A Visual Cueing Model for Terrain-Following Applications

Greg L. Zacharias* and Alper K. Caglayan*
Charles River Analytics, Inc., Cambridge, Massachusetts
and
John B. Sinacori†
John B. Sinacori Associates, Hollister, California

A model to account for the pilot's processing of visual flowfield cues during low-level flight over uncultured terrain is described. The model is predicated on the notion that the pilot makes noisy, sampled measurements on the spatially distributed visual flowfield surrounding him, and, on the basis of these measurements, generates estimates of his own linear and angular terrain-relative velocities which optimally satisfy, in a least-squares sense, the visual kinematic flow constraints. A subsidiary but significant output of the model is an "impact time" map, an observer-centered spatially scaled replica of the viewed surface. Simulation results are presented to demonstrate the potential for modeling relevant human visual performance data and evaluating candidate simulator configurations in terms of expected impact on the perceptual performance of the terrain-following pilot. Additional model applications are discussed, including interfacing with other human performance models and modeling other types of visually driven human task performance.

Introduction

THE current trend toward visual "stimulus realism" in flight simulators has been made possible by rapid advances in computer generated image (CGI) technology. Because of its appeal, there is a continuing trend in this direction, but it is not clear that such a strategy will guarantee effective training at reasonable cost. Currently, for example, there are efforts directed toward development of 200,000 polygon systems,[1] when it is possible that a more "cartoon-like" image will focus a trainee's attention on the critical task cues and result in more rapid training and efficient information transfer. Even if enhanced training is not achieved with such an impoverished (but directed) display, it may be possible to achieve significant system cost savings and, thus, improve overall simulator cost-effectiveness, if not absolute level of training effectiveness. In short, we are skeptical of the ultimate efficacy of full visual stimulus realism for training.

The development of effective "nonrealistic" displays requires, of course, an understanding of not only human visual processing capabilities and limitations, but also an understanding of "downstream," task-related capabilities (such as flight control), as well as learning behavior in training situations. We recognize that research advancements in these areas have not kept pace with the technological advances in CGI systems, and that the resulting "knowledge vacuum"' is perhaps as responsible for the trend toward realism as is the supporting technology itself. This trend, however, does not obviate the need for pushing forward in our effort to understand the pilot's perceptual capabilities.

To this end, we have undertaken an effort directed at understanding human visual processing in one specific flight task: terrain-following flight. We have deliberately limited the domain of applicability to assure a reasonable probability of success in modeling the processing function; we have chosen the specific flight task both on the basis of its current opertional relevance and its potential for tractable functional modeling.

Terrain-Following Flight

Terrain-following flight is performed when entering or leaving air space containing ground-to-air threats. The intent is the timely movement toward an objective while using the terrain for masking. This difficult and dangerous task results in high crew workload and demands a continuing awareness of one's own position and velocity with respect to the ground.

An informal review of "classical" visual cueing mechanisms[2] identifies few candidates which might subserve the required terrain-relative positioning function. The "apparent/familiar size" mechanism can provide cues to absolute distance, but in uncultured (devoid of man-made objects) terrain with tree, shrub, and grass cover, there are few, if any, objects that are familiar enough to support absolute distance estimation. A further examination of real terrain samples for interposition, texture density, and texture gradient cues showed us that these can be strikingly weak and/or sparse. Futhermore, while shading effects and aerial perspective effects are commonly seen, they are not associated with mechanisms that permit the perception of absolute distance; also flight at low level is easily performed without them.

It is well known, however, that pilots are able to fly at low-levels over unfamiliar terrain that they have never seen before. Indeed, they report no unusual difficulty except over "unfavorable" terrain; for example, snowscapes, desert sand dunes, or open, glassy water. This suggests that visual texture is an important factor. Since the static attributes of texture density and texture gradient did not appear to subserve the required distance cueing function adequately, we turned our attention to the dynamic attributes of visual texture, or the "flowfield."

Flowfield Cueing

Flowfield cues arise when an observer moves through a fixed visual world, causing lines of sight to viewed objects to change dynamically in response to the changing geometry between observer and object. These lines of sight and their

Presented as Paper 83-1081 at the AIAA Flight Simulation Technologies Conference, Niagara Falls, N.Y., June 13-15, 1983; submitted Aug. 11, 1983; revision submitted June 1, 1984.

*Principal Scientist. Member AIAA.
†President. Member AIAA.

corresponding rates of change define the observer's visual flowfield: a dynamic "map" which reflects the observer's translational and rotational motion with respect to the static visual world. Our concerns here are not so much with the generation of this map, but rather with the use of the information implicit in it relevant to the pilot's requirement for ongoing terrain-relative positioning information.

Gibson's[2] early pictorial representation of the flowfield initiated a series of studies directed toward a better understanding of the field characteristics. These studies (summarized and critiqued in detail by Warren[3]) provide significant insight into the basic cueing problem, but are limited by a lack of mathematical formalism and/or generality. Thus, some deal only with field speed and fail to recognize field velocity as a vector quantity (see, for example, Refs. 4 and 5); some present mathematically unverifiable results (see, for example, Ref. 6); some are restricted to the viewing of a single flat plane in the visual field (see, for example, Refs. 7 and 8); and most are restricted to considering only nonrotating rectilinear observer motion (see, for example, Refs. 4, 5, 7 and 9).

More recent work has focused on determining the "informational content" of the flowfield and how to extract it. Serveral studies have demonstrated, in principle at least, that flow cues alone can be used to determine not only observer (translational and rotational) velocity with respect to a fixed visual world, but also the geometric shape of that world. However, these studies also suffer from various shortcomings. Specifically, some deal only with nonrotating rectilinear motion[10]; some tie a proposed algorithm to a specific projection geometry,[11,12] when, in fact, the flowfield is independent of projection planes; some demand advanced computational capabilities for accomplishing what appears to be a relatively straightforward estimation function (for example, using first and second spatial gradients of the flowfield to estimate surface shape[10,11,13] when an interpolated "impact time" surface would seem to do as well); some focus concern on the "observability" or "solvability" of impoverished visual scenes having few visual objects,[12,14] when, in fact, most noncontrived scenes are likely to be characterized by an object count which is many orders of magnitude larger than the minimum required; and, finally, most fail to deal with the fundamental issue at hand: making accurate and reliable motion/surface estimates with what we know will be "noisy" flowfield cues.

Thus, we have directed our attention to the development of a model of flowfield cue processing, centered on a computational approach designed to overcome most of the above shortcomings. Our goal is for it to be quantitative, applicable to general viewing geometries, and observer motions, independent of projection geometry, and rational in its treatment of "redundant" and "noisy" flowfield cues. Our motivation stems from the terrain-following flight problem, but we are optimistic that such an approach has applicability to the larger family of viewer locomotion problems. Thus, we summarize herein the preliminary results of this effort.

Definition of Flowfield

We begin our definition of the flowfield by assuming that the external visual world can be modeled as an array of fixed, rigid, and opaque surfaces, which may or may not be connected together in some fashion to form visual "objects." Thus, the visual world may be a single surface, such as the rolling ground plane viewed by a pilot flying at low levels over the desert, or it may be a complex of connected flat surfaces, such as the array of building faces one might encounter in the center of a city.

For an observer moving with respect to this visual world, the problem geometry is as illustrated in Fig. 1. The observer's position $\underline{r}$ is referenced to a surface-fixed coordinate system (e.g., a conventional north-east-down local navigation frame), as are his linear and angular velocities, $\underline{v}$ and $\underline{\omega}$. He is shown "viewing" a point on the surface P_i defined by an observer-relative position vector $\underline{\rho}_i$. Denoting the magnitude of a vector $\underline{r}$ by r, the associated unit-length line of sight (LOS) vector $\underline{u}_i$ is then given by

$$\underline{u}_i = \underline{\rho}_i / \rho_i \tag{1}$$

In the observer's frame of reference, this LOS vector will appear to change with time, at a rate given by

$$\dot{\underline{u}} = \underline{\omega}_i \times \underline{u}_i \tag{2}$$

where the apparent rotation rate from the point of view of an observer both translating and rotating with respect to the fixed visual world is given by[15]

$$\underline{\omega}_i = \underline{u}_i \times (\underline{u}_i \times \underline{\omega}) - (1/\rho_i)(\underline{u}_i \times \underline{v}) \tag{3}$$

where $\times$ denotes the vector product. In Eq. (3), the first term is due to rotational motion, whereas the second term is due to rectilinear motion.

We can consider the visual "flow" associated with a single point P_i and LOS vector $\underline{u}_i$ to be simply the associated LOS rate vector $\dot{\underline{u}}_i$. The vector couple $(\underline{u}_i, \dot{\underline{u}}_i)$ thus defines single point location and flow. By direct extension, the "flowfield" associated with a set of N viewed points can be defined by the set of N corresponding LOS and LOS rate couples, or $\{(\underline{u}_i, \dot{\underline{u}}_i);\ i=1,\ldots,N\}$. As N increases, this spatially sampled field gradually approximates our intuitive notion of a (spatially continuous) visual flowfield.

Figure 2 gives a perspective rendition of the visual flow associated with observer motion over flat (textured) terrain, horizon, and (untextured) sky.

For simple translation (with no rotation), Fig. 2a shows the zero flow rate "expansion point," whose associated LOS vector is colinear with the observer's linear velocity vector $\underline{v}$. Assuming that the observer motion is toward the surface being viewed, then the "flow" of every other point P_i would be radially outward from this expansion point, with a flow rate given by the magnitude of the corresponding LOS rate vector, or, from Eq. (3),

$$\omega_i = \left| -\frac{1}{\rho_i}(\underline{u}_i \times \underline{v}) \right| = \frac{v}{\rho_i}\sin\beta_i \tag{4a}$$

where β is the eccentricity of P_i relative to the expansion point (i.e., the included angle between $\underline{u}_i$ and $\underline{v}$). Thus, an observer undergoing pure translation sees a radial flowfield, where flow rate increases with expansion point eccentricity.

For simple rotation (with no translation), Fig. 2b illustrates the flow for an observer undergoing a left roll ($\underline{\omega}$ out of page). The zero flow rate "rotation point" corresponds to the LOS vector colinear with the observer's angular velocity vector $\underline{\omega}$. The "flow" of every other point P_i is normal to the corresponding radial which emanates from this rotation

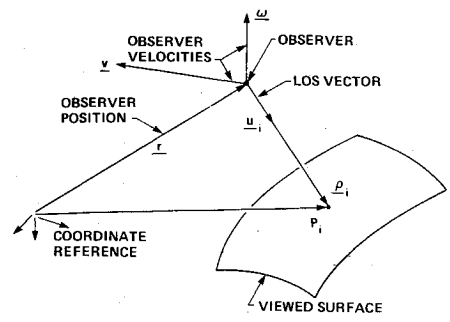

Fig. 1 Viewing geometry.

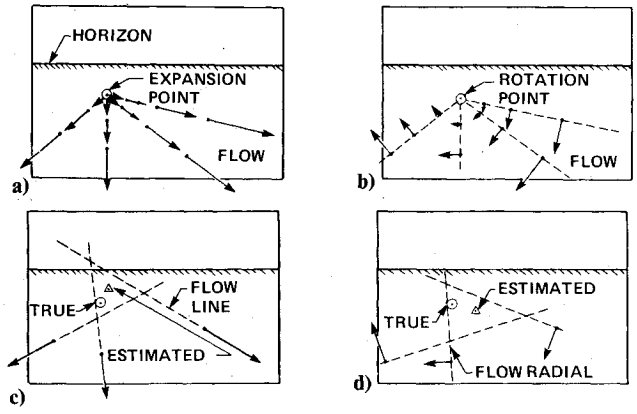

Fig. 2 a) Translational flow, b) rotational flow, c) aimpoint estimation, and d) spin axis estimation.

point, with a flow rate again given by the magnitude of the corresponding LOS rate vector, or, from Eq. (3)

$$\omega_i = |\underline{u}_i \times (\underline{u}_i \times \underline{\omega})| = \omega \sin\beta_i \tag{4b}$$

where, in this case, β_i is the eccentricity of P_i relative to the rotation point (i.e., the included angle between $\underline{u}_i$ and $\underline{\omega}$). Thus, an observer undergoing pure rotation sees a circumferential flowfield, where the flow rate increases with rotation point eccentricity.

Even with these simple constrained motions, the flowfield can show considerable richness and complexity. For example, a simple dive toward flat terrain can give rise to a complex distribution of flow rates within the field of view, and this distribution, in turn, will be a sensitive function of the dive angle itself.[15] Likewise, straight-and-level flight over hilly terrain can result in very complicated flow patterns; a graphic example is provided in Ref. 16. Finally, if we consider combined translational and rotational motions associated with normal aircraft maneuvers, we can generate corresponding flowfields which undergo dramatic and complex "deformations" over the course of the maneuver, even for simple maneuvers conducted over flat terrain.[15]

The potential complexity associated with these and other factors suggests that any self-motion estimator based on flowfield *pattern* recognition is likely to fail, except for fairly trivial and artificially constrained cases. Although qualitative discussions of flowfields and their interpretation in terms of the generating self-motion are useful in providing insight into the cueing situation, it does not appear that such an approach can provide a quantitative basis for self-motion estimation, given the wide scope of potential field patterns one might encounter during terrain-following flight. Thus, we are motivated to adopt a strategy which does not focus on flowfield *patterns* per se, but rather on local flowfield *measurements*; measurements which directly reflect the observer's motions through the cueing geometry involved.

Proposed Visual Cue Processing Strategy

We can specify a measurement strategy by assuming that the observer is limited to making only angular measurements on the visual world, and, hence, on the flowfield. In particular, we assume that he can measure, for each visible point P_i, only the LOS vector $\underline{u}_i$ and its angular rate of change $\underline{\omega}_i$. We specifically assume he is incapable of measuring the relative point range ρ_i.

For an N-point measurement set, the observer's self-motion estimation problem then becomes one of "solving" the set of N LOS rate equations given by Eq. (3) for the (unknown) linear and angular observer velocities, $\underline{v}$ and $\underline{\omega}$, and the (N unknown) relative point ranges $\{\rho_i\}$, given the (N known) visual LOS measurement couples $\{\underline{u}_i, \underline{\omega}_i\}$.

Direct inspection of Eq. (3) shows that it is not possible to "solve" for the unknowns $\underline{v}$ and ρ_i, since they enter only as the ratio $(\underline{v}/\rho_i)$, and thus can be known only to within a common scale factor. This motivates the introduction of two new unknowns, a heading vector $\underline{u}_v$ and an "impact time" τ_i defined for a nonstationary observer by

$$\underline{u}_v = \underline{v}/v \qquad \tau_i = \rho_i / v \tag{5}$$

The unit-length heading vector thus defines only the direction of the observer's motion, but not his speed. For a stationary observer v is zero, and these variables are undefined. This degenerate case easily can be treated separately. The impact time is so named because this would be the elapsed time before observer impact with the surface at point P_i *if* the observer were to head directly at P_i at his current speed v.

These variables allow us to express the $(\underline{v}/\rho_i)$ ratio of Eq (3) as:

$$(\underline{v}/\rho_i) = (v/\rho_i)\,\underline{u}_v = (1/\tau_i)\,\underline{u}_v \tag{6}$$

Clearly, this ratio will be determined by the $(\underline{u}_v, \tau_i)$ values and will be indifferent to the particular $(\underline{v}, \rho_i)$ values, as long as they vary according to

$$\underline{v} = c\underline{u}_v \tag{7a}$$

$$\rho_i = c\tau_i \tag{7b}$$

where c is an arbitrary (positive) common scale factor. These joint constraints are of course central to the design of any terrain-board visual simulation: the speed scale factor used in driving the camera gantry, c_{gantry}, must closely match the terrain map scale factor, c_{map}, if the flowfield cueing seen by an in-simulator pilot is to closely match the one seen by an in-flight pilot.

The observer-related unknowns of absolute velocity and point range $(\underline{v}, \rho_i)$ are thus fundamentally "unobservable" from the flowfield measurements alone; consequently, we are motivated to focus our attention on the flow-related unknowns of heading and impact time $(\underline{u}_v, \tau_i)$.

With the definitions for $(\underline{u}_v, \tau_i)$ given by Eq. (5), the measurement equation (3) becomes

$$\tau_i (A_i \underline{\omega} + \underline{\omega}_i) = \underline{u}_v \times \underline{u}_i \tag{8}$$

where we have introduced the (3×3)rotation matrix

$$A_i = I - \underline{u}_i \underline{u}_i^T \tag{9}$$

where I denotes that (3×3) identity matrix, and the superscript T denotes a transpose. The observer's motion estimation problem thus becomes one of solving the set of N LOS rate equations given by Eq. (8) for the unknown heading and angular velocity, $\underline{u}_v$ and $\underline{\omega}$, and the (N unknown) impact times $\{\tau_i\}$, given the (N known) visual LOS measurement couples $\{\underline{u}_i, \underline{\omega}_i\}$.

The structure of this three-dimenional vector measurement equation becomes clearer by decomposing it into its three scalar components. If we choose an orthogonal coordinate system defined by the point-specific orthonormal triad $(\underline{u}_i^L, \underline{u}_i^M, \underline{u}_i^N)$, defined by

$$\underline{u}_i^L = \underline{u}_i; \quad \underline{u}_i^M \perp \underline{u}_i^L; \quad \underline{u}_i^N = \underline{u}_i^L \times \underline{u}_i^M \tag{10}$$

where $\underline{u}_i^L$ is the LOS vector to P_i and $\underline{u}_i^M$ and $\underline{u}_i^N$ is any arbitrary pair of unit vectors in the plane perpendicular to $\underline{u}_i$, then Eq. (8) can be decomposed into the following three L-,

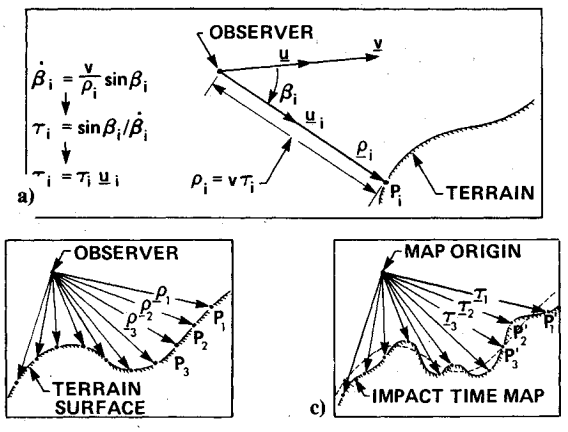

Fig. 3 a) Flight over rolling terrain, b) terrain surface, and c) impact time map.

M-, and N-axis scalar equations[15]:

$$L: \quad 0=0 \tag{11a}$$

$$M: \quad \tau_i = (\underline{u}_i^N \cdot \underline{u}) / \underline{\ }_i^M \cdot (\underline{\omega}+\underline{\omega}_i) \tag{11b}$$

$$N: \quad \underline{u}_v \cdot (A_i\underline{\omega}+\underline{\omega}_i) = 0 \tag{11c}$$

where $(\cdot)$ denotes the scalar product. Thus, the vector measurement equation is equivalent to 1) a trivial identity, 2) an impact time equation which defines τ_i in terms of the other variables, and 3) a "residual" measurement equation constraining the observer velocity states $\underline{u}_v$ *and* $\underline{\omega}$. This separation of the point-dependent unknown τ_i, from the point-independent unknowns $\underline{u}_v$ and $\underline{\omega}$, immediately suggests a stepwise solution. First, by some means, "solve" the N measurement residual equations of Eq. (11c), for $\underline{u}_v$ and $\underline{\omega}$; then, with the measurement $\underline{u}_i$, and the triad vectors defined by Eq. (10), directly compute a τ_i for each of the N viewed points using Eq. (11b).

Once impact time is found in this manner, we can compute a corresponding impact time vector from

$$\underline{\tau}_i = \tau_i \underline{u}_i \tag{12}$$

From Eqs (1) and (5), this vector is merely the speed-scaled observer-relative range vector to the viewed surface $(\underline{\rho}_i/v)$. Thus, the set of impact time vectors $\{\underline{\tau}_i\}$ provides us with an observer-centered spatially sampled and scaled replica of the viewed surface. Simply stated, viewed surface shape is potentially available from the flowfield cues, i.e., the LOS vectors and their associated angular rates.

Least-Squares Estimator

Since computation of the impact time map is straightforward once we know $\underline{u}_v$ and $\underline{\omega}$, the basic estimation problem reduces to finding a solution $(\underline{u}_v,\underline{\omega})$ to the N residual equations of Eq. (11c), an overdetermined equation set in any normal viewing situation (where N is normally very large). Since there will be errors in the measurements $(\underline{u}_i,\underline{\omega}_i)$, this equation set will not be satisfied exactly, and, in fact, will take on the form:

$$\underline{u}_v \cdot (A_i\underline{\omega}+\underline{\omega}_i) = \epsilon_i \tag{13}$$

where ϵ_i is the measurement residual associated with the ith measurement couple $(\underline{u}_i,\underline{\omega}_i)$ and the estimate $(\underline{u}_v,\underline{\omega})$. The set of N residuals thus provides a natural measure of how well the single estimate fits the N "noisy" measurements. With this view of the problem, we are thus motivated to take a least-squares approach, and generate an estimate which minimizes the sum of the squared residuals, by choosing $(\underline{u}_v,\underline{\omega})$ to minimize the cost J where

$$J = J(\underline{u}_v,\underline{\omega}) = \frac{1}{2}\sum_{i=1}^{N} \epsilon_i^2 \tag{14}$$

A graphical interpretation of this formulation of the problem is given in Fig. 2, which we used earlier to illustrate two simple flow patterns arising from uncoupled observer motion over flat terrain. The radial flow pattern of Fig. 2a suggests that the expansion point (or heading vector $\underline{u}_v$) may be estimated from the intersection point of a number of backward-traced flow vectors. Figure 2c shows how this might work for a three-point noisy measurement case. Errors in specifying point location (LOS $\underline{u}_i$) and flow rate (LOS rate $\underline{\omega}_i$) preclude a single point intersection of the three extended flow vectors, or "flow lines"; instead, we obtain a triangular region of uncertainty. As shown, any estimate, in general, will be some residual distance away from each of the three flow lines. A reasonable estimation strategy would be to choose the estimate which minimizes some joint metric of the residuals: the sum of the squared distance residuals naturally comes to mind. Since it is a direct matter to show that the distance residuals (for small errors) are proportional to the measurement residuals of Eq. (13), minimizing the cost J of Eq. (14) becomes roughly equivalent to choosing a heading estimate which is "optimally near" all of the observed flow lines, in a least-squares sense.

The circumferential flow pattern of Fig. 2b suggests a similar strategy for estimating the rotational point, here based on the intersection of the flow radials normal to the flow vectors themselves. Again, it is straightforward to show that this strategy (illustrated in Fig. 2d) yields distance residuals which are proportional to the measurement residuals of Eq. (13), so that minimizing J in this situation is roughly equivalent to finding a rotational estimate which is "optimally near" all of the flow radials. Note, however, that this graphic strategy provides no direct means of inferring total angular velocity $\underline{\omega}$, but only spin direction $\underline{\omega}/\omega$. Additional features of the field must be considered to estimate spin magnitude.

This discussion has attempted to provide a graphical interpretation of the formal minimization problem we introduced with Eq. (14). It should be recognized that the illustrations presented here have been limited to simple uncoupled observer motion (pure translation or pure rotation), have considered only a three-point measurement solution, rather than the general N-point solution, and have lacked somewhat in rigor, because of the use of perspective geometry projections in place of a more appropriate projection-free spherical geometry representation. Our formal minimization problem, of course, suffers from none of these limitations.

We have thus specified a least-squares estimator, which generates an estimate $(\underline{u}_v,\underline{\omega})$ to minimize the sum-squared measurement residual cost function of Eq. (14). The particular method by which this minimizing estimate is generated formally defines the estimator, but we will not go into the details here since they are covered in Ref. 17. We will note, however, that the estimator "solves" the nonlinear constrained least-squares problem by direct minimization of the residual cost. It does this recursively, in effect, by taking alternating minimizing steps in the individual $\underline{u}_v$ and $\underline{\omega}$ solution subspaces to arrive at a total solution $(\underline{u}_v,\omega)$ which globally minimizes the residual cost. Because the algorithm depends on simple matrix functions (e.g., inverse, eigenvalue) which operate on matrices no larger than 3×3, convergence of the algorithm is rapid and reliable.

It is worth noting again that once we have obtained an estimate $(\underline{u}_v,\underline{\omega})$, we are then in a position to directly calculate, from Eqs. (11b) and (12), the set of impact time vectors $\{\underline{\tau}_i\}$. As noted, these define an observer-centered spatially sampled and scaled replica of the viewed surface. Because of noise in the measurements $(\underline{u}_i,\underline{\omega}_i)$ and errors in the estimates $(\underline{u}_v,\underline{\omega})$,

the impact time vector set, and, thus, the impact time surface, will be a noisy approximation to a perfect surface replica. The accuracy of the approximation, of course, will depend on the level of the driving measurement noise and the viewing geometry.

Figure 3 illustrates the construction of this map for the special case in which an observer is flying straight-and-level over rolling terrain. Figure 3a sketches the viewing geometry in the plane containing $\underline{v}$ and the position vector $\underline{\rho}_i$ to the ith viewing point P_i. The heading eccentricity β_i is determined from the LOS $\underline{u}_i$ and the estimated heading $\underline{u}_v$; eccentricity rate $\dot{\beta}_i$ is the flow rate determined by Eq. (3), which, from Eq. (4a), simplifies to the first expression shown in Fig. 3a. From this we can solve for the impact time τ_i by simplifying Eq. (11b) which, with $\underline{u}_i$, allows us to compute the impact time vector $\underline{\tau}_i$ as shown.

Figure 3b sketches the situation for a number of coplanar viewed teriann points, each associated with an observer-referenced position vector $\underline{\rho}_i$. For each we can compute, via the process just outlined, an impact time vector $\underline{\tau}_i$. The resulting "map" of this set $\{\underline{\tau}_i\}$ is shown in Fig. 3c. Note that each terrian point P_i becomes a projected point P_i' in impact time space, although perfect registry of the two point sets is precluded by errors in the estimated vector set $\{\underline{\tau}_i\}$ due to flowfield measurement errors. Also note that we have indicated, in Fig. 3c via a solid line, an interpolated or continuous impact time surface; such interpolation naturally requires some sort of surface model, a modeling issue which goes beyond the scope of the estimation problem under discussion.

Although the flowfield pattern is used above to demonstrate the computation of the impact time map for the case of pure translational motion, these flow patterns are not used by our developed visual model which treats the general case of combined translational and rotational motion. For the case of general motion, the simple graphical interpretation of an "expansion point" and "center of rotation" would no longer be possible. In this case, the flow seen by the observer would consist of a set of curved flow lines ("streamers"), as discussed in Ref. 18.

Simulated Estimator Performance

The performance of the estimator has been investigated through digital simulation. We present herein some preliminary results which demonstrate estimation performance trends over various problem factors of interest.

Noisy observation of the visual flow was simulated by additive measurement noise corrupting the LOS measurement vector couple $(\underline{u}_i, \underline{\omega}_i)$. Vector noise magnitudes were normally distributed with zero mean; directions were uniformly distributed with the LOS noise constrained to ensure a unit length "noisy" measurement vector. Standard deviation (SD) of the LOS noise magnitude σ_{u0} was set at 1 arc-min based on human static threshold considerations.[19] The SD of the LOS rate noise magnitude σ_{ω_i} was defined as a composite of a fixed value and a value proportional to actual LOS rate ω_i, or

$$\sigma_{\omega i} = [\sigma_{\omega 0}^2 + (f_{\omega i})^2]^{1/2} \tag{15}$$

where $\sigma_{\omega 0}$ was set at 1 arc-min/s, and the rate noise/signal (N/S) ratio f was set at 10% or -20 dB, based on human dynamic threshold considerations.[19]

Results presented herein were obtained by computing the ensemble statistics of the estimation errors across 100 Monte Carlo simulation runs.

Sensitivity to Display Parameters

Simulations were conducted for a pilot flying straight-and-level at 400 knots, 100 ft above flat featureless textured terrain. Out-the-window field of view (FOV) was nominally limited to ±24 deg laterally, 24 deg down, and 12 deg up. A finite distance horizon was modeled to keep the textural point count finite; this distance was about 3500 ft, based on an assumption of a maximum 5 s preview or "look-ahead" distance. A number of surface decoration and visual sampling strategies were studied; the results shown here assume the observer visually samples the terrain surface in such a manner as to ensure that the resulting LOS vectors are uniformly spherically distributed over the solid angle subtended by the viewable terrain.

Figure 4 shows the standard deviation of the error in estimated heading to "aimpoint" $\underline{u}_v$. Error trends with N, the number of points viewed, are sketched in for lateral (left/right) error, vertical (up/down) error, and total error. As might have been expected from the least-squares structure of the estimator, significant error reductions are to be had as viewed point count N increases over the range from 10 to 1000. Point counts less than 10 occasionally led to algorithm convergence problems when certain "solvability conditions" were violated[15]; counts greater than 1000 were not run to conserve computational resources. Note, however, that the rate of these reductions diminish rapidly with increasing N. The implication for CGI design would seem to be that, in simulating this task, diminishing returns are to be had in terms of improving the simulator pilot's perception of heading for counts much greater than about 1000. Naturally, the use of such a low edge count to represent a high-density visual scene will lead to a considerable lack of "simulator realism"; however, the model results would indicate that this shortcoming in appearance would have little impact on the perceptual performance of the simulator pilot (for this task, terrain type, etc.).

Figure 5 shows the standard deviation of the error in estimating the total angular rate and the error breakdown by body axis using the conventional axis assignment, with roll along $\underline{v}$, yaw pointing down to the surface, and pitch normal

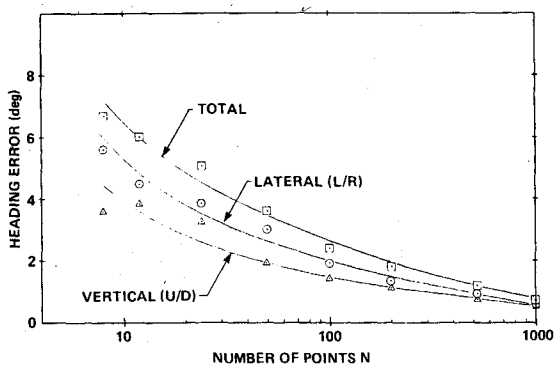

Fig. 4 Heading error vs array size.

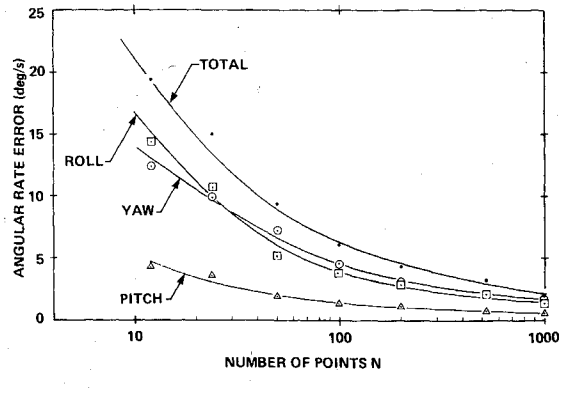

Fig. 5 Angular rate error vs array size.

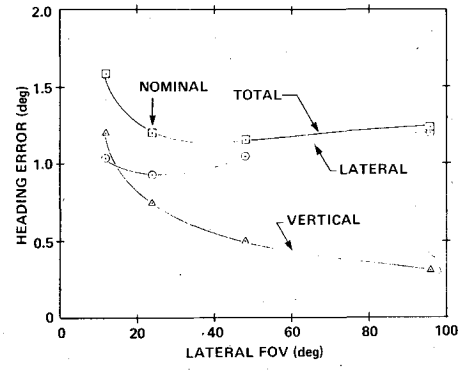

Fig. 6 Heading error vs lateral FOV.

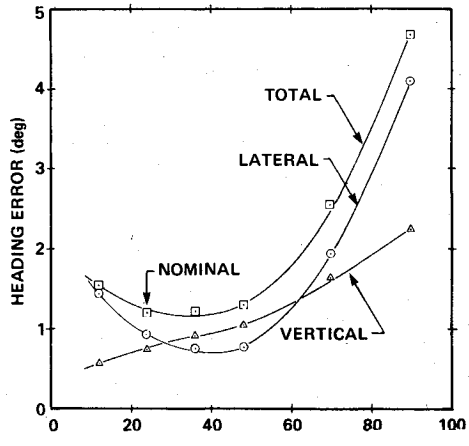

Fig. 7 Heading error vs vertical FOV.

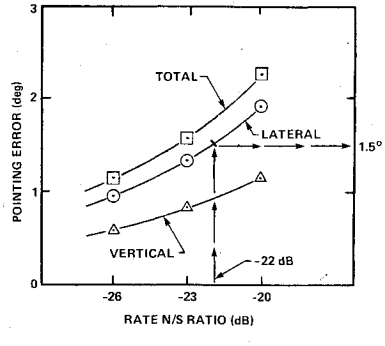

Fig. 8 Pointing error vs noise level.

to the other two. The trends with point count are similar to those just seen for heading error, and would imply a similar point of diminishing returns in the scene density, in terms of potential improvements in perceiving angular rates.

Figures 6 and 7 show heading estimation performance as a function of lateral and vertical FOV limits. Viewed point count N was maintained at approximately 500 points across all FOV conditions indicated.

Figure 6 shows the effect of lateral (half-width) FOV variations. As the curves indicate, once the FOV exceeds roughly the nominal 24-deg half-width, the opposite trends of lateral and vertical heading estimation errors yield a total heading error function which is fairly insensitive to lateral FOV. For the geometry we are considering, this suggests that aimpoint can be fairly well localized in the absence of the peripheral visual flow provided by wide FOV displays, assuming, as we have in these simulations, that the FOV is centered on the aimpoint.

Figure 7 shows the effect of variations in the vertical (downward) FOV. As seen, vertical errors grow directly with increasing FOV throughout the FOV range; lateral errors also increase with FOV, but only for the larger FOV limits. Because of the relative minimum in the lateral error curve, both errors combine to yield a total heading error function which is minimized for an FOV of about 36 deg. We suspect that this behavior reflects the fact that an FOV that is too small overly constrains the viewing geometry and provides an inadequate measurement sample for accurate estimation, whereas an FOV that is too large encompasses very high-rate, and, thus, high-noise measurements which degrade the overall heading estimate. Hence, an intermediate vertical FOV limit should thus ensure the best estimation performance.

Human Performance Simulation

Currently, we are conducting model simulations of human estimation performance. We report herein the results of one such simulation of a psychophysical experiment conducted by Warren.[20]

Subjects were "flown" over flat terrain decorated by a uniformly random array of luminous dots. The flight path was straight and level with an altitude of d units and speed $1.25d$ units/s, where d was the average interdot spacing of the terrain texture. Lateral FOV was ± 26 deg; vertical FOV was 26 deg below the displayed horizon, which was depressed about 1.2 deg below the true horizon to maintain a finite texture point count (of about 1100 viewable points). After viewing the display, subjects were required to indicate their lateral heading aimpoint, and were able to do so with an accuracy of 1.5-deg standard deviation (across subjects).

Figure 8 shows the results of a model simulation of this experiment. Using the above problem parameters, and visual thresholds of 1 arc-min and 1 arc-min/s, heading estimation performance was "swept out" as a function of N/S ratio f, as shown. In modeling terms, the observed lateral pointing accuracy of 1.5 deg is simply explained by a -22 dB rate noise ratio, or, equivalently, a dynamic discrimination capability of about 8%. This is in remarkable agreement with a demonstrated 10% figure obtained by independent threshold measurements.[19]

Conclusions

We have developed and exercised a model of visual flowfield cue processing for direct application to low-level flight. Our consideration of the terrain-following cueing environment allowed us to focus specifically on flowfield cueing and develop a mathematical description of the kinematic flow constraints. A least-squares solution satisfying these constraints was proposed as the basis for the estimator model, which we showed was capable of estimating not only observer motion states, but viewed surface shape as well.

The least-squares estimator formulation of the cueing model is applicable to general observer motions and viewing geometries, is projection-plane independent, and is rational in its treatment of redundant and noisy flow cues. Simulations of its performance demonstrate its utility in investigating the sensitivity of task-relevant perceptual performance to simulator-related visual display parameters, such as edge count and field of view. Preliminary simulation results of human visual performance are encouraging, and have motivated an ongoing experimental validation effort. Additional applications of visually driven task performance currently are being pursued in an effort to widen the scope of applicability of the flowfield cueing model.

Acknowledgments

This work was performed under USAF Contract F33615-81-C-0515 while Drs. G.L. Zacharias and A.K. Caglayan

were with Bolt Beranek and Newman Inc.; technical monitor was Dr. Grant R. McMillan of USAF AMRL. A series of technical discussions with Drs. G. R. McMillan and R. Warren of AMRL throughout the course of this study are gratefully acknowledged.

References

[1] "Boeing Imagery Apes Reality," *Photonic Spectra,* Jan. 1983, pp. 73-74.

[2] Gibson, J.J., *The Perception of the Visual World,* Houghton Mifflin, Boston, Mass., 1950.

[3] Warren, R., "Optical Transformation During Movement," AFOSR-TR-82-1028, NTIS AD-A122 275/1, Oct. 1982.

[4] Havron, M.D., "Information Available from Natural Cues During Final Approach and Landing," Human Sciences Research Inc., Arlington, Va., HSR-RR-62/3-MK-X, March 1962.

[5] Whiteside, T.C.D. and Samuel, G.D., "Blur Zone," *Nature,* Vol. 225, 1970, pp. 94-95.

[6] Gordon, D.A., "Static and Dynamic Visual Fields in Human Space Perception," *Journal of the Optical Society of America,* Vol. 55, 1965, pp. 1296-1303.

[7] Gibson, J.J., Olum, P., and Rosenblatt, F., "Parallax and Perspective During Aircraft Landings," *American Journal of Psychology,* Vol. 68, 1955, pp. 372-385.

[8] Harker, G.S. and Jones, P.D., "Depth Perception in Visual Simulation," Air Force Human Resources Lab., Brooks Air Force Base, Tex., AFHRL-TR-80-19, Aug. 1980.

[9] Lee, D.N., "A Theory of Visual Control of Braking Based on Information About Time to Collision," *Perception,* Vol. 5, 1976, pp. 437-459.

[10] Clocksin, W.F., "Perception of Surface Slant and Edge Labels from Optical Flow: A Computational Approach," *Perception,* Vol. 9, 1980, pp. 253-269.

[11] Hoffman, D.D., "Inferring Shape from Motion Fields," Artificial Intelligence Laboratory, MIT, Cambridge, Mass., AI Memo 592, Dec. 1980.

[12] Ullman, S., *The Interpretation of Visual Motion,* MIT Press, Cambridge, Mass., 1979.

[13] Longuet-Higgins, H.C. and Prazdny, K., "The Interpretation of a Moving Retinal Image," *Proceedings of the Royal Society of London,* Vol. B-208, 1980, pp. 385-397.

[14] Prazdny, K., "Egomotion and Relative Depth Map from Optical Flow," *Biological Cybernetics,* Vol. 36, 1980, pp. 87-102.

[15] Zacharias, G.L., "Flow-Field Cueing Conditions for Inferring Observer Self-Motion," Bolt Beranek and Newman, Inc., Cambridge, Mass., Rept. 5118, Sept. 1982.

[16] Ballard, D.H. and Brown, C.M., *Computer Vision,* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1982.

[17] Zacharias, G.L., Caglayan, A.K., and Sinacori, J.B., "A Model for Visual Flow-Field Cueing and Self-Motion Estimation," *Proceedings of 1983 American Control Conference,* June 1983; also, to appear in *IEEE Transactions on Systems, Man, and Cybernetics.*

[18] Grunwald, A.J. and Merhav, S.J., "Vehicular Control by Visual Field Cues – Analytical Model and Experimental Validation," *IEEE Transactions on Systems, Man, and Cybernetics,* Vol. 6, Dec. 1976, pp. 835-845.

[19] Statler, I.C., "Characteristics of Flight Simulator Visual Systems," AVRADCOM Research and Technology Lab., NASA Ames Research Center, Moffett Field, Calif., Rept. 81-A-8, April 1981.

[20] Warren, R., "The Perception of Egomotion," *Journal of Experimental Psychology: Human Perception and Performance,* Vol. 2, 1976, pp. 448-456.