

# VISION-AUGMENTED GNC: PASSIVE RANGING FROM IMAGE FLOW

Matthew D. Markel \*  
Sverdrup Technology

Juan Lopez †  
Sverdrup Technology

Glenn Gebert ‡  
Sverdrup Technology

Johnny Evers §  
U.S. Air Force Research Laboratory

## Abstract

Knowledge of range and closing velocity is known to enhance missile guidance, improve launch envelope, and reduce miss; however, scenarios exist where the range-to-target is unknown to the weapon. This paper provides the mathematical framework and algorithm for calculating range-to-target from sequential optical (intensity only) images. Such use of additional information from imaging seekers to assist GNC falls into the general category of Vision-Augmented GNC (VA-GNC), and a brief overview of this emerging field is provided.

It has been shown that through use of the optical flow constraint (OFC), a range expression at each image pixel may be directly calculated from spatial-temporal gradients, known pixel angular location, and ownship motion. However, these optical flow/gradient-based passive ranging algorithms are subject to singularities, noise, and bogus range estimates for certain image content. In this work, a local median filtering approach is applied to the raw range estimates to remove outliers before segmentation. Using this approach provides the additional benefit of both intensity and *range* information for segmentation. Once segmented, the range to the target of interest is extracted and available for guidance or range-rate estimation.

Finally, simulation performance results are presented using synthetic, analytically de-

rived images for the case of a body-fixed camera. These results indicate that, for stationary ground targets, the technique yields range accuracy which is adequate for guidance and object tracking.

## 1. Introduction to VA-GNC

Vision-Augmented guidance, navigation, and control (VA-GNC) is a primary topic of research at the Air Force Research Laboratory (AFRL) at Eglin AFB. Loosely stated, the field of VA-GNC can be defined as the study and development of performance enhancements to weapon GNC functions by incorporating additional information from a weapon's imaging seeker beyond the "typical" tracking (heading) errors. To date, the primary focus of this research has been on improvements to GNC through optical flow, i.e. the apparent motion of brightness contours in an image. In this work, we present an approach to enhance guidance for a weapon system through the passive calculation of range and range rate solely from image data (intensity) and ownship motion.

VA-GNC research has roots in the fundamental idea that animals and humans rely significantly on visual inputs for survival. Motion relative to an object or surface is perhaps the most critical of visual perceptions. Image flow is fundamental to the perception of depth and detection of objects; things that move are likely threats to be avoided, or sources to be pursued. Visual perception of motion and the estimation of range and velocity are important for survival. The rate of apparent motion of a stationary object induced by translation of the eye is inversely proportional to the distance from the eye - close objects produce large image flow. The challenge for

\*Associate Principal Engineer

†Chief Scientist, Senior Member AIAA

‡Tech Fellow for Aerodynamics, Senior Member AIAA

§Senior Member AIAA

This paper is declared a work of the U.S. Government and is not subject to copyright protection in the United States. Drs. Matthew Markel, Juan Lopez, and Glenn Gebert are with Sverdrup Technology, P.O. Box 1935, Eglin AFB, FL 32542. Johnny Evers is with the U.S. Air Force Research Laboratory Munitions Directorate.

# Potential VA-GNC Applications

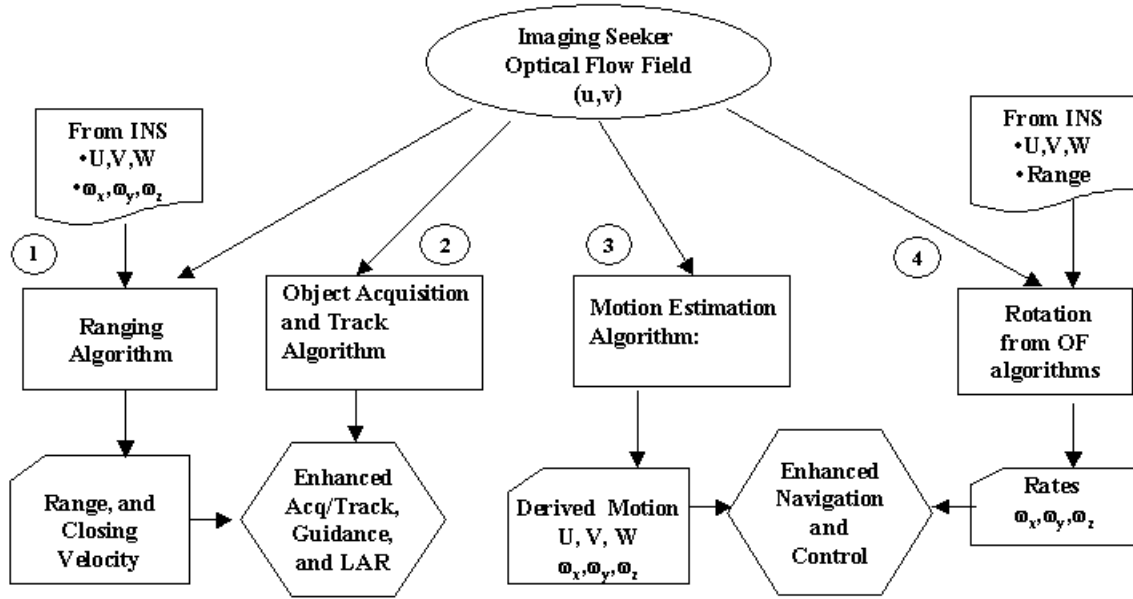


Figure 1: Vision-Augmented GNC (VA-GNC) potential applications.

VA-GNC weapon designers is to effectively employ this additional “visual” information from the seeker to complement the traditional GNC function.

The problem of determining sensor platform motion (velocity and angular rates) from a sequence of images has been the topic of many research papers in the area of machine vision, with much of this work centering on the optical flow. Various research papers have investigated the accuracy of various optical flow techniques. For example, Barron, Fleet, and Beauchemin [1] made an attempt to quantitatively evaluate several existing optical flow estimation methods. Kumar, Tannenbaum, and Balas [2] employed a curve evolution approach to the computation of optical flow. It should be noted that most work on egomotion estimation has been restricted to relatively slow translational and rotational motion; little work has been done exploring the types of motion and speeds typical of weapon applications. The relationship between the optical flow velocity and the dynamic motion of the sensor is based on the original approach of Bruss and Horn [3], and Negahdaripour and Horn [4]. The approach to obtain a closed-form range expression is based on [4] and the work of Raviv and Albus [5]. This approach derives an expression for range using the optical flow constraint of Horn and Schunk [6] to omit the explicit computation of optical flow. This way, range is estimated from image spatial-temporal gradients,

the known seeker velocities and angular rates, and the object location in the image. However, no insight was provided in [5], (which based the derivation on spherical coordinates), as to the accuracy of the gradient-based algorithm to estimate range.

Knowledge of egomotion and object awareness derived from image flow data has the potential to enhance GNC algorithms in modern weapons. The feasibility of using additional vision-based seeker information in a VA-GNC architecture (Figure 1) has several potential areas of application. From a GNC perspective, an easily identifiable application is the tracking of objects in clutter. However, other uses include improving guidance and trajectory shaping, tracking objects and obstacle avoidance, complementing the inertial navigation by compensating for gyro drift. As the environment of modern weapons increases in complexity and requirements, redundant / low-cost information such as image flow data, may be useful to enhance or complement other sensors performance. As an initial step in the implementation of VA-GNC concepts, this paper develops an approach to augment guidance with range-to-target (and range-rate) information derived from a non-ranging seeker via image flow.

It is well known that guidance can be improved over a wide range of operating conditions by using knowledge of range and closing velocity. Potential benefits of an augmented guidance law and trajec-

tory shaping include an extended launch acceptability region (LAR) due to better energy management and an improved terminal miss distance against maneuvering targets. An example may be a guided missile (with an imaging seeker) launched without a range cue. Using the approach presented in this work the weapon could possibly get an autonomous range estimate from the seeker image.

The concepts presented in this work utilize the same general ranging approach presented in [5] to develop a raw range estimate to each pixel in the image, but are novel in several areas. First, the approach provided here offers increased robustness over that of [5] in terms of mitigating the effects of singularities (outliers) in the ranging equation and bogus range estimates. Second, this approach enables an architecture for enhanced segmentation capability over that available via the intensity images alone by including the range image in the segmentation process.

The remainder of this paper is as follows. Section 2 reviews the nomenclature used throughout the work. Section 3 presents the mathematical model and algorithm that develops raw range estimates from image data and ownship motion for each pixel. Section 4, develops a generic image processing architecture using the results and limitations of the algorithm developed in Section 3. Section 5 contains simulation results of the algorithm using synthetic image data. Finally, Section 6 contains the conclusions and recommendations for further research.

## 2. Nomenclature

$\varepsilon_y$	pixel angular location along y dimension
$\varepsilon_z$	pixel angular location along z dimension
$U, V, W$	camera translational velocities
$\omega_x, \omega_y, \omega_z$	camera rotational velocities
$X, Y, Z$	location of target wrt camera
$I(\varepsilon_y, \varepsilon_z, t)$	pixel intensity at time $t$
$u, v$	optical flow velocity
$I_y, I_z$	spatial intensity gradients
$I_t$	temporal intensity gradient
$\Delta T$	time between frames
$\lambda_s$	spatial wavelength
$\Delta P$	pixel angular size

## 3. Range-From-Image Algorithm

The mathematical model for relating motion of the vehicle to sequential frames of optical flow (OF) image data follows the form of [3]. In addition, this derivation assumes that the imaging sensor is attached to the body frame of reference. The more general problem of a gimballed sensor will involve an additional angular transformation between the vehicle-fixed frame and the image plane making use of the seeker gimbal angles, and is not addressed in this paper. The range estimation problem will be formulated in rectangular coordinates. Following [4] and [5], the optical flow constraint is used to by-pass the tedious calculation of the optical flow velocities. In this case, the range is determined directly from the image brightness gradients and other known parameters, the image pixel location, and the ownship velocities and angular rates.

### Mathematical Model and Geometry

Figure 2 describes the geometry, an imaging sensor looking at an object at an unknown downrange, denoted by  $X$ . The object is projected to the image plane and its location given by a pixel location. As sequential images are processed, the OF algorithms provide a time-dependent object velocity in the 2-dimensional image, at each pixel at time  $t$ . Having that information, the desire is to formulate a mathematical model relating the optical flow components and the known object location in the image plane to the translational and rotational motion of the vehicle. The model describes a single object, however the concepts may be applied to several segmented portions of the image. For this derivation the object is assumed stationary. The translational velocity and angular rates of the camera (in the camera frame of reference) are given by the column vectors:

$$\vec{V}_t = [U, V, W]^T \quad (1)$$

$$\vec{\omega} = [\omega_x, \omega_y, \omega_z]^T \quad (2)$$

where  $[\cdot]^T$  represents the vector transpose.

For a stationary target, the total velocity due to camera translation and rotation in the camera frame of reference is given by

$$\vec{V}_p = -\vec{V}_t - \vec{\omega} \times \vec{R}_p \quad (3)$$

where

$$\vec{R}_p = [X, Y, Z]^T \quad (4)$$

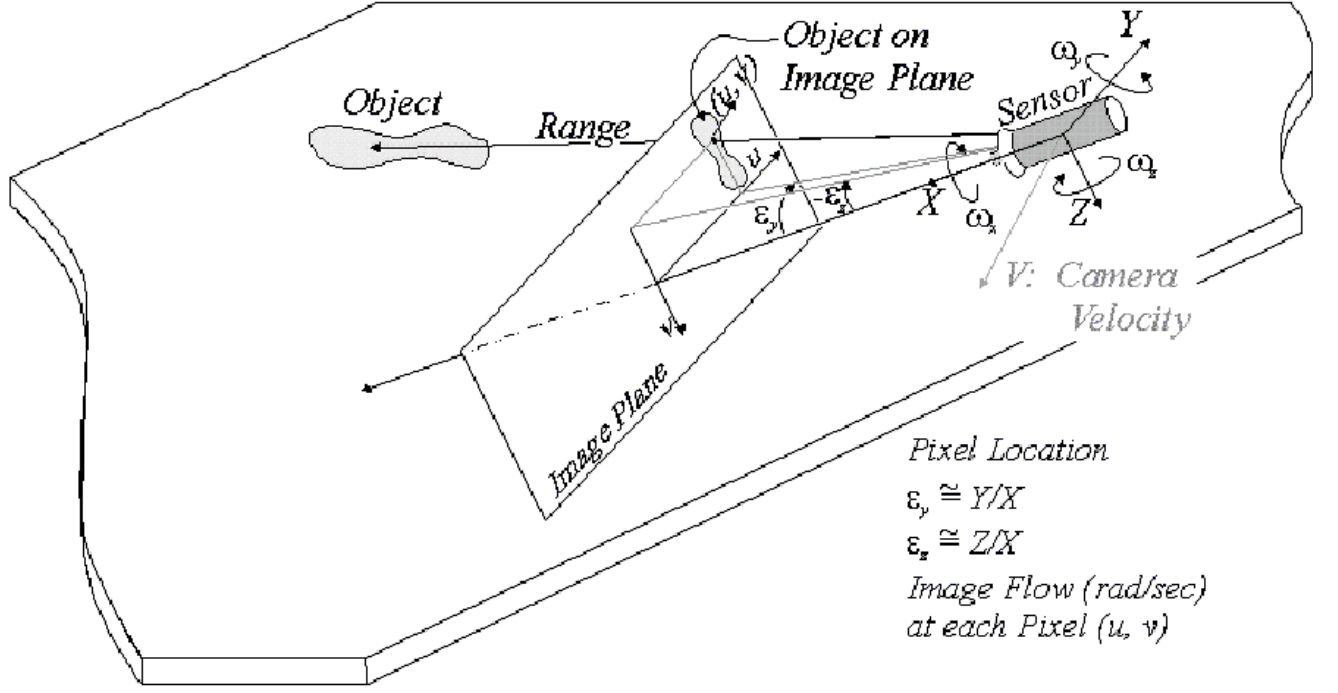


Figure 2: Geometry Description

is the relative position of the object “p” with respect to the camera range-to-target, with the  $X$  axis along the optical axis, and

$$\vec{V}_p = [\dot{X}, \dot{Y}, \dot{Z}]^T \quad (5)$$

is the total velocity in the camera frame of reference. In component form the total velocity is given by

$$\begin{bmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{bmatrix} = - \begin{bmatrix} U \\ V \\ W \end{bmatrix} - \begin{bmatrix} Z\omega_y - Y\omega_z \\ X\omega_z - Z\omega_x \\ Y\omega_x - X\omega_y \end{bmatrix} \quad (6)$$

The location of the object in the image (pixel location) is given by the angles  $\varepsilon_y$  and  $\varepsilon_z$ , as a function of range components

$$\frac{Y}{X} = \tan(\varepsilon_y) \cong \varepsilon_y \quad (7)$$

$$\frac{Z}{X} = \tan(\varepsilon_z) \cong \varepsilon_z \quad (8)$$

where  $\cong$  represents the standard “small angle” approximate equality. Because the seeker tracking loops of guided munitions center objects on the focal plane, the small angle approximation employed here is reasonable. Using the definition of the optical flow velocity components ( $u, v$ ) at the angular

pixel location  $(\varepsilon_y, \varepsilon_z)$ , and taking the derivative of (7) and (8) yields

$$\begin{aligned} u &\triangleq \dot{\varepsilon}_y \\ &= \frac{\dot{Y}}{X} - \frac{Y\dot{X}}{X^2} \end{aligned} \quad (9)$$

$$\begin{aligned} v &\triangleq \dot{\varepsilon}_z \\ &= \frac{\dot{Z}}{X} - \frac{Z\dot{X}}{X^2} \end{aligned} \quad (10)$$

Substituting (6), (7), and (8) in the OF components given by (9) and (10), then separating the translation and rotational components yields the desired relationship between the optical flow velocities and the motion parameters for the case of stationary target.

$$\begin{aligned} u &= \frac{1}{X} (-V + U\varepsilon_y) + \omega_x\varepsilon_z \\ &\quad + \omega_y\varepsilon_y\varepsilon_z - \omega_z(1 + \varepsilon_y^2) \end{aligned} \quad (11)$$

$$\begin{aligned} v &= \frac{1}{X} (-W + U\varepsilon_z) - \omega_x\varepsilon_y \\ &\quad - \omega_z\varepsilon_y\varepsilon_z + \omega_y(1 + \varepsilon_z^2) \end{aligned} \quad (12)$$

Again, in the above two equations (11) and (12), all parameters are assumed known except the down

range component,  $X$ , and the optical flow components  $u$  and  $v$  for each pixel. If the optical flow components  $u$  and  $v$  are available, we have two estimates of range for each pixel in the image.

A rich literature exists on algorithms to process optic flow, i.e. produce  $u$  and  $v$  for each pixel in a focal plane. Additional work on silicon circuitry designed to emulate insect eye processing for optical flow calculation is also ongoing. The implication of these approaches for optical flow calculation for air vehicle motion lies in future work. If the optical flow components are not available, but reliable spatial and temporal image gradients are, then, an expression for range can be found by using the optical flow constraint, as shown below.

### Incorporation of the Optical Flow Constraint

A common equation used in sequence image processing is the Optical Flow Constraint (OFC) or Spatio-Temporal Constraint Equation, formulated in a classical paper by B. Horn and B. Schunck [6]. Loosely stated, the optical flow constraint requires that the intensity of objects in an image is preserved throughout their motion. The OFC is formulated as follows: Given 1) the velocity vector field  $u(\varepsilon_y, \varepsilon_z), v(\varepsilon_y, \varepsilon_z)$  at each pixel location  $(\varepsilon_y, \varepsilon_z)$ , which provides information on how, at that pixel, the brightness is moving across the image, and 2) the image intensity  $I(\varepsilon_y, \varepsilon_z, t)$ , which is a function of position in the image  $(\varepsilon_y, \varepsilon_z)$  and time,  $t$ . Then the OFC states that at a small displacement away, and after a short time, the intensity is maintained. After a first order Taylor series expansion [6] this may be written concisely as:

$$\frac{\partial I}{\partial \varepsilon_y} u + \frac{\partial I}{\partial \varepsilon_z} v + \frac{\partial I}{\partial t} = 0 \quad (13)$$

and holds for all pixels in the focal plane array satisfying the ‘‘small angle’’ assumption.

### Pixel Range Calculation

To obtain the closed-form for range, we substitute the optical flow velocity components,  $u$  and  $v$ , given by (11) and (12), into the optical flow constraint (13). Solving for the downrange component given by each pixel yields:

$$X(\varepsilon_y, \varepsilon_z) = \frac{I_{\varepsilon_y}(-V + \varepsilon_y U) + I_{\varepsilon_z}(-W + \varepsilon_z U)}{- (I_{\varepsilon_y} A + I_{\varepsilon_z} B + I_t)} \quad (14)$$

where

$$A = \omega_x \varepsilon_z + \omega_y \varepsilon_y \varepsilon_z - \omega_z (1 + \varepsilon_y^2) \quad (15)$$

$$B = -\omega_x \varepsilon_y - \omega_z \varepsilon_y \varepsilon_z + \omega_y (1 + \varepsilon_z^2) \quad (16)$$

and

$$I_k = \frac{\partial I}{\partial k}, \quad k = \varepsilon_y, \varepsilon_z, t \quad (17)$$

are the spatial and temporal intensity gradients.

### Observations

It is difficult to analyze the range expression (14) since it depends on many parameters, each varying at different rates. However, some observations are worth making on the expected behavior of the range estimate. Although the range expression is mathematically correct, in practice the pixel ranges are often plagued with outliers or erroneous estimates. This algorithm relies on the spatial and temporal gradients as the means to extract information from the image, and problems may occur for several reasons. First, spatial / temporal intensity gradient estimates are susceptible to errors from noise, quantization, and resolution. Second, a combination of parameters could cause the denominator of (14) to approach zero, resulting in a singular expression. A third possibility for erroneous results is if there is insufficient motion in the image frame. Finally, highly textured areas of the image or areas with essentially homogeneous intensity do not contain the information required to develop the range estimate reliably. To challenge these issues, we propose the median-filtered, range-image method described in the next section.

## 4. Processing Flow

In this section we provide the general architecture for the proposed ranging from image processing flow. The description that follows is designed as a separate processing flow path from the *conventional* seeker processing employed, to highlight the VA-GNC functions. In practice, the conventional and VA-GNC processing could be combined for processing efficiency.

*Step 1 - Pixel Range Calculation:* The algorithm begins with the three translational and three rotational velocities from the INS, and the intensity image, all at time  $t$ . Using the three velocity terms, the three angular rates, and equation (14), calculate the raw range to each pixel in the image. In this step, wild points and outliers may appear, depending on image content. Care must be taken to guard against possible singularities in (14). Simulation results have shown that performance of the algorithm

in (14) is increased by averaging the spatial gradients over the time interval used to estimate the temporal gradient.

*Step 2-Local Outlier Rejection:* The pixel range calculation is known to produce wildpoints and outliers. These may occur due to homogeneous areas of intensity, at range discontinuities, gradient approximation error, or just due to statistical image noise attributes. In the case of the latter two, the wildpoints may be isolated to a small number of pixels in any given local area. Averaging (or other linear filtering methods) over the local area containing wildpoints may produce disastrous results, however *non-linear* filtering can effectively remove wildpoints leaving the resulting pixel ranges relatively unchanged [7]. The approach employed in this work is the median filter, which operates by replacing the raw range of the pixel under test by the median of the pixel ranges in a small neighborhood around the pixel under test. The neighborhood size used in this work is a  $5 \times 5$  mask centered about the pixel under test. The collection of pixel ranges after outlier rejection is referred to as the filtered range image.

*Step 3-Segmentation:* In order to calculate the range to the target, the portion of the image containing the target must be identified. This *segmentation* may be performed using the intensity image alone, or a *dual segmentation* approach using both the intensity and the range images may be used.

*Step 4-Extract Range of Interest:* Once the portion of the range image corresponding to the target is determined, the corresponding pixels are used to get a single range to the target. In the following examples, the range to the target is simply the average of the corresponding pixel ranges, excluding a border region, since the estimated pixel ranges at the border, even after outlier rejection, may be suspect due to the range discontinuity.

*Step 5-Use in Guidance:* Once the range to target is estimated using Steps 1 – 4, this range is incorporated in the guidance solution. This may be achieved by using a two-state Kalman filter to estimate the range rate, and using these to enhance proportional navigation, schedule guidance gains, and improve trajectory shaping.

## 5. Simulation Results

To provide initial performance results the scenario chosen is a simple non-rotational case designed to provide a proof of concept for this particular VA-GNC construct.

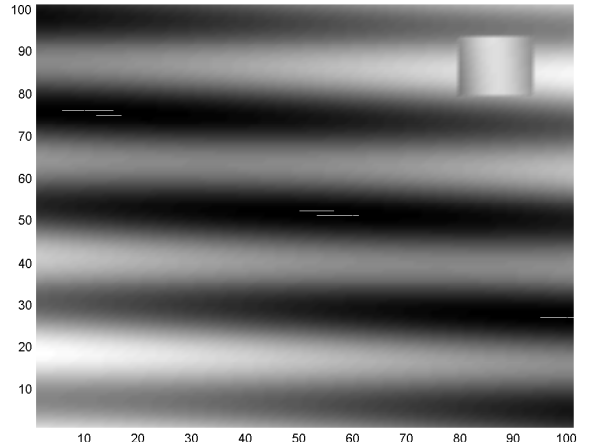


Figure 3: Intensity pattern at a point in flight.

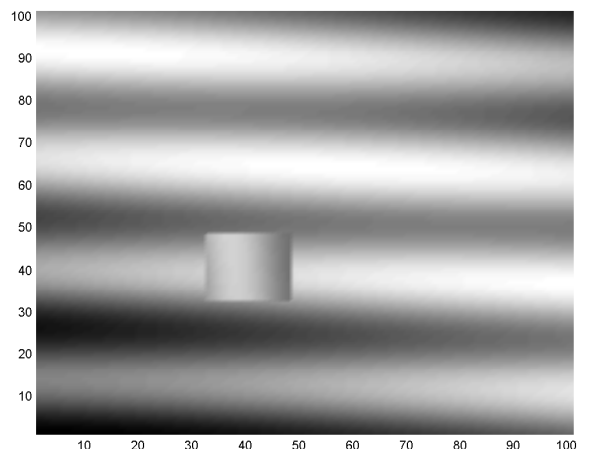


Figure 4: Intensity pattern at a later point in flight. Notice the ‘target’ has moved in the image as the camera moves across the ground.

The scene model consists of two types of returns: ‘background’ and ‘target.’ The background and target are both at a constant (but different) altitude, and both use the following model for intensity:

$$I(y_g, z_g) = \beta_1 \sin(\alpha_1 y_g + \alpha_2 z_g) + \beta_2 \sin(\alpha_3 y_g + \alpha_4 z_g) + n \quad (18)$$

where  $y_g$  and  $z_g$  represent the appropriate pixel locations on the ground for either background or target. The constants  $\beta_i$  and  $\alpha_i$ , which determine the amplitude and spatial wavelength of the intensity patterns, may be different for background and target. The target is 500 distance units above the background.

The seeker uses a  $101 \times 101$  pixel square array, with an angular span of  $\pm 2$  degrees in both dimensions.

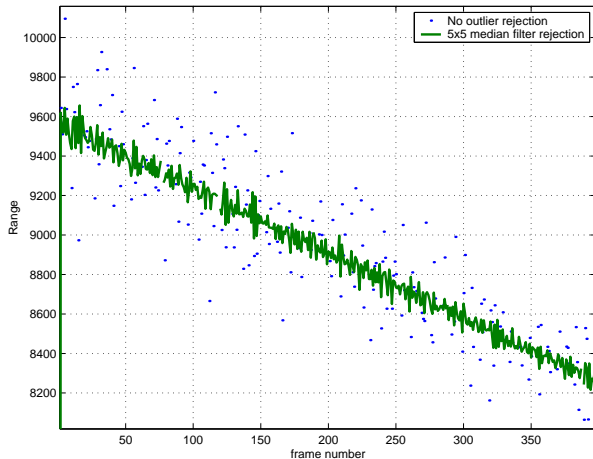


Figure 5: Range calculated with and without new outlier rejection. The local median filter (solid line) provides significantly better outlier rejection.

The intensity is quantized to 12 bits, and the frame rate is 30 frames per second.

In the scenario used for this study, the camera (seeker) is pointed directly at the ground. It is translating with velocities of 100, 25, and 25 distance units / second in the  $X$ ,  $Y$ , and  $Z$  directions, respectively (i.e. closing on the ground while translating over it). The initial altitude of the seeker is 10000 distance units. Figure 3 contains the intensity image at one point in flight. Notice the “target” near the upper right corner. At the later point in flight depicted in Figure 4, the camera has translated in all three dimensions, resulting in the target appearing both larger and in a different location.

Figure 5 shows the estimated range to the target, both with and without the median filter outlier rejection. The range estimate is the average of the pixel ranges corresponding to the target, excluding the border of the target. The benefits of the outlier rejection logic are immediately apparent. In addition, the range estimated with this approach provides a quality sufficient for tracking and guidance improvement.

An illustration of algorithm accuracy and its sensitivity are shown in Figure 6 using the same intensity model and camera motion described in Section 4. The percent range error (i.e. the ratio of the range error to the true range) is illustrated for various spatial sampling rates and camera velocities (using the total velocity normal to the camera axis). Spatial sampling is often measured in pixels per spatial wavelength ( $P/\lambda_s$ ), however, the spatial wavelength is difficult to define for the superposition of sinusoids (Equation 18) that compose our terrain

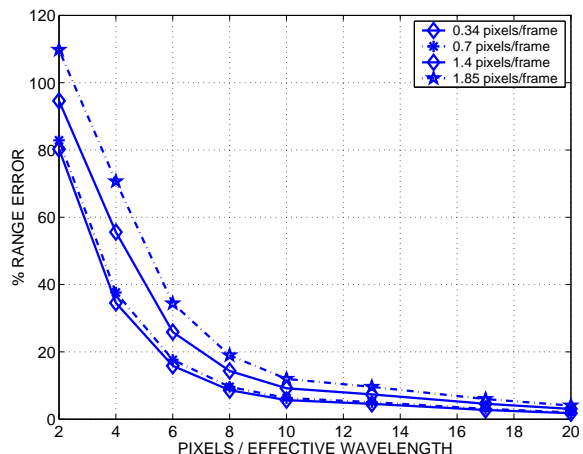


Figure 6: Percent range error vs. pixels per effective wavelength. Performance curves are shown for four different cross-range camera velocities.

intensity model. For this figure the spatial wavelength corresponds to the sinusoid with the smallest wavelength. Figure 6 shows the percent range error vs.  $P/\lambda_s$  for four cross-range (orthogonal to camera axis) velocities. The cross-range velocity is expressed in pixels traversed per frame,  $P/F = \frac{V_c \Delta T}{R \Delta P}$  where  $V_c$  is the cross-range velocity in distance per time,  $\Delta T$  is the frame period,  $R$  is the range to target, and  $\Delta P$  is the pixel size.

For the simple synthetic images used in this example, range accuracies of approximately 15% were observed when the spatial sampling  $P/\lambda_s$  was greater than about 10. The accuracy decreases as the cross-range velocity (and therefore  $P/F$ ) increases. For comparison, the extensive work described in [1] investigating different optical flow methods used values for  $P/\lambda_s$  of 6 to 16, and speeds corresponding to 1.02 to 1.63 pixels/frame. It should be noted that for simple terrain/target intensity models used in this investigation, the measures of performance should be taken as optimistic bounds. A more detailed analysis using more realistic images is in progress.

## 6. Conclusions

In this work we have presented an introduction to the concept of Vision-Augmented GNC, and have developed an initial mechanization of one VA-GNC approach. We have explored and improved a range-to-target estimation algorithm using a non-ranging seeker and non-linear filtering. Simulation results

using analytically generated images indicate that range estimation is possible to a fidelity required to increase guidance performance.

Future work in this area will investigate more realistic images and more complex camera motion profiles. In addition, this approach will be applied to the more interesting case of moving targets, and closed-loop dynamics with and without the VA-GNC approach evaluated. In addition, the effects of a gimbal-mounted seeker on this algorithm will be evaluated. As discussed in the “Observations” section, this algorithm is sensitive to the case of small motion in the image plane. Therefore this effort will focus on determining what seeker and scenario parameters are required to provide sufficient image motion for reliable algorithm performance. Future work in the general VA-GNC field is focused on quantifying the relationship between seeker capabilities and the additional information the seeker may provide to GNC.

### Acknowledgements

This research was sponsored by the Air Force Research Laboratory at Eglin Air Force Base, Florida. Simulation analysis was supported by Obaid Naveed Siddiqi of Sverdrup Technology.

### References

- [1] J.L. Barron, D.J. Fleet, and S.S. Beauchemin, “Systems and experiment: Performance of optical flow techniques,” *Intern. Journal of Computer Vision*, vol. 12, no. 1, pp. 43–77, December 1994.
- [2] A. Kumar, A.R. Tannenbaum, and G. Balas, “Optical flow: A curve evolution approach,” *IEEE Transactions on Image Processing*, vol. 5, no. 4, pp. 598–610, April 1996.
- [3] A.R. Bruss and B.K.P. Horn, “Passive navigation,” vol. 21, pp. 3–20, 1983.
- [4] S. Negahdaripour and B.K.P. Horn, “Determining 3-d motion of planar objects from image brightness patterns,” *Proceedings of the 9th Intern. Joint Conference on Artificial Intelligence*, pp. 898–901, August 1985.
- [5] D. Raviv and J. Albus, “A closed-form massively-parallel range-from-image-flow algorithm,” *IEEE Trans. On Systems, Man, and Cybernetics*, vol. 22, no. 2, pp. 322–327, March/April 1992.
- [6] B.K. Horn and B.G. Schunck, “Determining optical flow,” *Artificial Intelligence*, vol. 17, pp. 185–203, October 1981.
- [7] Milan Sonka, Vaclav Hlavac, and Roger Boyle, *Image processing, analysis, and machine vision*, chapter 4, pp. 74–76, Brooks / Cole Publishing Company, second edition, 1999.