters such as the field of view and the amount of visual noise that can be tolerated in these displays.

Obviously visual motion is an important source of information for man and many animals [7, 8, 9, 10]. However, in the context of rotorcraft flight the question remains as to just how much information about surface layout and obstacles can be derived from motion cues ? We know that heading information can be gained from the motion field [7, 11] and that relative motion and parallax can provide important information about depth discontinuities [1, 12]. What is not so well established concerns the case of an observer moving forward with a fixed line of sight. This is supposedly the simplest situation for the structure from motion problem (since there is no rotational component), yet it not obvious in this case as to how well the environmental layout can be inferred on the basis of motion information alone. If a pilot is viewing FLIR imagery of a very unstructured scene (e.g. foliage or trees) such that the usual cues of linear perspective, intensity gradients and size gradients are missing, how well can they perceive the scene layout on the basis of motion alone ? It is a useful exercise to examine some of the issues involved in this relatively simple situation. From this we can begin to see some of the important variables that need to be considered and which factors are worth manipulating experimentally. It also helps define the scope of the problem and puts a cap on what can realistically be obtained from machine vision applications such as remote sensors for autonomous flight. Identifying potential problem areas for computer vision systems can provide useful insights into what may also prove problematical to humans.

EFFECTIVE RANGE OF MOTION INFORMATION

One point that is often ignored is that the effective range over which motion information is even detectable is quite limited. There has been some work on this area in relation to the extraction of heading direction [11, 13] but this hasn't been carried over much to the structure from motion problem. Figure 1 shows the theoretical flow field for a craft moving at 3 eyeheights/sec over a flat plane. The figure shows a .25 sec "snapshot" of the motion field. Each vector is for a point on a square grid with the inter-point distance equal to 1 eyeheight. The thing to notice is that the length of the velocity vectors fall off rapidly with distance. If for simplicity we limit ourselves to points lying along the median plane, the angular velocity of the points can be found from:

$$\dot{EL} = -\frac{\dot{x}}{z}\sin^2 el \tag{1}$$

where is the forward velocity parallel to the ground plane, z is the height of the observer above the ground plane, and el is the elevation angle (measured from the horizon) of the point on the ground. (Using equation 22 from Warren [14] with = 0). Substituting z/ for sin(el) we have:

$$\dot{EL} = -\dot{x}\frac{z}{x^2 + z^2} \tag{2}$$

This shows that angular velocity basically falls off as the square of the distance from the observer. The angular velocity is also a function of the height of the observer, but the effect of the z term is less than the distance factor. Figure 2 shows a plot of angular velocity (in min of arc/sec) against distance (in height units) for an eyepoint moving at 3 eye-heights/sec. Points higher up in the field (smaller z) will have an even smaller value of and would lie below the curve shown in the figure. It is difficult to set a threshold level for velocity detection and it is based on many factors [15]. However, if we use the value based on practical complex situations [16] i.e 40 min of arc/sec we find that absolute velocity information is becoming subthreshold at about 15 to 16 eyeheight units away. This is the length of the "headlight beam" defined by motion information alone. At a speed of 3 eyeheight units/sec, this only gives about 5 seconds to respond to features on the ground that are revealed by the motion process. These critical distances and times are shorter for objects above the ground that lie closer to the eye-level plane.

This analysis is only supposed to provide a qualitative feeling for the limitations to the structure from egomotion problem. It ignores the issue of how the structure is actually extracted and is based on absolute motion thresholds rather than the perhaps more relevant relative motion thresholds. However it does help to limit the domain over which motion information can be considered to contribute to the perception of surface layout. It could also be an important factor when the visual motion information is being relayed to the pilot via some form of CRT display device, such as those used in FLIR systems. The limited resolution of these displays degrades the motion information even further and thus puts a further cap on the effective range of the structure from motion process.

This limited range of utility for motion information also becomes an important issue when certain environmental features are present in the field of view. Most terrain is not perfectly flat like the example shown in Figure 1. Rather NOE flight is often over sloping terrain with many hills and valleys. The correct perception of the slope and orientation of this terrain is important to the pilots perception of his own spatial orientation. Any misperception of surface layout can affect the flight path chosen by the pilot and, in theory, could lead to disorientation in some cases.

It is therefore of interest to examine the motion information that is available during flight towards sloped terrain.

Motion Information and Sloped Terrain

It turns out that this situation generates an interesting pattern of flow information and raises important theoretical issues as to how humans actually infer layout from the motion flow field. Figure 3 shows the theoretical flow field for pure translatory motion towards a planar hill, slanted relative to a ground plane. The angular velocity of points on the hill decrease as a function of distance from the observer as before, but also as a function of height in the field. This is the z term in the numerator and denominator of equation 2 above. The z term decreases as a function of the distance as well as a function of the tangent of the slant angle of the hill. This means that for situations such as that shown in Figure 3, the angular velocity is low over a large area of the field, especially in the important region directly ahead of the observer. This makes it difficult for any system attempting to infer the slope of the terrain using the angular velocity of points in the field.

There are many ways by which the slant could be recovered from the motion information. It is interesting to consider some of the techniques that an artificial vision system could use to tackle this problem. One technique would be to assign a depth value to each of the points based on its angular velocity (i.e. the length of the vectors in fig 3). Since we assume that the actual observer speed is unknown, this amounts to the derivation of a relative range map, or a time to impact map which is independent of the actual speed [17]. The slant would then be found by "fitting" a plane to the distribution of relative distances derived from the impact times. However in order to convert the angular velocities into impact times, the magnitude of each velocity vector must be divided by a factor proportional to the square of the distance in the image plane, of the point from the focus of expansion.

Because the angular velocities are so small for a large proportion of the points in the field, we cannot expect the impact time map to be very accurate in these regions, especially when the velocity estimates are noisy.

The other general approach would be to use relative velocities (e.g. [1]) and find the slant of a local patch and then integrate over the entire surface. However since this requires taking differences of already small 2-D velocity estimates, we can again expect errors to occur over a large part of the field. The problem is that the relative difference in velocities in small local patches can be quite small, and it is only when the differences in velocities of very disparate points are considered, that such differences may reach threshold. Techniques which use local differencing methods are good for finding depth discontinuities in the field, but with slanted surfaces the change can be very gradual.

There is an interesting parallel between the problem of judging slant from motion with the problem of detecting slant using stereoscopic vision. Gillam et al., [18] have tried to argue that stereoscopic process requires changes (2nd derivatives) in the slant to perceive the slant correctly. They found that subjects took very long times to accurately judge the slant of flat planes, but much shorter times if the plane contained changes in slant or a discontinuity in depth. If it can be shown that the structure from motion system is also dependent upon such changes, then we can expect problems in situations where they are absent or subthreshold. Fortunately, most natural scenes and terrain are not perfectly flat like figure 3. There are instances however, such as with snow covered terrain, where many of the small details and features which can provide information about changes in slant are lacking. In such situations, the unique pattern of velocity vectors produced by slanted terrain during forward translation may become important. Relatively small differences in speed are produced locally and this could result in problems for a system that is designed for the detection of the large local speed differences. Such large differences are much more common in the visual field since they result whenever we move through and environment made up of objects occupying different depth planes.

There is more to the problem of detecting surface layout in the presence of slant however, than just the slow change in speed values. Braunstein, [19] showed that surface slant could be judged reasonable accurately for surfaces slanted only 20 to 30 deg from vertical. In this case, the local differences in speed are small. However Braunstein used motion parallel to the image plane so the the image motion was completely unidirectional. With forward motion, the velocity flow field consists of vectors of many different directions. Figure 4 shows what the flow field for the surfaces in figure 3 would look like in the case of translation parallel to the image plane. We are currently testing experimentally whether the difference in the flow fields between the two situations can account for any differences in the ability to extract surface slant under the two types of translation.

Comparison of the two flow fields in Figs 3 and 4 draws attention to another interesting feature. In the case of forward translation (fig 3) the perspective indicated by the vectors conflicts with the perspective indicating the static layout of the surfaces (the rectangular layout of the points helps to define the static perspective cues). The pattern defined by the motion vectors is similar to what would be obtained if a curved surface was covered with poles set normal to the surface. In the case of motion towards a flat plane (wall, cliff) (figure 5) the perspective suggested by the vectors is similar to that produced by a forward slanting surface like a ceiling.

If a display is used with a slow phosphor such that streaking of the image is present, one would predict some misperceptions of the surface layout to occur under these situations. We have noticed exactly these effects in our experimental displays when the image contrast is high and the room illumination low. It would be of interest to see if any of these effects can occur in cockpit mounted displays or FLIR displays. This is an instance of motion interacting with the pattern processing system. Motion artifacts such as streaking seem to be creating a misperception of surface orientation. This introduces a slightly different question: Can motion override a familiar "illusion" of surface orientation that occurs under static viewing conditions?

Motion and Surface Slant Underestimation

Many situations have been identified in which surface orientation under static viewing conditions, can be greatly misperceived by observers. Some of the stimulus features that contribute to these errors have been identified [20, 21, 22, 23]. One particular situation relevant to N.O.E flight conditions is the perceived slope (slant) of the terrain immediately in front of the craft. Under both laboratory and environmental testing conditions the slant of a surface is, in most cases, perceived to be closer to the observers frontoparallel plane than its true position [20, 21, 23]. This has mainly been tested under monocular static conditions although there is some evidence that slant underestimation still occurs for stereoscopic displays [24].

One question that has not been addressed fully in the research literature in this field is whether or not motion information can override this tendency to misperceive the slant of surfaces. For stimulus displays in which the motion is parallel to the image plane, Braunstein, [19] has shown that motion information provides strong cues for slant and that very little underestimation occurs under the motion conditions. It is not clear whether or not this same state of affairs would exist for the case of motion towards the surface. We have begun a series of experiments aimed at answering this question.

Although it is fairly well established now that much information about surface layout can be gained from motion cues, it is not so clear as to what information humans can use and what specific information they should be provided with. The various theoretical analyses tell us that the information is there in the stimulus. It will take many more experiments to verify that this information can be used by humans to extract surface layout from the 2-D velocity flow field. Pilots obviously can use the information efficiently in most situations. This paper has tried to draw attention to some of the visual motion factors that can affect the pilot's ability to control his craft and to infer the layout of the terrain ahead of him.

REFERENCES

1. Clocksin, W. F. Perception of surface slant and edge labels from optical flow: a computational approach. 9: 253-269, 1980.

2. Longuet-Higgins, H. C. A Computer Algorithm for Reconstructing a Scene From Two Projections. 293: 133-135, 1981.

3. Prazdny, K. Determining the Instantaneous Direction of Motion From Optical Flow Generated by a Curvilinearly Moving Observer. : 109-114, 1981.

4. Tsai, R. Y. and T. S. Huang. Uniqueness and Estimation of 3-D Motion Parameters and Surface Structures of Rigid Objects. Computer Science. : 1-34, 1983.

5. Zacharias, G. L., A. K. Caglayan and J. B. Sinacori. A visual cuing model for terrain-following applications. J. Guidance. 8(2): 201-207, 1985.

6. Perrone, J. A. In search of the elusive flow field. Proceedings of the IEEE Workshop on Visual Motion. 181-188, 1989.

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

8. Nakayama, K. Biological image motion processing: a review. Vision Res. 25: 625-660, 1985.

9. Hildreth, E. C. and C. Koch. The Analysis of Visual Motion: From Computational Theory to Neuronal Mechanisms. Ann. Rev. Neurosci. 10: 477-533, 1987.

10. Horridge, G. A. The evolution of visual processing and the construction of seeing systems. Proc. R. Soc. Lond. 230: 279-292, 1987.

11. Cutting, J. E. "Perception with an eye for motion." 1986 Bradford. Cambridge.

12. Nakayama, K. and J. M. Loomis. Optical velocity patterns, velocity-sensitive neurons, and space perception: A hypothesis. Perception. 3: 63-80, 1974.

13. Palmer, E. A. Experimental Determination of Human Ability to Perceive Aircraft Aim Point From Expanding Gradient Cues. 40th Annual Scientific meeting of the AMA. 176-177, 1969.

14. Warren, R. Optical Transformation During Movement: Review of the Optical Concomitants of Egomotion. O.F.O.S.R. report 81-0108. 1982.

15. Salvatore, S. The Perception of Real Motion A Literature Review. Injury Control Research Laboratory Report ICRL-RR-70-7. 1972.

16. Michaels, R. M. and L. W. Cozan. Perceptual and field factors causing lateral displacements. Highway Research Record. 25: 1963.

17. Lee, D. N. "Visual information during locomotion." Perception: Essays in honor of J.J. Gibson. Macleod and Pick ed. 1974 Cornell U. Press. Ithaca.

18. Gillam, B., T. Flagg and D. Finlay. Evidence for disparity change as the primary stimulus for stereoscopic processing. Perception and Psychophysics. 36: 559-564, 1984.

19. Braunstein, M. L. Motion and texture as sources of slant information. J. Exp. Psychology. 78: 247-253, 1968.

20. Gibson, J. J. The Perception of Visual Surfaces. Am. J. of Psych. 63: 367-384, 1950.

21. Clark, W. C., A. H. Smith and A. Rabe. The Interaction of Surface Texture, Outline Gradient, and Ground in the Perception of Slant. Canad. J. Psychol. 10(1): 1-8, 1956.

22. Gogel, W. C. Equidistance Tendency and its Consequences. 64(3): 153-163, 1965.

23. Perrone, J. A. Visual slant underestimation: a general model. Perception. 11: 641-654, 1982.

24. Gillam, B. J. Perception of Slant when Perspective and Stereopsis Conflict: Experiments with Aniseikoniclenses. J. of Exp. 78: 299-305, 1968.

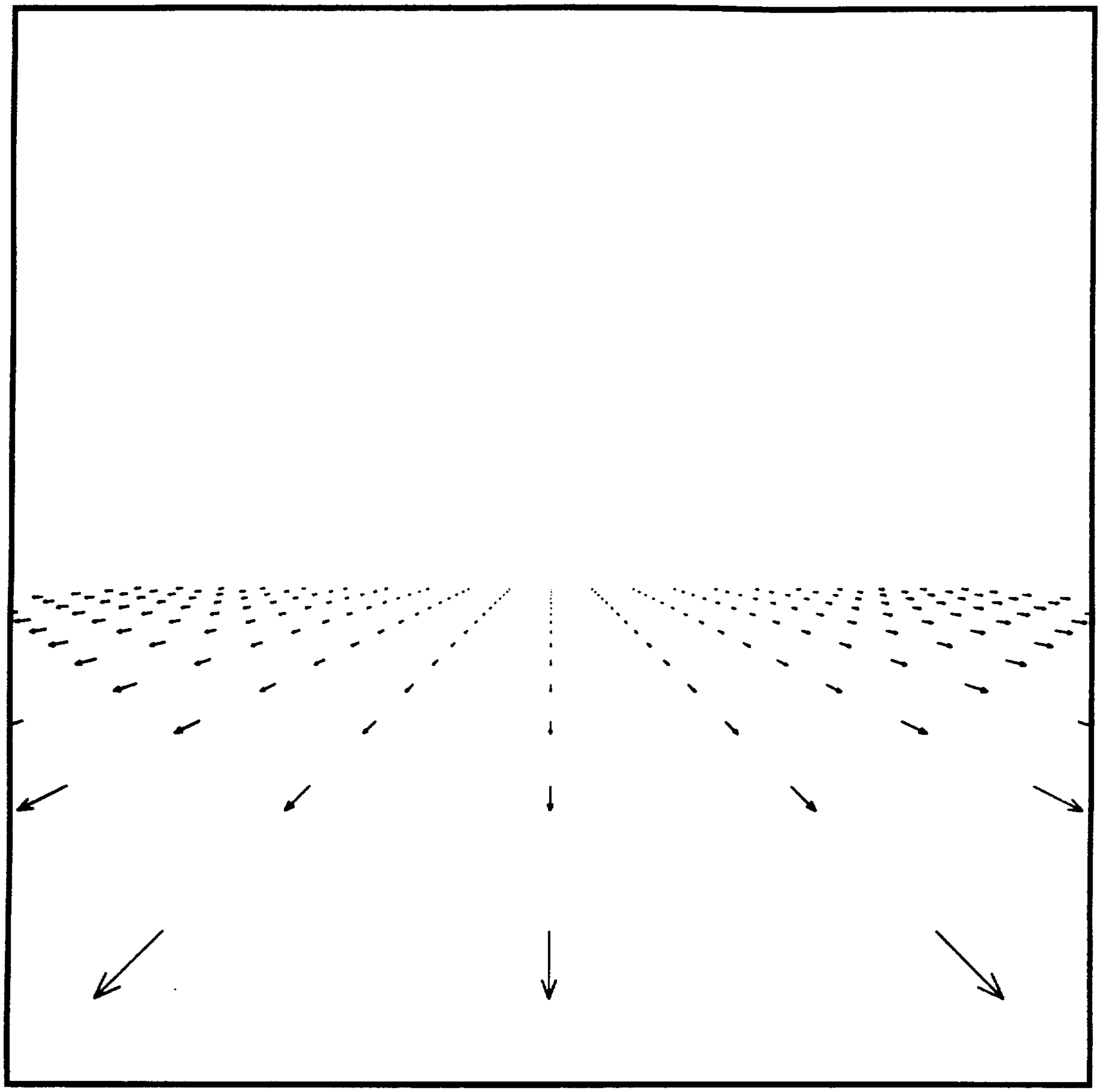

Figure 1. Theoretical flow field for motion over a flat ground plane.

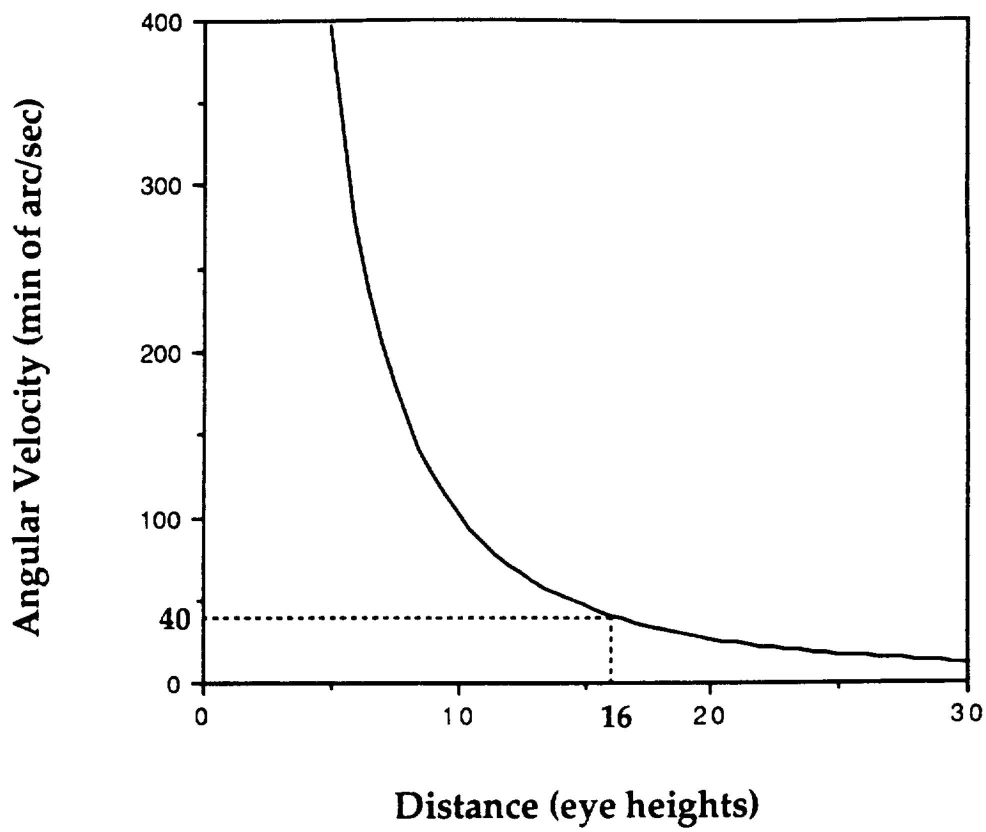

Figure 2. Plot of angular velocity versus distance along ground plane.

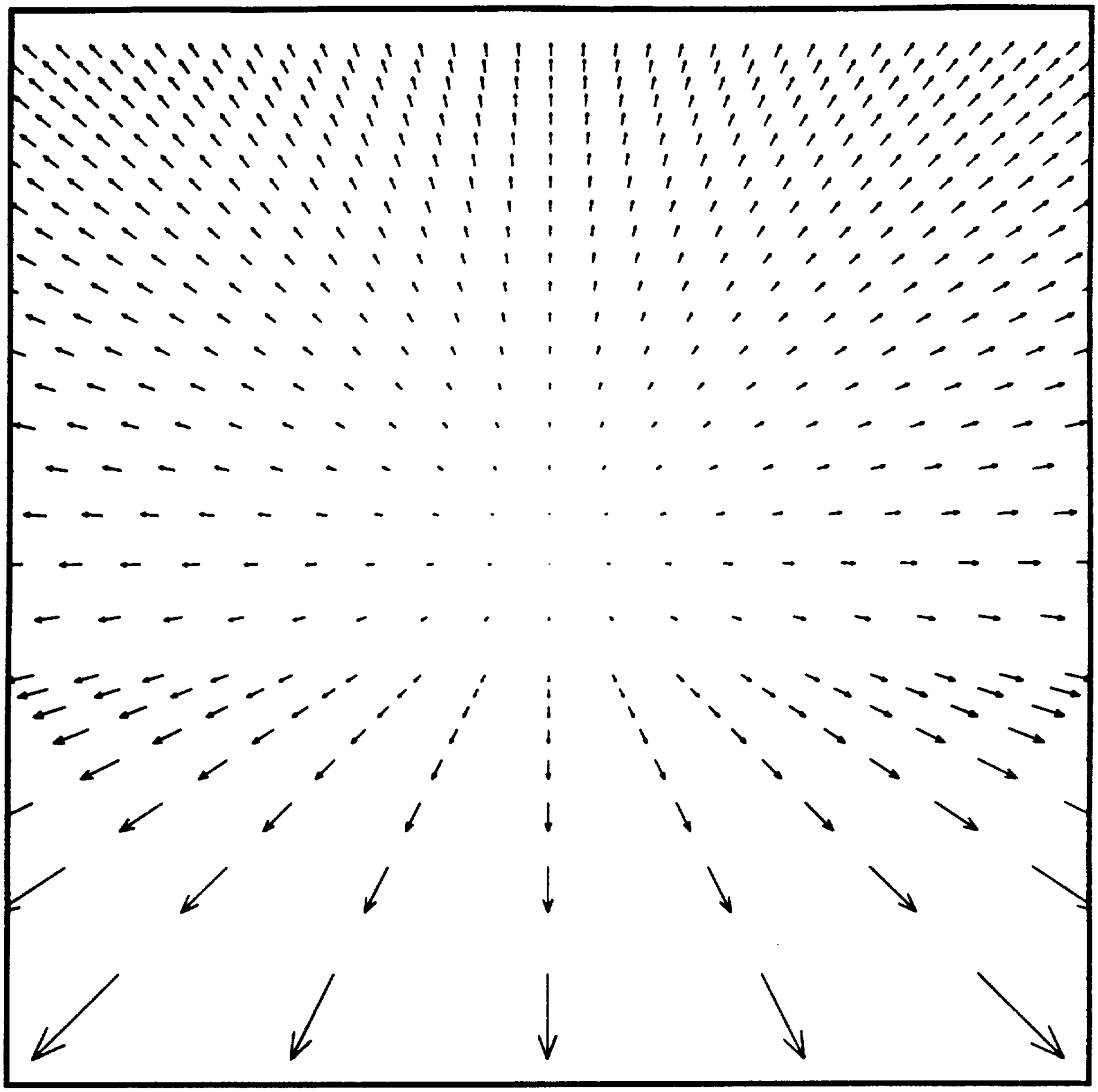

Figure 3. Flow field for motion toward a plane slanted 60 degrees from the horizontal.

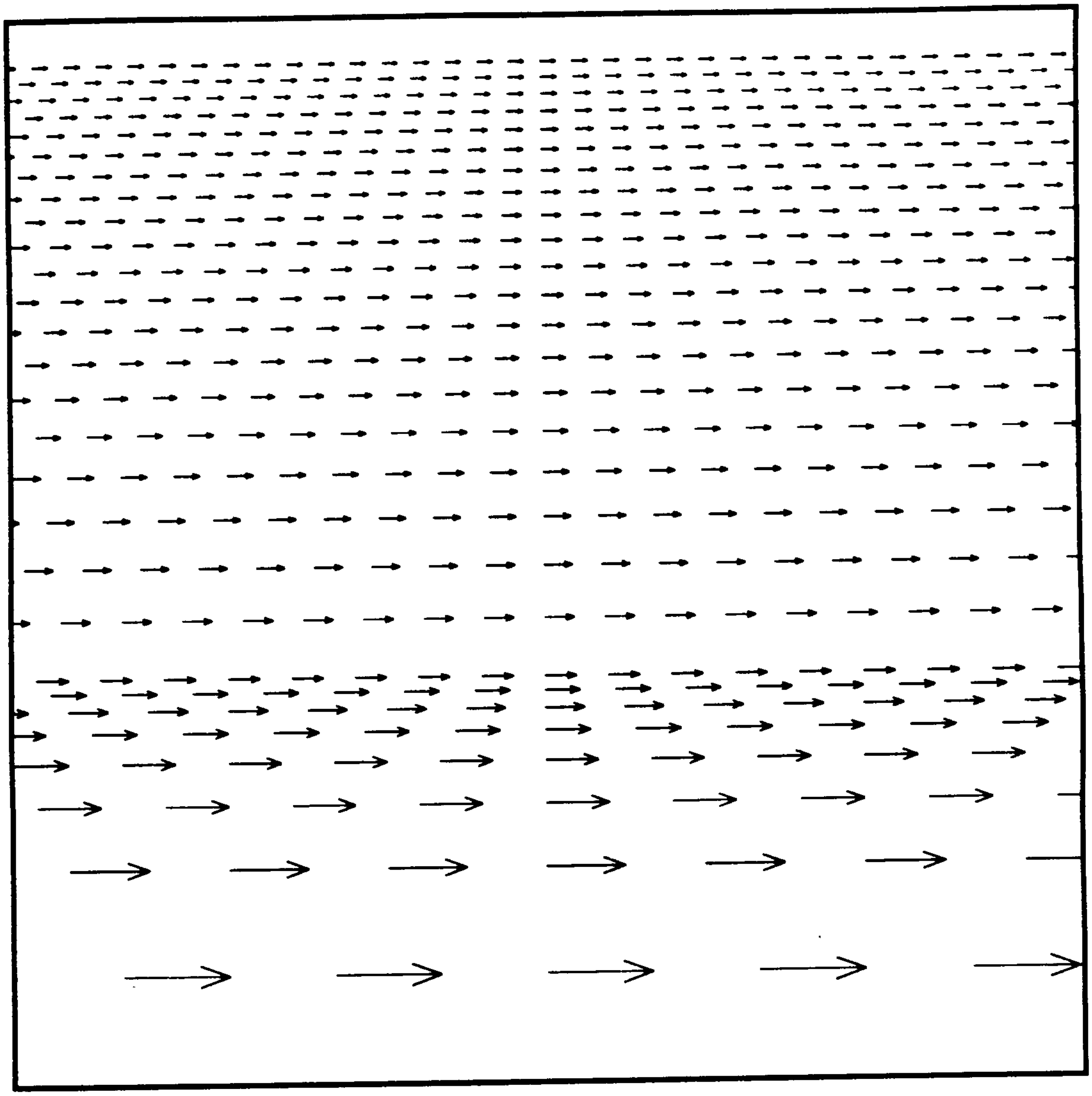

Figure 4. Flow field for slanted plane when motion is parallel to image plane.

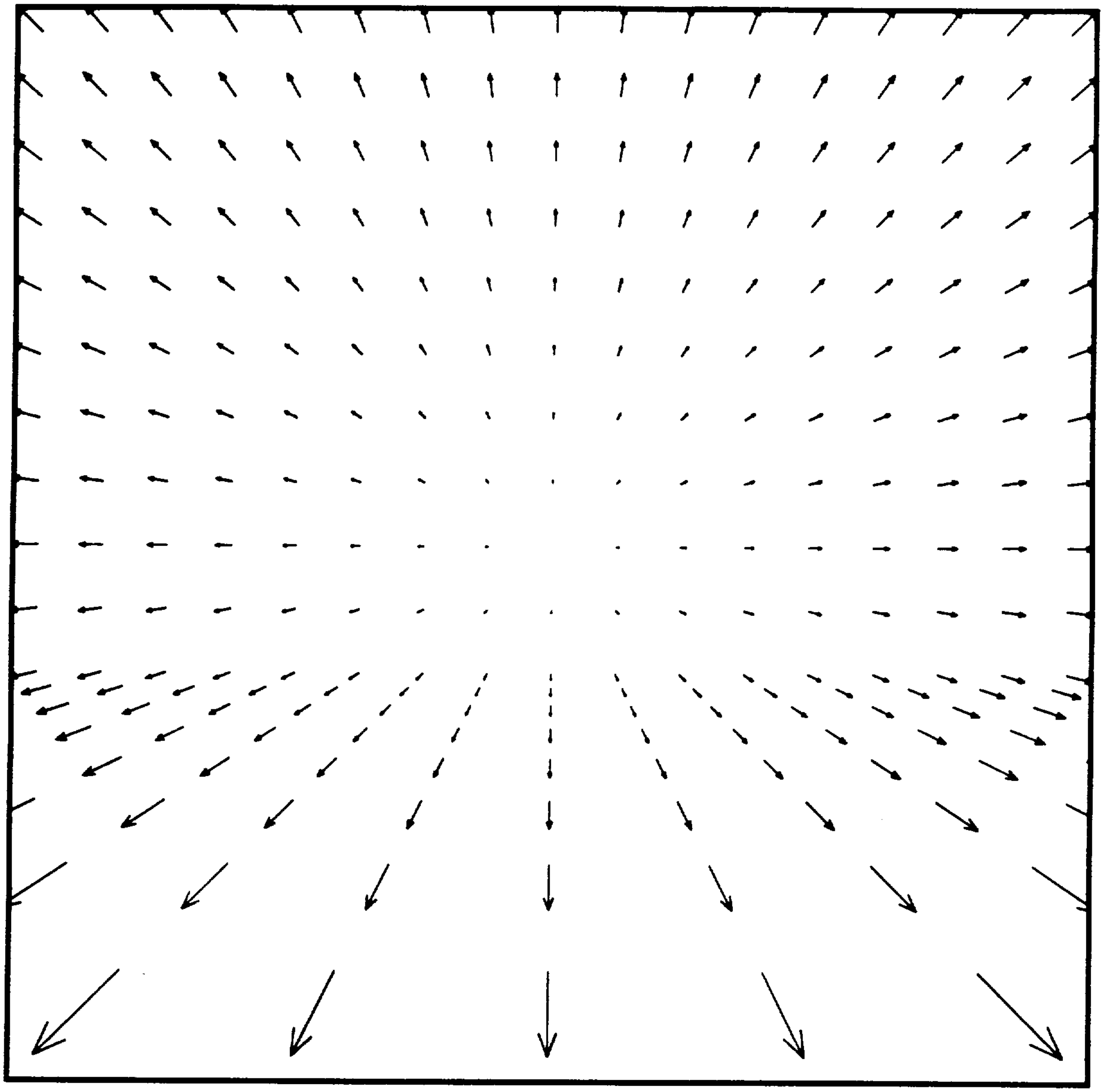

Figure 5. Flow field for motion toward a vertical surface.