TWO ALGORITHMS FOR RECONSTRUCTING SHAPES

L. L. Kontsevich, M. L. Kontsevich, and A. Kh. Shen

Avtometriya, No. 5, pp. 72-77, 1987

UDC 528.722.024

The problem of reconstructing the shape of a three-dimensional object from several plane projections is analyzed. A general approach is pointed out which leads to recurrence algorithms for solving various versions of this problem. This approach is illustrated with two examples: 1) reconstructing the shape of an object from an arbitrary set of orthogonal projections which are known within a scale factor; 2) reconstructing the shape of an object from two central projections with known parameters of the internal orientation.

INTRODUCTION

In this paper we wish to examine some mathematical questions which arise in the course of solving problems involving the interpretation of depth on the basis of the parallax of motion and binocular parallax.

We start by formulating a general problem. We assume that a finite set of points $A=\{a_i\}$ has been singled out in a three-dimensional Euclidean space S. This set (which contains n points) is called the "object." We assume that m projection operators p_j act on S in two-dimensional Euclidean spaces F_j. In each F_j, a point $a_i \in A$ corresponds to a point $a_{ji} = p_j(a_i)$.

Let us assume that we know neither $A \subset S$ nor p_j. We know only the a_{ji} and their correspondence (i.e., we know which points are different projections of a common point). We also know a class of possible mappings of p_i. We are to reconstruct A.

We will describe the solutions of the following problems of this type:

1) reconstructing the shape of an object from three or more orthogonal projections;

2) reconstructing the shape of an object from two central projections in the case in which the base of the perpendicular drawn from the center of the projection to the projection plane and its length are known for each projection.

We will not discuss questions related to distinguishing a figure from a background or establishing a correspondence between point in different projections.

The first of the problems listed above corresponds to the observation of a rotating object whose linear dimensions are much smaller than the distance to it. In this case the object must contain at least four points of general position. A solution which is found is determined within isometric transformations (i.e., the distance between any pair of points can be reconstructed). In almost the same way we solve the problem of reconstructing the depth in the case in which the projections are known within a similarity (this situation corresponds to large displacements along the depth of a small and remote object). In this case the solution is determined within an affine transformation which preserves angles (i.e., we can reconstruct the ratio of distances between any two pairs of points).

The second of these problems corresponds to a binocular observation of an object whose linear dimensions are comparable to the distance to the object. For the method which we are proposing here for solving this problem, we must require that the object contain at least eight points (although five points of general position would, generally speaking, be sufficient for reconstructing a shape). The solution of this problem is again determined within an angle-preserving affine transformation. Furthermore, if we do not know the orientation of the projections, two different solutions may appear.

Solutions based on the traditional apparatus of projective geometry have been proposed in many places for problems of this class [1-3]. That approach, however, has several shortcomings:

1) the algorithms turn out to be very complicated;

2) the points of an object are nonequivalent in the calculations. Specifically, the order in which the points are processed must conform to certain conditions, otherwise, incorrect results may be found;

3) the algorithms which are found do not in principle make use of the integral information on all points. In order to improve the accuracy of the algorithms it is necessary to partition the sets of points of the object into subsets (which may be intersecting) and to solve in each resulting subset the problem of reconstructing the depth. The results are then averaged.

It is pleasant to note an exceptional case. Longuet-Higgins [4], in solving the second of these problems, proposed an algorithm which is free of the shortcomings just listed and which agrees with our own in certain regards. On the whole, however, that algorithm seems to have been found more by a bit of luck than by applying a universal approach to a particular problem.

The method which we are proposing here makes it possible to uniformly utilize the information on all points simultaneously in the reconstruction of object A (there is no upper bound on the number of these points). The effect is to substantially improve the accuracy of the solution of the metric-reconstruction problem on the basis of inexact information about a_{ji}. The fact that the problem can be linearized completely, i.e., that the solution of the problem can be reduced to standard operations on vectors and matricds, should also be counted as an important advantage of this method. Additional information about the object (the existence of a symmetry plane in the object, information about the absolute or relative orientation of a projection, images on projections of the centers of other projections ("core points"), etc.) can easily be employed with appropriate modifications of the proposed calculation methods.

In developing specific algorithms we strive to make use of recurrence procedures, which make it possible to add new points rapidly and to subsequently correct the parameters of the affine (projective) structure of the object. For the second problem, this requirement is met completely. The program reconstructs the projective structure of the object and the metric by making use of a memory volume which does not depend on the number of points being processed (the coordinates of the points are stored in a memory which does not pertain to the program).

The resulting algorithms have proved to be extremely fast. In the second, and more complicated problem, for example, the addition of a new point and the cycle of recalculations by the recurrence procedure requires about four hundred operations with real numbers.

Some of the problems have not been solved in the course of constructing solutions:

1) in problem 1, it has not been found possible to construct a recurrence method for adding new points and new projections;

2) in problem 2, we ran into serious difficulties in moving up to a larger number of projections (more than two) or moving down to a smaller number of points (less than eight).

We realize that not all of the readers who need to use these algorithms will be interested in their mathematical underpinnings. Consequently, for each problem we will go immediately from the general discussion to an explicit description of a possible algorithm.

All the algorithms have been tested on a computer, and all have proved to be quite stable with respect to errors in the localization of points on the projections during the reconstruction of an affine (projective) structure of an object. They have also proved to be highly accurate in reconstructing a Euclidean structure in a case in which these errors are sufficiently small. Both of the algorithms include intermediate checks, so they are stable with respect to errors of a "malfunction" type during the entry of the coordinates of points of the projections.

1. RECONSTRUCTION OF THE DEPTH OF AN OBJECT FROM SEVERAL PARALLEL PROJECTIONS

1.1. Reconstruction of an object within an affine transformation. We consider the mapping $p=p_1\oplus\ldots\oplus p_m:S\to F_1\oplus\ldots\oplus F_m$. If m > 2, in the case of a general position, p is injective, and S can be identified as an affine space with p(S). In using S below, we will have p(S) in mind.

The problem thus reduces to one of drawing a three-dimensional affine subspace through a finite set of points p(A). If this problem is to have a single-valued solution, the object must not be a plane figure.

1.2. Reconstruction of a Euclidean structure of an object. We recall that S and F_j have a Euclidean structure. We transform from affine spaces to linear spaces, choosing as the origin of coordinates one of the points of the set and its images. We consider the case in which the p_j are orthogonal projections. Imposing these conditions on p_j is equivalent to requiring that the conjugate images $p_j^*:F_j^*\to S^*$ be isometric injectives. Each p_j determines three independent linear conditions on the quadratic form Q*, which specifies a Euclidean structure in S* which is the conjugate of the scalar product in S. The matrix of the corresponding scalar product in S is the inverse of the matrix of the form Q* in the conjugate basis.

It is not difficult to see that in order to reconstruct a Euclidean structure of an object we need at least three projections (although a Euclidean structure of S is determined by six parameters, and two projections give us six equations for these parameters; the rank of the system of linear equations which is constructed would be five).

If the orthogonal projections p_j are given within a scaling factor, one could work in a corresponding way to reconstruct a scalar product within a factor. The only difference is that in this case each p_j determines two independent linear conditions on the quadratic form Q*. In this case, the form is reconstructed within a homothetic transformation, again on the basis of three or more projections.

1.3. Algorithm. 1. For points a_{ji} on the projections F_j we write coordinates (x_{ji}, y_{ji}) in a 2m × n matrix M:

$$M=\begin{pmatrix} x_{11}\ldots x_{1n}\\ y_{11}\ldots y_{1n}\\ \ldots \quad \ldots\\ x_{m1}\ldots x_{mn}\\ y_{m1}\ldots y_{mn}\end{pmatrix}$$

2. We find a three-dimensional affine subspace which is constructed (approximately) on the column vectors of this matrix. One way to do this is as follows: from all the columns of M we subtract the first (we perform a linearization p(S)). Using the standard procedure, we carry out the singular decomposition $M = O_1DO_2$, where O_1 and O_2 are orthogonal matrices, and D is diagonal [5]. We assume $D=\mathrm{diag}(\lambda_1, \lambda_2, \ldots)$, where $\lambda_1\geqslant\lambda_2\geqslant\ldots\geqslant 0$. If $\lambda_3\gg\lambda_4$ (i.e., if the matrix M is approximately a third-rank matrix), we assume that the system of projections is matched, and we set $M := O_1D'O_2$, where $D'=\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3, 0, \ldots, 0)$, and rank(M) = 3. We denote by B the 2m × 3 matrix which consists of the first three columns of O_1. The columns of B form an orthonormal basis (in the usual metric R^{2m}) in the image of the operator M.

3. We decompose the columns of M in accordance with the basis of B. The object is thereby injective in coordinate space R^3. If the basis of B is chosen as outline above, the three-dimensional vector associated with the point $a_i\in A$, consists of the first three components of column i of the matrix $D'O_2$. Our remaining task is to find the matrix of the scalar product in R^3.

4. We decompose the matrix B into m blocks $\{B_j\}$ of size 2 × 3. We find linear equations for the symmetric 3 × 3 matrix Q* from the relations $B_jQ^*B_j^T=\mathrm{diag}(1, 1)$ for the orthogonal projection, and $B_jQ^*B_j^T$ is equal to a matrix of the form $\mathrm{diag}(\lambda, \lambda)$ for the version with scaling. The resulting system of 3m (or, respectively, 2m) linear equations is then solved by any suitable method.

5. $Q = (Q^*)^{-1}$. Diagonalizing Q, we find a Euclidean model of the object.

2. RECONSTRUCTION OF AN OBJECT FROM TWO CENTRAL PROJECTIONS

2.1. **Reconstruction of an object within a projection transformation.** We denote by V an affine space. With it we associate the vector product $\tilde{V}$, in which V is identified with a hyperplane which does not pass through the origin. The projective space $P(\tilde{V})$ consists of points V and infinitely remote points. The mapping of the central projection, $p_j: S \to F_j$, can be raised to a linear mapping $\tilde{p}_j: \tilde{S} \to \tilde{F}_j$ within a factor. The points a_i in the space S and a_{ji} in the projection F_j are associated with one-dimensional spaces $\tilde{a}_i$ and $\tilde{a}_{ji}$ in $\tilde{S}$ and $\tilde{F}_j$, respectively. Here we have $\tilde{p}_j: \tilde{a}_i \to \tilde{a}_{ji}$.

Let us consider the mapping $\tilde{p} = \tilde{p}_1 \oplus \tilde{p}_2: \tilde{S} \to \tilde{F}_1 \oplus \tilde{F}_2$. If the centers of the projection do not coincide we have $\mathrm{Ker}(\tilde{p}) = \mathrm{Ker}(\tilde{p}_1) \cap \mathrm{Ker}(\tilde{p}_2) = \tilde{f}_1 \cap \tilde{f}_2 = 0$ (here $f_j \in S$ is the center of projection j). Consequently, $\tilde{S}$ can be identified as a linear space with $\tilde{p}(\tilde{S})$. The reconstruction of the projective structure of the object thus reduces to finding, in the six-dimensional space $\tilde{F}_1 \oplus \tilde{F}_2$, a four-dimensional subspace $\tilde{S}$ such that we have $\dim(\tilde{S} \cap (\tilde{a}_{1i} \oplus \tilde{a}_{2i})) \geqslant 1$ for arbitrary i. This condition is equivalent to the existence of a point a_i which projects onto a_{1i} and a_{2i}. We show below that these restrictions on $\tilde{S}$, with a sufficient number of points in A, determine a single-parameter family of possible solutions which are equivalent from the standpoint of projective structure. In the specific calculations, it is sufficient to construct simply one of the solutions.

Since the codimensionality of $\tilde{S}$ is two, it corresponds to a 2-form $\omega \in \Lambda^2((\tilde{F}_1 \oplus \tilde{F}_2)^*)$, defined within a scalar factor. We denote by T the component ω in the term $\tilde{F}_1^* \oplus \tilde{F}_2^*$ in the decomposition into a direct sum, $\Lambda^2((\tilde{F}_1 \oplus \tilde{F}_2)^*) = \Lambda^2(\tilde{F}_1^*) \oplus (\tilde{F}_1^* \otimes \tilde{F}_2^*) \oplus \Lambda^2(\tilde{F}_2^*)$. It can be shown that the tensor T determines the space S within the action of independent homothetic transformations in $\tilde{F}_1$ and $\tilde{F}_2$. The limitations on S can be rewritten in the form

$$\langle T, a_{1i} \wedge a_{2i} \rangle = \langle \omega, a_{1i} \wedge a_{2i} \rangle = 0.$$

Let us write the conditions on T explicitly. The subspace is determined by the two linear equations

$$(\alpha_1^T, \alpha_2^T)(v_1, v_2)^T = 0;$$

$$(\beta_1^T, \beta_2^T)(v_1, v_2)^T = 0.$$

Here $v_j \in \tilde{F}_j$; $\alpha_j^T, \beta_j^T \in \tilde{F}_j^*$. In matrix form, we would have $T = \beta_2 \alpha_1^T - \alpha_2 \beta_1^T$. The conditions on the subspace $\tilde{S}$ imply

$$\det \begin{pmatrix} \alpha_1^T a_{1i} & \alpha_2^T a_{2i} \\ \beta_1^T a_{1i} & \beta_2^T a_{2i} \end{pmatrix} = 0 = (\beta_2^T a_{2i})(\alpha_1^T a_{1i}) - (\alpha_2^T a_{2i})(\beta_1^T a_{1i}) = \mathrm{tr}((a_{2i}^T \beta_2 \alpha_1^T a_{1i}) -$$
$$- (a_{2i}^T \alpha_2 \beta_1^T a_{1i})) = \mathrm{tr}((\beta_2 \alpha_1^T - \alpha_2 \beta_1^T)(a_{1i} a_{2i}^T)) = \mathrm{tr}(T(a_{1i} a_{2i}^T)).$$

We thus find linear equations which determine the tensor T. Since rank(T) = 2, T can be written as the difference between two matrices of rank 1: $T = T_1 - T_2$. Decomposing T_1 and T_2 into a product of a column and a row, we find the coefficients of the equations which specify one of the possible subspaces of S (it was mentioned above that the solution of the problem is not single-valued).

2.2. **Reconstruction of the Euclidean structure of the object.** We introduce a nondegenerate scalar product in $\tilde{F}_j$ in the following way. Since F_j is determined by an optical system with a known focal length, $\tilde{F}_j$ can be identified canonically with a three-dimensional physical space in which the focal length would naturally be regarded as the unit of length.

We can show that a Euclidean structure (within a scalar factor) in S determines the quadratic form Q^* of signature (+++0) in $\tilde{S}^*$ (within a factor), and vice versa. Specifically, $\tilde{S}$ contains a linear subspace L which is parallel to S and which therefore has a nondegenerate quadratic form. We have thus defined a scalar product in $L^* = \tilde{S}^*/\mathrm{Ann}(L)$ and thus in $\tilde{S}^*$. Inversely, the quadratic form Q^* of this signature has a one-dimensional nucleus whose annulator is a hyperplane L in $\tilde{S}$. The metric in L can be extended to any hyperplane parallel to L and thus, within a factor, to the region $P(\tilde{S}) \setminus P(L)$ of the projective space.

As in the preceding problem, $\tilde{p}_j^*: \tilde{F}_j^* \to \tilde{S}^*$ are isometric injective functions within a factor.

The space $M = \tilde{p}_1^*(\tilde{F}_1^*) \cap \tilde{p}_2^*(\tilde{F}_2^*)$ is two-dimensional, and the two quadratic forms in it which appear from Euclidean structures in $\tilde{F}_1^*$ and $\tilde{F}_2^*$ must be proportional. We multiply one of the quadratic forms in $\tilde{F}_j^*$ by the corresponding coefficient to make their limitations on M compatible. In M we choose an orthonormal basis e_1^* and e_2^*. To it we add vectors e_3^* and e_4^* (from $\tilde{p}_1^*(\tilde{F}_1^*)$ and $\tilde{p}_2^*(\tilde{F}_2^*)$, respectively) to form orthonormal bases in these subspaces. In the basis $e_1^*, e_2^*, e_3^*, e_4^*$ the Gram matrix is

$$Q^* = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \lambda \\ 0 & 0 & \lambda & 1 \end{pmatrix},$$

where λ is an unknown number. From the condition $\det(Q^*) = 0$ we find $\lambda = \pm 1$. Choosing one of these values of λ, we find two solutions of the problem.

With a point from p(S) with the coordinates (x_1, x_2, x_3, x_4) in the basis e_k we associate a point of the space R^3 with the coordinates $x_1/(x_3 - \lambda x_4)$, $x_2/(x_3 - \lambda x_4)$, $(x_3 + \lambda x_4)/(x_3 - \lambda x_4)$. The mapping p(S) into R^3 which is constructed is an isometry.

2.3. Algorithm. 1. On projection F_j we introduce an orthogonal coordinate system. As the origin we choose the point in the projection plane which is closest to the center of projection. We choose the corresponding focal length as the unit of length in the coordinate system associated with projection j. We denote the coordinates of points a_j in the resulting coordinate system as x_{ji}, y_{ji}.

2. We find the matrix T from a system of linear equations of the type

$$\mathrm{tr}\left(T\begin{pmatrix} x_{1i}\,x_{2i} & x_{1i}\,y_{2i} & x_{1i} \\ y_{1i}\,x_{2i} & y_{1i}\,y_{2i} & y_{1i} \\ x_{2i} & y_{2i} & 1 \end{pmatrix}\right) = 0.$$

This is a homogeneous sytem of equations with nine unknowns. A nonvanishing solution is determined within a linear factor by eight points.

To solve the indefinite system of linear equations it is convenient to use a recurrence method based on a Greville pseudoinversion algorithm [6].

This pseudoinversion algorithm makes it possible to reconstruct T within about $n^{-1/2}$ (n is the number of points of the object) at a fixed error in the coordinates of the points on the projections. The matrix T is determined by the relative positions of the centers and the projection planes and determines it within a scaling factor. We can find T much more accurately than we can find the shape of the object itself, and doing so may prove to be a problem of independent interest; however, it goes beyond the scope of the present paper.

If the predeterminant of the matrix T which is found turns out to be zero (within a given error, of course), the calculations can be pursued. Otherwise, the projections are incompatible.

3. We decompose T into the difference between rank-1 matrices T_1 and T_2 in an arbitrary way. For example, we could use the standard procedure of singular decomposition. We write the rank-1 matrices T_1 and T_2 as products of a column and a row (in an arbitrary way): $T = T_1 - T_2 = \beta_2\alpha_1^T - \alpha_2\beta_1^T$.

We thus find columns of coefficients $\alpha_1, \alpha_2, \beta_1, \beta_2$. Since we have used a singular decomposition, we have $(\alpha_1, \beta_1) = (\alpha_2, \beta_2) = 0$ and $(\alpha_1, \alpha_1) = (\beta_1, \beta_1) = 1$. The condition of the compatibility of quadratic forms means $(\alpha_2, \alpha_2) = (\beta_2, \beta_2) = c$. This condition must be checked: if this condition does not hold, the problem does not have a solution.

4. The basis in the space $\tilde{S}^*$, thought of as the factor $R^6 \Big/ \left(* \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} + * \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \right)$, is

$$e_1^* = \begin{pmatrix}\alpha_1\\0\end{pmatrix},\ e_2^* = \begin{pmatrix}\beta_1\\0\end{pmatrix},\ e_3^* = \begin{pmatrix}\alpha_1\times\beta_1\\0\end{pmatrix},\ e_4^* = \begin{pmatrix}0\\\alpha_2\times\beta_2\end{pmatrix}\Big/c.$$

5. For each point a_i we construct its representative $v_i \in \mathcal{S} \subset R^6: v_i = c_1(x_{1i}, y_{1i}, 1, 0, 0, 0)^T + c_2(0, 0, 0, x_{2i}, y_{2i}, 1)^T$. We choose the coefficients c_1 and c_2 in such a way that we have $(\alpha_1^T, \alpha_2^T)v_i = 0$. For example, we could choose $c_1 = \alpha_2^T(x_{2i}, y_{2i}, 1)^T$ and $c_2 = -\alpha_1^T(x_{1i}, y_{1i}, 1)^T$.

6. For each point a_i we calculate four coordinates $x_k(a_i) = e_k^* v_i$. Using the formula given above, and choosing $\lambda = +1$ or $\lambda = -1$, we find two Euclidean models of the objects.

We wish to thank M. N. Vaintsvaig, P. P. Nikolaev, and A. M. Bonch-Osmolovskii, whose advice was of no small assistance in the appearance of this paper.

REFERENCES

1. S. Ullman, Principles of thePerception of Moving Objects [Russian translation], Radio i Svyaz, Moscow, 1983.
2. S. A. Smirnov, Stereo Views in Photogrammetry [in Russian], Nedra, Moscow, 1982.
3. A. N. Lobanov, Photogrammetry [in Russian], Nedra, Moscow, 1984.
4. H. C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," Nature, vol. 293, p. 133, 1981.
5. I. M. Gel'fand, Lectures on Linear Algebra [in Russian], Nauka, Moscow, 1966.
6. T. Kokhonen, Associative Memory [Russian translation], Mir, Moscow, 1980.

7 April 1987 Moscow

Kontsevich, L.L.; Kontsevich, M.L.; Shen, A.Kh.
[Two algorithms for reconstructing shapes].
Avtometriya, 1987 (no.5):72-7.
Translation: Optoelectronics, Instrumentation and Data Processing, 1987 (no.5):76-81.