Image Matching Using Corresponding Point Measurements

Nils N. Haag
Science Applications International Corporation, 700 South Babcock Street, Melbourne, FL 32901

ABSTRACT: This article describes a relationship which enables one to determine whether the image of a three-dimensional object in one photograph matches the image of an object in another photograph. The image matching algorithm requires the measurement of what appear to be nine corresponding image points in the two photographs. The relationship depends only upon the image point coordinates and does not require knowledge of the object space point coordinates or the interior and exterior orientation of the two cameras. The relationship may be used for change detection or classification, and it is particularly useful if there is a large disparity in the perspectives of the two photographs. A by-product of the development is the equation of the epipolar line which is useful in automated image search and correlation algorithms.

THIS STUDY ADDRESSES THE PROBLEM of matching the image of a three-dimensional object in two photographs taken from different perspectives. The formulated technique is evaluated by matching the simulated reconnaissance camera image coordinates of two ships, neither of whose object space coordinates is known.

Perhaps the best method for determining whether two photographs match is to measure what appear to be five or more corresponding image points in the two photographs and then to perform a conventional relative orientation using the standard photogrammetric coplanarity equations. If the images match, the measurement residuals will be consistent with the known measurement error variance. A practical difficulty in applying this technique is in determining an initial estimate of the relative orientation parameters of the two cameras. This is required to rigorously solve the coplanarity equations which are non-linear in the relative camera position coordinates and in the orientation angles defining the rotation of one camera relative to the other. Initial values are particularly difficult to estimate when there is a large, unknown disparity in the perspectives of the two photographs. Another practical difficulty is encountered when the relative orientation adjustment is weakly determinant because of narrow image ray bundles which expose a small image compared to the photograph format. The photographs simulated for the image matching presented here contain both large perspective differences and narrow image ray bundles.

An alternative approach to the conventional method is to compute rather than to guess at the initial estimates of the relative orientation. This can be accomplished by using eight corresponding image points (Longuet-Higgins, 1981). This technique is unique in that it provides a closed form solution by solving eight equations for coefficients containing linear combinations of the camera position parameters and the rotation matrix elements. The camera position parameters and orientation angles are then extracted from the coefficients. This approach also fails for the simulations presented here, apparently because of its sensitivity to small, random measurement errors and narrow image ray bundles. Another solution for the approximate orientation is offered by Hinsken (1988). It is based upon an iterative solution for the quaternions defining the orientation matrix. The sensitivity of this algorithm was not investigated.

This study develops a new image matching algorithm using nine corresponding image points to overcome the problems stated above. The derived relationship among the image coordinate mesurements in the two photographs is invariant, never changing regardless of the relative orientation of the two cameras. This means that the relationship is always satisfied if the image measurements in the two photographs are of the same object. If the image measurements are of different objects, the relationship is not satisfied.

Recent studies by Barrett *et al.* (1990, 1991) pointed out general methods for determining invariant relationships in imagery. The second study also addresses the practical problem of matching images of aircraft where the object points are contained in a plane of symmetry defined by the wing tips, elevator tips, etc. This study also investigates images of three-dimensional objects for the special case in which the image planes of the stereo pair are coincident. A similar study by Haag *et al.* (1991) develops an invariant relationship in side-looking synthetic aperture radar imagery applicable to radar images of three-dimensional objects.

In the study presented here the general problem of matching a three-dimensional object appearing in two frame camera photographs is addressed by first deriving an invariant mathematical relationship using Barrett's method of homogeneous equations and then formulating the condition equation of the epipolar line. This is followed by a description of the image matching algorithm. An example of applying the algorithm is then presented followed by a discussion of the attributes of the relationship and certain degenerate cases.

FORMULATION OF THE INVARIANT RELATIONSHIP

Let (x_{ij}, y_{ij}) represent the frame camera image coordinates of point j in photograph i which have been reduced to the principal point, and let f_i represent the camera focal length. The relationship can be shown to be independent of the principal point location by including its offset in the image coordinates, but its introduction unduly complicates the development which follows.

Define the column vector

$$\mathbf{V}_{ij} \equiv (x_{ij}, y_{ij}, -f_i)^{\mathrm{T}}. \tag{1}$$

Next, define an arbitrary reference coordinate system so that (X_i, Y_i, Z_i) are the camera coordinates in this sytem and $\mathbf{T}_i$ is the 3 by 3 orthogonal rotation matrix which rotates the camera axes to the reference system axes. The direction cosines of an

PHOTOGRAMMETRIC ENGINEERING & REMOTE SENSING,
Vol. 58, No. 12, December 1992, pp. 1693–1697.

0099-1112/92/5812–1693$03.00/0

image ray from camera i to object point j in this system are given by

$$(l_{ij}, m_{ij}, n_{ij})^T = (1/\mathbf{r}_{ij})\mathbf{T}_i\mathbf{V}_{ij} \tag{2}$$

in which

$$\mathbf{r}_{ij} = (\mathbf{V}_{ij}^T\mathbf{V}_{ij})^{\frac{1}{2}}. \tag{3}$$

A unit vector along the image ray is denoted by

$$\mathbf{R}_{ij} = [l_{ij}, m_{ij}, n_{ij}], \tag{4}$$

and the vector between the two cameras is similarly denoted by

$$\mathbf{P} = [(X_2 - X_1),(Y_2 - Y_1), (Z_2 - Z_2)] \equiv [p_1,p_2,p_3]. \tag{5}$$

The familiar coplanarity condition in photogrammetry requires that the unit vectors along the image rays and the vector between the two cameras lie in the same plane so that the scalar triple product

$$\mathbf{P} \cdot (\mathbf{R}_{1j} \times \mathbf{R}_{2j}) = 0 \tag{6}$$

or, what is the same, the determinant

$$\det \begin{bmatrix} p_1 & p_2 & p_3 \\ l_{1j} & m_{1j} & n_{1j} \\ l_{2j} & m_{2j} & n_{2j} \end{bmatrix} = 0. \tag{7}$$

Expanding the determinant results in

$$(l_{1j}, m_{1j}, n_{1j})\mathbf{S}(l_{2j}, m_{2j}, n_{2j})^T = 0 \tag{8}$$

where $\mathbf{S}$ is the 3 by 3 skew symmetric matrix given by

$$\mathbf{S} = \begin{bmatrix} 0 & p_3 & -p_2 \\ -p_3 & 0 & p_1 \\ p_2 & -p_1 & 0 \end{bmatrix}. \tag{9}$$

If one take the origin of the coordinate system at camera 1 (i=1), the coordinates (X_1, Y_1, Z_1) become zero and the matrix $\mathbf{T}_1$ becomes the identity matrix. Using this assumption, substituting Equation 2 into Equation 8, and multiplying through by $\mathbf{r}_{1j}\mathbf{r}_{2j}$ results in

$$\underset{1\times 3}{\mathbf{V}_{1j}^T} \; \underset{3\times 3}{\mathbf{A}} \; \underset{3\times 1}{\mathbf{V}_{2j}} = 0 \tag{10}$$

in which

$$\mathbf{A} = \mathbf{S}\mathbf{T}_2 \tag{11}$$

Equation 6 is traditionally used for determining the relative orientation. It is non-linear in the orientation angles and camera coordinates and must be solved iteratively. Equation 10, however, contains nine unique, non-zero elements of $\mathbf{A}$ (a_{pq}) which themselves define the relative orientation of the two cameras. Longuet-Higgins (1981) derives Equation 10, eliminates one of the variables, and solves eight equations using eight points. He then extracts the orientation angles and camera coordinates from his solution, thus providing a non-iterative approach to the problem. We will depart from his development here.

We first substitute Equation 1 into Equation 10 and collect terms, resulting in

$$\underset{1\times 9}{\mathbf{U}_j} \; \underset{9\times 1}{\mathbf{W}} = 0 \tag{12}$$

where

$$\mathbf{U}_j = (x_{1j}, y_{1j}, x_{2j}, y_{2j}, x_{1j}x_{2j}, x_{1j}y_{2j}, y_{1j}x_{2j}, y_{1j}y_{2j}, 1) \tag{13}$$

and

$$\mathbf{W} = (-f_2a_{13}, -f_2a_{23}, -f_1a_{31}, -f_1a_{32}, a_{11}, a_{12}, a_{21}, a_{22}, f_1f_2a_{33})^T. \tag{14}$$

It is important to note that the row vectors $\mathbf{U}_j$ which we will be concerned with contain only the image coordinates of the point j in the two photographs.

Now, consider nine points and write

$$\underset{9\times 9}{\begin{bmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{U}_9 \end{bmatrix}} \underset{9\times 1}{\mathbf{W}} \equiv \mathbf{UW} = \underset{9\times 1}{\begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{bmatrix}} \tag{15}$$

where $\mathbf{U}$ is defined as the matrix on the left-hand side of the equation. This sytem of nine linear homogeneous equations clearly has a solution in the camera related parameters $\mathbf{W}$ (degenerate cases are considered later). The coefficient matrix $\mathbf{U}$ is singular and the determinant

$$\mathbf{D} \equiv \det \mathbf{U} = 0. \tag{16}$$

Because the rows $\mathbf{U}_j$ of $\mathbf{U}$ in this equation contain only image point coordinates in two photographs, we have a relationship which is invariant (equal to zero) with respect to the interior and exterior orientation of the two cameras.

Equation 16, in principle, is satisfied if the image measurements in the two photographs are of points on the same object. If the image measurements in one photograph are of a slightly different object, the relationship is not satisfied and the value of the determinant is not zero. Equation 16 therefore forms the basis for the image matching algorithm.

FORMULATION OF THE EPIPOLAR LINE CONDITION EQUATION

It is useful at this point to introduce an alternate method for evaluating the determinant in Equation 16. Let c_{kl} represent the cofactor of the kth row and lth column of the matrix $\mathbf{U}$, then the determinant may be evaluated by cofactors of row k, resulting in

$$\mathbf{D} = \det \mathbf{U} = \underset{1\times 9}{\mathbf{U}_k} \; \underset{9\times 1}{\mathbf{C}_k} = 0 \tag{17}$$

in which

$$\mathbf{C}_k = (c_{k1}, c_{k2}, \ldots, c_{k9})^T. \tag{18}$$

Expanding Equation 17 and collecting terms results in the equation of a line (E.B. Barrett, private communication, 1991) which is demonstrated by simulation to be the epipolar line. The resulting equation of the epipolar line is given by

$$\mathbf{D} = A^\circ x_{2k} + B^\circ y_{2k} + C^\circ = 0 \tag{19}$$

where the coefficients are

$$A^\circ = c_{k3} + x_{1k}c_{k5} + y_{1k}c_{k7}, \tag{20}$$

$$B^\circ = c_{k4} + x_{1k}c_{k6} + y_{1k}c_{k8}, \text{ and} \tag{21}$$

$$C^\circ = c_{k9} + x_{1k}c_{k1} + y_{1k}c_{k2}. \tag{22}$$

This means that if one is given the measurements of eight corresponding image points j ($1 \le j \le 9$) in the two photographs and the measured coordinates of a ninth point k not equal to j ($1 \le k \le 9$) in photograph 1 (x_{1k}, y_{1k}) then the corresponding coordinates of the point in photograph 2 (x_{2k}, y_{2k}) must lie on the epipolar line given by Equation 19. It is helpful to think of the process described above as follows. Take a set of eight corresponding image points and perform a "relative orientation" (compute the cofactors c_{kl}). Next, take another image point in photograph 1 and use the "relative orientation" to find the equation of the epipolar line in photograph 2.

IMAGE MATCHING ALGORITHM

Equation 16 (or, equivalently, Equation 19) is the condition equation used to determine whether the two images match. Even if the images do match, the value of the determinant **D** is never zero because all of the image measurements contain random measurement error. The determinant actually takes on quite large values because of the large order of the matrix (nine). Introducing normally distributed error into the measurements results in values of the determinant which are *not* normally distributed. This precludes comparing the difference between the value of the determinant and its expected value (zero) to its standard deviation to determine the probability of a match.

An alternative test is to determine whether one can "move" any one of the nine image points around within some specified radius (r) on the photograph being tested in order to satisfy the condition equation. This is achieved in the following manner. Let $(x^\circ, y^\circ)_{2k}$ represent the measured coordinates of point k in photograph 2 and $(x', y')_{2k}$ represent the unknown coordinates which satisfy Equation 19. Ignoring the subscripts for a moment, one can write

$$A^\circ x' + B^\circ y' + C^\circ = 0 \tag{23}$$

and

$$A^\circ x^\circ + B^\circ y^\circ + C^\circ = \mathbf{D}^\circ \tag{24}$$

where $\mathbf{D}^\circ$ is the value of the determinant computed from the measured coordinates. Subtracting the two equations results in

$$A^\circ(x' - x^\circ) + B^\circ(y' - y^\circ) = -\mathbf{D}^\circ. \tag{25}$$

Define r as the unknown distance to move the point to satisfy the relationship. Then

$$x' - x^\circ = r\cos\theta \tag{26}$$
$$y' - y^\circ = r\sin\theta \tag{27}$$

where the angle θ which minimizes r is to be determined. Substituting Equations 26 and 27 into Equation 25 and solving for r yields

$$r = -\mathbf{D}^\circ/(A^\circ\cos\theta + B^\circ\sin\theta). \tag{28}$$

The angle θ which minimizes r is found by equating the derivative of r with respect to θ to zero and solving for θ. Substituting the result into Equation 28 produces

$$r_k = \pm\mathbf{D}^\circ(A^{\circ 2} + B^{\circ 2})^{-\frac{1}{2}} \tag{29}$$

which is the distance of the point k from the epipolar line. Therefore, if we move the point $(x^\circ, y^\circ)_{2k}$ the distance r_k, the relationship is satisfied.

Unfortunately, r_k is not normally distributed either. One now has the choice of abandoning the image matching altogether or adopting a less rigorous criterion. Fortunately, the value of r_k has a physical meaning and allows one to construct a heuristic approach. The accuracy of this approach is detailed in the simulation results presented later.

Let σ_x and σ_y represent the standard deviations of the image measurement and point transfer error in x° and y°, respectively. The error is normally distributed with mean zero. One usually assumes that the standard deviation on each axis is the same, so $\sigma_x = \sigma_y = \sigma$. Using this assumption, and the assumption that the equation of the epipolar line is error free (one of the nine "relative orientations" achieves its expected true value), it is easily shown using Equations 24 and 29 that the standard deviation of r_k is given by $\sigma_r = \sigma$. One would then expect the value of r_k to be less than three times σ_r, or

$$\frac{r_k}{\sigma_r} < 3. \tag{30}$$

This will serve as our test. It should be pointed out here that, because r_k is a physical quantity and has an intuitive meaning as well, one can devise other tests.

An algorithm for determining whether the two images match can now be stated:

(a) given the image measurements (x_{ij}, y_{ij}) of nine points in the two photographs form the row vectors $\mathbf{U}_j$ given by Equation 13 and the matrix **U** given by Equation 15;
(b) compute the determinant $\mathbf{D}^\circ = \det \mathbf{U}$ and the nine cofactors c_{k1} of **U**;
(c) for each point k, $k=1, 2, \ldots, 9$
   (1) compute A° and B° using Equations 20 and 21,
   (2) compute r_k using Equation 29,
   (3) if $r_k/\sigma_r < 3$ the images match;
(d) if $r_k/\sigma_r > 3$ for all k, the images do not match.

SIMULATION RESULTS

Consider the two sister ships U.K. Ark Royal and U.K. Eagle depicted in Figure 1 (Jane's, 1970). The ships are identical except for the placement of the masts, minor superstructure differences, and the angle of the flight decks. Nine points are considered and numbered in the figures. Table 1 gives the object space coordinates of each point as scaled from the drawings. The image coordinates used in the simulations are computed from these points.

We will consider two cases:

- two photographs taken of the same ship (left photograph, Ark Royal, and right photograph, Ark Royal), and

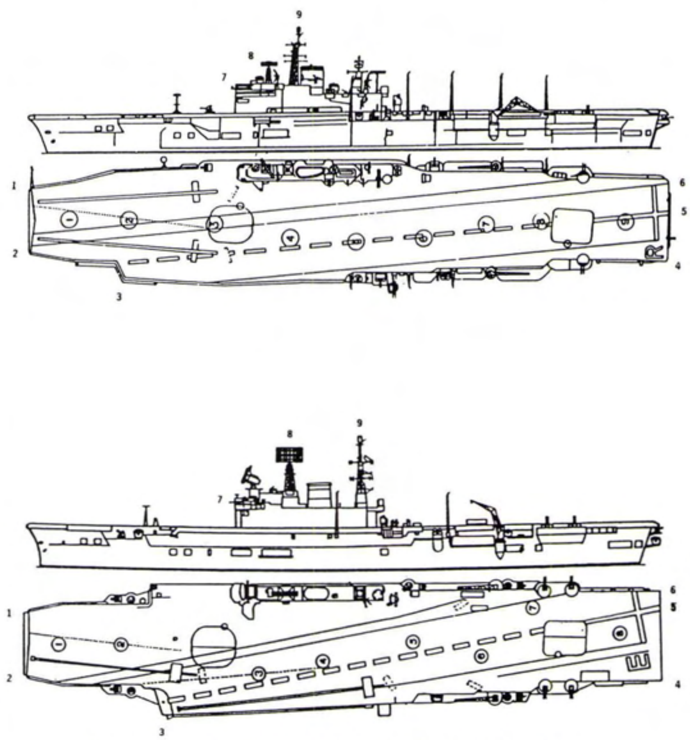

FIG. 1. U.K. Ark Royal (top) and U.K. Eagle (bottom).

TABLE 1. OBJECT SPACE COORDINATES OF SHIPS

U.K. Ark Royal Point Number (j)	X_j (metres)	Y_j (metres)	Z_j (metres)
1	0.0	48.0	13.0
2	0.0	24.0	13.0
3	32.0	16.0	13.0
4	220.0	14.0	13.0
5	220.0	39.0	13.0
6	220.0	50.0	13.0
7	72.0	50.0	24.0
8	84.0	55.0	29.0
9	93.0	55.0	43.0
U.K. Eagle Point Number (j)	X_j (metres)	Y_j (metres)	Z_j (metres)
1	0.0	48.0	13.0
2	0.0	24.0	13.0
3	48.0	11.0	13.0
4	220.0	14.0	13.0
5	220.0	45.0	13.0
6	220.0	50.0	13.0
7	72.0	50.0	24.0
8	92.0	55.0	39.0
9	116.0	55.0	47.0

TABLE 2. PHOTOGRAPH IMAGE COORDINATES OF MATCHING IMAGES

Left Photograph Image Coordinates (Arc Royal) Point Number (j)	x_{1j} (mm)	y_{1j} (mm)
1	−3.843	1.166
2	−1.925	0.884
3	−0.819	2.865
4	1.862	14.254
5	−0.018	14.506
6	−0.839	14.615
7	−2.936	6.305
8	−3.156	7.355
9	−3.044	8.601
Right Photograph Image Coordinates (Ark Royal) Point Number (j)	x_{2j} (mm)	y_{2j} (mm)
1	7.850	2.399
2	3.919	2.073
3	3.567	−1.108
4	9.459	−20.427
5	13.918	−19.929
6	15.937	−19.868
7	10.552	−2.604
8	11.898	−2.937
9	12.371	−2.078

- two photographs taken of different ships (left photograph, Ark Royal, and right photograph, Eagle).

In each case we will try to determine if the two photograph images match.

The left photograph in all cases has a scale of 1:12,000 and is assumed to be taken with a reconnaissance camera having a 6-inch focal length. The right photograph has a scale of 1:6,000 and is assumed to be taken with a camera having a 12-inch focal length. Each camera is located one nautical mile from the ship.

The simulated image coordinates in the left photograph contain normally distributed random mensuration error of 3 micrometres. The image coordinates in the right photograph contain the same error in addition to a normally distributed point transfer error of 82 micrometres (which amounts to 1.5 metres or 5 feet three sigma in object space). The point transfer error results from the inability to precisely locate a feature in the right photograph which is measured in the left photograph.

A number of simulation runs were made for each case by changing the perspectives of the photographs. The case having the largest perspective difference is presented here. A summary of all of the simulations is presented later.

For the two cases presented here, the left photograph of the ship has the camera depressed 55 degrees from the horizon (the aircraft flying high) while the right camera is depressed 35 degrees from the horizon (the aircraft flying low). The convergence angle (perspective disparity) between the two lines of sight is 88.4 degrees.

Table 2 presents the photograph image coordinates of matching images where both photographs are of the Ark Royal. Using the image matching algorithm, a minimum value of r_k of 0.035 mm was found for point $k=8$. This results in a ratio r_k/σ_r of 0.4 which, being less than three, correctly classifies the images as a match.

Table 3 presents the photograph image coordinates of non-matching images where the left photograph is of the Ark Royal and the right photograph is of the Eagle. In this case a minimum value of r_k of 0.293 mm. was found for point $k=8$. This results in a ratio r_k/σ_r of 3.6 which indicates the images do not match.

A total of 84 simulation runs were made for each of the two cases by varying the depression angles of the cameras to the horizon from 85 to 35 degrees and changing the direction of the cameras from abeam the ship by plus or minus 80 degrees. This resulted in a range of convergence angles between the lines of sight of from 13 to 88 degrees. These simulations resulted in 96 percent of the classifications being correct when both photographs were of the Ark Royal (4 percent false negative classifications) and 87 percent of the classifications being correct when one photograph was of the Ark Royal and the other of the Eagle (13 percent false positive classifications). The overall rate of success was 91 percent.

TABLE 3. PHOTOGRAPH IMAGE COORDINATES OF NON-MATCHING IMAGES

Left Photograph Image Coordinates (Eagle) Point Number (j)	x_{1j} (mm)	y_{1j} (mm)
1	−3.844	1.166
2	−1.920	0.886
3	−0.816	2.864
4	1.863	14.253
5	−0.023	14.511
6	−0.840	14.615
7	−2.929	6.305
8	−3.154	7.358
9	−3.047	8.602
Right Photograph Image Coordinates (Eagle) Point Number (j)	x_{2j} (mm)	y_{2j} (mm)
1	7.824	2.388
2	4.027	2.179
3	3.317	−2.741
4	9.443	−20.564
5	14.956	−19.897
6	15.980	−19.817
7	10.509	−2.657
8	12.178	−2.475
9	12.936	−3.722

DEGENERATE CASES

The image matching algorithm presented here is based on the premise that the system of equations represented by Equation 15 is homogeneous. If any relative orientation of the two cameras causes an element a_{pq} of the matrix **A** to become zero, then the algorithm fails. The cases for which this is true are easily seen by examining Equation 11 in which **A** is the product of the skew symmetric matrix **S** and the orthogonal rotation matrix $\mathbf{T}_2$. From this it is easily shown that, if any two of the three angles defining the relative orientation of the two cameras are zero, then one of the elements a_{pq} is also zero. This precludes the algorithm being applied to conventional vertical photography.

Another consideration is the distribution of object space points. If all of the object points are collinear, then the image points must also be collinear. In the case the y_{ij} terms in Equation 13 disappear. An additional constraint which is not at all obvious is that no more than seven of the nine object points may lie in a plane.

CONCLUSIONS

Although the conventional method of relative orientation is the best technique for matching images, it has serious difficulties with large perspective differences in the two photographs because it requires *a priori* estimates of the orientation angles defining the relative orientation of the cameras. It also requires knowledge of the cameras interior orientation. The relative orientation solution is also unstable when the image ray bundles are narrow.

The algorithm presented here for matching images in two photographs is potentially useful when there is a large, unknown disparity in the perspectives of the two photographs and the points are well distributed in three dimensions in object space. It is also useful in that it is independent of the principal point coordinates and focal lengths of the cameras and may be used for non-metric photography.

REFERENCES

Barrett, E. B., P. M. Payton, M. H. Brill, and N. N. Haag, 1990. Invariants Under Image Perspective Transformations: Theory and Examples, *International Journal of Imaging Systems and Technology*, Vol. 2, pp. 296–314.

Barrett, E. B., P. M. Payton, N. N. Haag, and M. H. Brill, 1991. General Methods for Determining Projective Invariants in Imagery, *Computer Vision, Graphics, and Image Processing*, Vol. 53, No. 1, pp. 46–65.

Haag, N. N., M. H. Brill, and E. B. Barrett, 1991. Invariant Relationships in Side-Looking Synthetic Aperture Radar Imagery, *Photogrammetric Engineering & Remote Sensing*, Vol, 57, No. 7, pp. 927–931.

Hinsken, L., 1988. A Singularity Free Algorithm for Spatial Orientation of Bundles, *International Society of Photogrammetry and Remote Sensing Comm. V, Kyoto*, pp. 262–272.

Jane's Fighting Ships 1969–70. McGraw-Hill Book Co., New York, New York.

Longuet-Higgins, H. C., 1981. A Computer Algorithm for Reconstructing a Scene from Two Projections, *Nature*, Vol. 293, No. 5828, pp. 133–135.

(Received 27 December 1991; Revised 5 May 1992; Accepted 9 April 1992)