A CLOSED-FORM MASSIVELY-PARALLEL RANGE-FROM-IMAGE-FLOW ALGORITHM

Daniel Raviv

Florida Atlantic University
The Robotics Center and
The Electrical Engineering Department
Boca Raton, FL 33431

James S. Albus

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

NIST

A CLOSED-FORM
MASSIVELY-PARALLEL
RANGE-FROM-IMAGE-FLOW
ALGORITHM

Daniel Raviv

Florida Atlantic University
The Robotics Center and
The Electrical Engineering Department
Boca Raton, FL 33431

James S. Albus

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

October 1990

A CLOSED-FORM MASSIVELY-PARALLEL
RANGE-FROM-IMAGE-FLOW ALGORITHM

Daniel Raviv* and James S. Albus**

*The Robotics Center and The Electrical Engineering Department
Florida Atlantic University, Boca Raton, FL 33431; and
Robot Systems Division
National Institute of Standards and Technology (NIST)
Bldg 220, Room B124, Gaithersburg, MD 20899

**Robot Systems Division
National Institute of Standards and Technology (NIST)
Bldg 220, Room B124, Gaithersburg, MD 20899

ABSTRACT

We provide a closed-form solution for obtaining 3-D structure of a scene for a given six degree of freedom motion of a camera. The solution is massively parallel, i.e., the range that corresponds to each pixel is dependent on the spatial and temporal changes in intensities of that pixel, and on the motion parameters of the camera. The measurements of the intensities are done in *a priori* known directions. The solution is for the general case of camera motion.

The derivation is based upon representing the image in the spherical coordinate system, although a similar approach could be taken for other image domains, e.g., the planar coordinate system. We comment on the amount of computations, errors and singular points of the solutions. We also suggest a practical way to significantly reduce and implement them.

1. INTRODUCTION

The perception of depth is critical to survival for creatures of nature. Depth perception is necessary for locomotion without collision in a complex natural 3-D world. It is also advantageous for recognizing objects. In a natural environment, image segmentation based on brightness boundaries of stationary objects is extremely difficult. It is confounded by shadows, surface texture, and markings that often produce brightness and color boundaries that are much more distinct than true object or surface boundaries. In the natural world, segmentation based on motion discontinuities is much more reliable than segmentation based on brightness discontinuities.

Many methods have been proposed for computing depth from visual input. Of these, only the methods of stereo and image flow can compute depth directly without any knowledge of, or understanding about, the contents of the image. Only image flow can compute depth from a single retina without simultaneously viewing the scene from two separate eyes. Depth from image flow is thus extremely important to prey animals, such as fish, most birds, and rabbits because their eyes are on the opposite sides of their heads. If they have stereo vision at all, it covers very little of the visual field. Yet such animals can swim, fly, or run rapidly through complex natural environments (such as reefs, forests, or thickets) filled with obstacles without collisions. They can also effectively flee predators and manuever relative to other members in flocks, herds, and schools.

Image flow is motion of an image on the retina of the eye, or on the photodetector array of a camera. Image flow can be caused by two phenomena: 1) motion of objects in the world, and 2) motion of the eye (camera) through the world. Motion of the eye can be of two types: translation and rotation. The algorithm described here applies to image flow produced by translation and rotation of the eye in a stationary world.

It is well known that image flow due to translation is by itself a sufficient visual cue for safe movement through a natural environment filled with complex 3-D obstacles. Helmholtz [1] observed as early as 1925 that monocular image flow provides direct and immediate perception of 3-D space. The rate of apparent motion of a stationary object in the world induced by translation of the eye through the world is inversely proportional to the distance from the eye to the object. Close objects which are potential collision hazards produce large image flow. Distant objects or close objects near the instantaneous axis of motion produce little or no image flow.

The psychophysics of image flow was extensively studied by Gibson [2], who treated the perception of space as the perception of a collection of surfaces, and considered motion relative to a surface to be the most fundamental of visual perceptions. Rogers and Graham [3] have shown that random-dot displays of monocular flow patterns produce the perception of solid oriented surfaces, and that image flow is adequate for shape and depth perception with no other depth information.

Methods for computing image flow from time-varying images have been proposed by Horn and Schunck [4], Longuet-Higgins and Prazdny [5], Hildreth [6], Waxman and Wohn [7], Waxman, Kamgar-Parsi, and Subbarao [8], and others. In most of this

work, camera motion is assumed to be unknown. Overcoming this assumption requires complex mathematical analyses that are subject to error from quantization noise, and are too computationally intensive for practical real-time implementation. Recent work by Bolles [9], Matthies, Szeliski and Kanade [10], Albus and Hong [13], Rangachar, Hong, and Herman [16], and Heel [15] has been based on knowing camera motion. Combined with Kalman filtering techniques this approach has produced dense (iconic pixel based) range maps with relative RMS error values of less than three percent. Except for Heel, the approach has been defined only for translation without rotation with the camera scan lines aligned with an epipolar plane.

The algorithm given here treats the general case of simultaneous camera translation and rotation. It assumes both to be known. We assert this to be a good assumption, because most creatures of nature (and robot vehicles as well) possess non-visual vestibular and proprioceptive sensors, and other sensory-motor systems that enable them to estimate three-axis eye rotation, as well as ground speed, air speed, or water speed independently of visual input. This estimate can be refined by the addition of global visual data integrated over the entire visual field. This paper thus assumes that there exists, either in the brain or the robot system, a mechanism by which both visual and non-visual information is fused into a single best estimate of eye translation and rotation.

This paper provides a closed form solution to the computation of range to every pixel, and suggests a massively-parallel computational structure that could make range from image flow a practical mechanism for enabling high speed robot mobility. The paper begins by describing the camera-centered coordinate systems that are used here, followed by deriving expressions that relate optical flow and distance of a point to six-degree-of-freedom camera motion. When the optical flow constraint (as described by Horn and Schunck [4]) is added to these equations, a pixel-based closed-form solution for the 3D location of a point in the camera coordinate system is obtained.

2. EQUATIONS OF MOTION AND OPTICAL FLOW

This section describes the equations that relate a point in 3-D space to the projection of that point in the image for general six-degree-of-freedom motion of the camera. Some of the equations are well known [11].

Assume a moving camera in a stationary environment with the coordinate system fixed with respect to the camera as shown in Figure 1. Assume a pinhole camera model, such that the pinhole of the camera is at the origin of the coordinate system. We derive the optical flow components in the spherical coordinate system $(R\,\theta\phi)$. In this frame, angular velocities ($\dot\theta$ and $\dot\phi$) of any point in space, say P, are identical to the optical flow values $\dot\theta$ and $\dot\phi$ in the image domain. See Figure 2.

The relationship between the image flow rates in the spherical domain of Figure 2 and the velocity of a point in the cartesian coordinates of Figure 1 is given by the following equations:

Let the instantaneous cartesian coordinates of the point P be $\mathbf{R} = (X,Y,Z)^T$ (where the superscript T denotes transpose). Let the instantaneous translational velocity of the camera be $\mathbf{t} = (U,V,W)^T$ and the instantaneous angular velocity be $\omega = (A,B,C)^T$. Then the velocity vector $\mathbf{V}$ of the point P with respect to the XYZ coordinate system is (see [11]):

$$\mathbf{V} = -\mathbf{t} - \omega \times \mathbf{R} \tag{1}$$

or:

$$V_X = -U - BZ + CY \tag{2}$$

$$V_Y = -V - CX + AZ \tag{3}$$

$$V_Z = -W - AY + BX \tag{4}$$

where V_X, V_Y, and V_Z are the components of the velocity vector $\mathbf{V}$ along the X, Y, and Z directions respectively. Let $s_\theta = \sin\theta$, $c_\theta = \cos\theta$, $s_\phi = \sin\phi$, $c_\phi = \cos\phi$, and $R = |\mathbf{R}|$. To convert from XYZ to $R\theta\phi$ coordinates we use the relations:

$$X = R\, c_\phi\, c_\theta \tag{5}$$

$$Y = R\, c_\phi\, s_\theta \tag{6}$$

$$Z = R\, s_\phi \tag{7}$$

In order to find the optical flow of a 3-D point in $R\theta\phi$ coordinates, we use the following relations and transformations (see [12] and Figure 1). Let V_R, V_θ, and V_ϕ be the components of the vector $\mathbf{V}$ in spherical coordinates, and

$$\begin{bmatrix} V_R \\ V_\theta \\ V_\phi \end{bmatrix} = \begin{bmatrix} c_\phi & 0 & s_\phi \\ 0 & 1 & 0 \\ -s_\phi & 0 & c_\phi \end{bmatrix} \begin{bmatrix} c_\theta & s_\theta & 0 \\ -s_\theta & c_\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} V_X \\ V_Y \\ V_Z \end{bmatrix} \tag{8}$$

Also from [12]

$$V_R = \dot{R}, \tag{9}$$

$$V_\theta = R\,\dot{\theta}\, c_\phi, \tag{10}$$

$$V_\phi = R\,\dot{\phi}, \tag{11}$$

where dot denotes derivative with respect to time.

Using equations (2)-(11) the following expressions are obtained:

$$\begin{bmatrix} R\,\dot{\theta}\, c_\phi \\ R\,\dot{\phi} \end{bmatrix} = \begin{bmatrix} -s_\theta & c_\theta & 0 \\ -s_\phi\, c_\theta & -s_\phi\, s_\theta & c_\phi \end{bmatrix} \begin{bmatrix} -U - B\,R\,s_\phi + C\,R\,c_\phi\, s_\theta \\ -V - C\,R\,c_\phi\, c_\theta + A\,R\,s_\phi \\ -W - A\,R\,c_\phi\, s_\theta + B\,R\,c_\phi\, c_\theta \end{bmatrix} \tag{12}$$

There are three unknowns in equation set (12): R, $\dot{\theta}$, and $\dot{\phi}$. For each pixel, θ and ϕ are known. The motion parameters (A,B,C,U,V,W) are also known.

- 3 -

3. SOLUTIONS

3.1 GRADIENT-BASED METHOD

3.1.1 THE OPTICAL FLOW CONSTRAINT

If brightness I varies smoothly with θ, ϕ, and t, then we obtain the optical flow constraint equation (see [4]):

$$\frac{\partial I}{\partial \theta} \dot{\theta} + \frac{\partial I}{\partial \phi} \dot{\phi} + \frac{\partial I}{\partial t} = 0 \tag{13}$$

or

$$I_\theta \dot{\theta} + I_\phi \dot{\phi} + I_t = 0 \tag{14}$$

where

$$I_\theta = \frac{\partial I}{\partial \theta} \qquad I_\phi = \frac{\partial I}{\partial \phi} \qquad I_t = \frac{\partial I}{\partial t}.$$

I_θ, I_ϕ, and I_t can be approximated from the image sequence as follows. For a given pixel (θ_i, ϕ_j) (Figure 3) in the image at time instant t_k

$$\frac{\partial I}{\partial \theta_i} \approx \frac{I(\theta_i, \phi_j, t_k) - I(\theta_{i-1}, \phi_j, t_k)}{\theta_i - \theta_{i-1}} \tag{15}$$

$$\frac{\partial I}{\partial \phi_j} \approx \frac{I(\theta_i, \phi_j, t_k) - I(\theta_i, \phi_{j-1}, t_k)}{\phi_j - \phi_{j-1}} \tag{16}$$

$$\frac{\partial I}{\partial t_k} \approx \frac{I(\theta_i, \phi_j, t_k) - I(\theta_i, \phi_j, t_{k-1})}{t_k - t_{k-1}} \tag{17}$$

3.1.2 DEPTH COMPUTATIONS

Equation set (12) together with equation (14) form a set of three equations with three unknowns R, θ, and ϕ. Solving these equations for R yields:

$$R = \frac{(Us_\theta - Vc_\theta)I_\theta + c_\phi(Us_\phi c_\theta + Vs_\phi s_\theta - Wc_\phi)I_\phi}{-c_\phi I_t - c_\phi(Bc_\theta - As_\theta)I_\phi + (Cc_\phi - s_\phi(Bs_\theta + Ac_\theta))I_\theta} \tag{18}$$

The meaning of solution (18) is that if the location of the pixel in the image, the motion parameters of the camera, and the spatial and temporal intensity changes are given, then the depth of the corresponding point in 3-D can be obtained.

At first inspection, 21 multiplications and divisions, and 10 summations and subtractions appear to be necessary for each pixel at each instant of time. However, for a given pixel, θ and ϕ are constants and therefore all functions of θ and ϕ such as $c_\theta c_\phi$ can be precomputed. Also, the rate of change of the motion parameters A, B, C, U, V, W is usually much lower than the rate of change of the image variables I_θ, I_ϕ and I_t. Thus the update rate of the motion parameters can be significantly smaller than the computation of R. For example, for a translation in the XY plane, and rotation about the Z axis (i.e., A, B, and W are zero), U, V, and C may be constants

for some period of time but I_θ, I_ϕ and I_t change. Practically, accelerations (either translational or rotational) are limited due to mass (or inertia) constraints. Thus, any function of θ,ϕ,A,B,C,U,V,W will typically have an update rate that an order of magnitude smaller than the intensity changes rate. Equation (18) can be rewriten as follows:

$$R = \frac{a_1 I_\theta + a_2 I_\phi}{a_3 I_t + a_4 I_\phi + a_5 I_\theta} \tag{19}$$

where $a_1 = Us_\theta - Vc_\theta$, $a_2 = c_\phi(Us_\phi c_\theta + Vs_\phi s_\theta - Wc_\phi)$, $a_3 = -c_\phi$, $a_4 = c_\phi(As_\theta - Bc_\theta)$, and $a_5 = Cc_\phi - s_\phi(Bs_\theta + Ac_\theta)$. By observing equation (19) it is clear that for given I_θ, I_ϕ and I_t, at the fast sampling rate there are only five multiplications, one division, and three summations per pixel.

For each point in space there is a set of three equations that result in a reconstruction for R. The expression for R that corresponds to a pixel in the image can be computed using a pixel-based processor, and special purpose computers may run in parallel to reconstruct all visible points in space. Figure 4 shows a possible multi-rate structure. For each pixel there is a "fast" computer, and a "slow" computer. The first part of the "fast computer" extracts I_θ, I_ϕ, and I_t, and the second part of this computer computes the distance R that corresponds to this pixel. The latter part accepts spatial and temporal intensity changes functions as well as motion parameter functions for this particular pixel (i.e., $a_1, \ldots, a_5$ in equation (19)), and combines them for instantaneous distance extraction. The "slow" computer generates functions of θ,ϕ,A,B,C,U,V,W at a low update rate.

It is difficult to analyze the error since there are eight parameters $(\theta,\phi,A,B,C,U,V,W)$ to consider. However, there are singular points, i.e., combinations of the parameters and variables values that cause the denominator to be zero. There are also points that cannot be reconstructed, e.g., those that lie exactly on the direction of the motion vector.

3.1.3 A SPECIAL CASE

In order to intuitively understand expression (18), we show a special case. Suppose that the camera undergoes translation U and V in the instantaneous XY plane and rotation C about the Z axis only, i.e., $A = B = W = 0$. Assume $\phi = 0$, i.e., the pixel which is analyzed lies on the XY plane. Substituting these values in (18) results in the following expression for R:

$$R = \frac{(Us_\theta - Vc_\theta)I_\theta}{-I_t + CI_\theta} \tag{20}$$

or (assuming $I_\theta \neq 0$)

$$R = \frac{Us_\theta - Vc_\theta}{-\dot{\theta} + C} \tag{21}$$

where $\dot{\theta} = \dfrac{I_t}{I_\theta}$ Given that $\dot{\theta} \neq C$, equation (21) is a solution for any point in the XY plane

- 5 -

when translating in this plane and rotating about the Z axis.

3.2 CORRELATION METHOD

Computation of image flow by cross-correlation is an alternative to using the optical flow constraint equation (13). Cross correlation also requires that the direction of the flow vector be known. Assume that for any point in space viewed by a pixel at an instant of time, the direction of the image flow vector can be found. i.e., assume that the ratio $\dfrac{\dot\theta}{\dot\phi}$ is known. A second pixel can then be selected at some angular distance ΔA from the first along the flow direction. One can then compute a cross-correlation function between the first and second pixels. This will peak at some delay interval τ_{max} and the image flow rate $\dfrac{dA}{dt}$ of the point in space is then given by (see [13]):

$$\frac{dA}{dt} = \lim_{\Delta A \to 0} \frac{\Delta A}{\tau_{max}} \tag{22}$$

It can be approximated by:

$$\frac{dA}{dt} \approx \frac{\Delta A}{\tau_{max}} \tag{23}$$

After computing $\dfrac{dA}{dt}$ and the ratio $\dfrac{\dot\theta}{\dot\phi}$, the values of $\dot\phi$ and $\dot\theta$ can be computed from:

$$\dot\phi = \frac{dA}{dt} \cos\left(\tan^{-1}\frac{\dot\theta}{\dot\phi}\right) \tag{24}$$

$$\dot\theta = \frac{dA}{dt} \sin\left(\tan^{-1}\frac{\dot\theta}{\dot\phi}\right) \tag{25}$$

The value of R can be computed from equation set (12).

For a general motion $\dfrac{\dot\theta}{\dot\phi}$ is unknown. Thus several pixels in the neighborhood of the first one has to be selected. The pixel for which the cross correlation function peaks provides $\dfrac{\dot\theta}{\dot\phi}$, ΔA, and τ_{max}, and the same (previously described) procedure can be used to compute R.

4. DISCUSSION

Computing range from image flow by the method of Horn and Schunck [4] has the advantage of simplicity. Unfortunately, it tends to produce noisy and inaccurate results. There are four sources of noise and errors:

First, the sampled data system approximations to $\dfrac{dI}{dt}$, $\dfrac{dI}{d\theta}$, and $\dfrac{dI}{d\phi}$, are subject to digitization noise and Nyquist sampling frequency limitations. If $I(\theta,\phi,t)$ contains spatial (or temporal) frequency components that are higher than half the spatial (or temporal) sampling frequency, aliasing may occur. Of course, the natural world is infinitely

- 6 -

rich in detail. Thus low-pass filtering must be performed on natural images to smooth them before the Horn and Schunck method can be used. The resultant blurring of the images removes sharp features that contain the most accurate information about image flow.

Second, except in the vicinity of brightness edges, $\frac{dI}{dt}$ is small, and may be zero. Division by small numbers magnifies errors.

Third, the sensitivity of photodetectors in any array is not uniform. The difference in signal from two adjacent photo-detectors thus is not necessarily an accurate measure of the difference in illumination. In fact, in the biological eye, where accomodation occurs in the retinal photodetectors, if the eye is fixated for an extended period, accommodation causes the image to fade away completely. In this case, difference in output of adjacent photodetectors provides no information at all about the spatial brightness gradient.

Fourth, thermal noise in photodetectors is not "white", but heavily biased toward low frequencies. The difference in signal from two adjacent photodetectors is thus subject to large amounts of differential low frequency noise. Thus range from image flow by the Horn and Schunck method tends to be inaccurate for smooth images, and unreliable for sharply focused natural images.

Computing range from image flow by the method of cross-correlation can be much more accurate and noise free. It is, however, more complicated and computationally intensive. The selection of which two pixels to correlate is a function of the translational and rotational velocities. For each pixel where correlation is performed, a delay line, a cross correlator, and a correlation peak detector are required [13]. The accuracy is limited by the ratio of the sampling rate to the delay τ_{max}. For one percent accuracy, the intersample interval must be 100 times less than τ_{max}.

In order to reduce noise and increase accuracy, temporal integration methods can be used. Methods for performing temporal integration include Kalman filtering [9],[10], and computing the running average in world coordinates [13]. Temporal integration by running averaging has been shown to reduce inaccuracies in the computation of depth from image flow by the Horn and Schunck method by about an order of magnitude, from about 30% absolute depth error in the raw data to about 3% absolute depth error in the averaged data [13]. This is comparable to the results achieved by the much more computationally intensive cross-correlation and Kalman filtering methods of Matties et al [10]. Recent results achieved by a new correlation-integration method [14] show absolute depth errors of less than 0.6%.

5. CONCLUSIONS

When eye velocity is known, image flow can be a simple, robust, and powerful method for generating dense range images. For complex natural scenes, dense range images computed from image flow should prove to be much easier to segment that brightness images. The combination of image flow techniques with more traditional methods thus promises to vastly improve the performance of machine vision systems.

The ability to compute dense range images without knowledge of, or information about, the contents of scenes may explain the ability of creatures of nature, such as flying insects, birds, fish, and mammals, to manuever in a complex natural world. Accuracy of range from image flow is proportional to accuracy of velocity. For many important robotics applications, such as driving unmanned ground vehicles, velocity is known well enough to compute useful range from image flow. Image flow may thus become a practical method for obstacle avoidance in unmanned vehicles.

When eye rotation is zero, the computation of image flow, and range from image flow, is quite simple. This may explain why many creatures of nature have a vestibular reflex that can mechanically stabilize the eye against rotation. There is also psychophysical evidence for neuronal mechanisms that scroll the retinal image so as to cause the visual world to be subjectively perceived as stationary despite eye rotation that causes the retinal image to move about.

This paper has shown how depth from image flow can be computed directly from local image parameters, given knowledge of eye translation and rotaion. The closed-form algorithm that has been presented is amenable to implementation on massively-parallel single-instruction multiple data (SIMD) computing machines. It could, in fact, be computed by a single layer of neurons of the type known to exist in the biological brain.

Future work will focus on measures of performance of the above algorithms in an environment filled with a variety of natural objects such as trees, bushes, rocks, streams, dirt, leaves, and grass. Tests will be made that simulate tasks such as driving, running, or flying at high speeds, like insects and birds routinely do, through complex natural environments such as woods and fields.

REFERENCES

[1] Helmholtz, H. von, *Treatise on Physiological Optics* (translated by J.P.C. Southall), Dover Publications, New York, 1925.

[2] Gibson, J.J., *The Ecological Approach Visual Perception*, Cornell University Press, Ithaca, N.Y., 1966.

[3] Rogers, B., and Graham, M., "Motion parallax as an independent cue for depth", Perception 8, pp. 125-134, 1979.

[4] Horn, B.K.P., and Schunck, B.G., "Determining Optical Flow", Artificial Intelligence, Vol. 17, pp.185-204, 1981.

[5] Longuet-Higgins, H.C., and Prazdny, K., "The interpretationof a moving retinal image", Proc. Royal Society of London, B208, pp. 385-397, 1980.

[6] Hildreth, E.C., *The Measurement of Visual Motion*, Ph.D. Thesis, MIT, August

1983.

[7] Waxman, A.M., and Wohn, K., "Image flow theory: a framework for 3-D inference from time-varying imagery", in *Advances in Computer Vision*, Vol. 1 (C. Brown, Editor), Lawrence Erlbaum Associates Publishers, 1988.

[8] Waxman, A.M., Kamgar-Parsi, B., and Subbarao, "Closed form solution to image flow equations", International Journal of Computer Vision, 1, pp. 239-258, 1987.

[9] Bolles, R.C., Baker, H.H., , and Marimont, D.H., "Epipolar-plane image analysis: an approach to determining structure from motion", International Journal of Computer Vision, 1, pp. 7-55, 1987.

[10] Matthies, L., Szelinski, R., and Kanade, T., "Kalman filtering based algorithms for estimating depth from image sequences", CMU Technical Report CMU-CS-87-185, 1987.

[11] Horn, B.K.P., *Robot Vision*, MIT Press, 1986.

[12] Meriam, J.L., *Dynamics*, John-Wiley & Sons, 1975.

[13] Albus, J.S., and Hong, T.H., "Motion, Depth, and Image Flow", Proceedings IEEE International Conference on Robotics and Automation, Cincinnati, OH, 1990.

[14] Raviv, D., "Parallel Algorithm for 3D Surface Reconstruction", Proceedings, SPIE Symposium on Advances in Intelligent Robotics Systems, Philadelphia, PA, Nov. 1989.

[15] Heel, J., "Direct Dynamic Motion Vision", Proceedings IEEE International Conference on Robotics and Automation, Cincinnati, OH, 1990.

[16] Rangachar, R., Hong, T., and Herman M., "Real-Time Differential Range Estimation Based on Time-Space Imagery Using PIPE", Proceedings, SPIE Symposium on Image Processing II: Algorithms, Architectures, and Applications, Vol. 1295, Orlando, FL, April 1990.

NIST-114A (REV. 3-89)	U.S. DEPARTMENT OF COMMERCE NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY	1. PUBLICATION OR REPORT NUMBER NISTIR 4450
		2. PERFORMING ORGANIZATION REPORT NUMBER
	BIBLIOGRAPHIC DATA SHEET	3. PUBLICATION DATE OCTOBER 1990

4. TITLE AND SUBTITLE

A Closed-Form Massively-Parallel Range-From-Image-Flow Algorithm

5. AUTHOR(S)

Daniel Raviv and James S. Albus

6. PERFORMING ORGANIZATION (IF JOINT OR OTHER THAN NIST, SEE INSTRUCTIONS)	7. CONTRACT/GRANT NUMBER
U.S. DEPARTMENT OF COMMERCE NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY GAITHERSBURG, MD 20899	
	8. TYPE OF REPORT AND PERIOD COVERED

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

10. SUPPLEMENTARY NOTES

☐ DOCUMENT DESCRIBES A COMPUTER PROGRAM; SF-185, FIPS SOFTWARE SUMMARY, IS ATTACHED.

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

We provide a closed-form solution for obtaining 3D structure of a scene for a given six degree of freedom motion of camera. The solution is massively parallel, i.e., the range that corresponds to each pixel is dependent on the spatial and temporal changes in intensities of that pixel, and on the motion parameters of the camera. The measurements of the intensities are done in a priori known directions. The solution is for the general case of camera motion.

The derivation is based upon representing the image in the spherical coordinate system, although a similar approach could be taken for other image domains, e.g., the planar coordinate system. We comment on the amount of computations, errors and singular points of the solutions. We also suggest a practical way to significantly reduce and implement them.

12. KEY WORDS (6 TO 12 ENTRIES; ALPHABETICAL ORDER; CAPITALIZE ONLY PROPER NAMES; AND SEPARATE KEY WORDS BY SEMICOLONS)

3-D reconstruction, optical-flow range-from-image-flow, visual motion, global-flow, mobility

13. AVAILABILITY	14. NUMBER OF PRINTED PAGES
[X] UNLIMITED	14
☐ FOR OFFICIAL DISTRIBUTION. DO NOT RELEASE TO NATIONAL TECHNICAL INFORMATION SERVICE (NTIS).	
☐ ORDER FROM SUPERINTENDENT OF DOCUMENTS, U.S. GOVERNMENT PRINTING OFFICE, WASHINGTON, DC 20402.	15. PRICE
[X] ORDER FROM NATIONAL TECHNICAL INFORMATION SERVICE (NTIS), SPRINGFIELD, VA 22161.	A02

ELECTRONIC FORM