

Three-Dimensional Reconstruction from Projections: A Review of Algorithms

RICHARD GORDON

Mathematical Research Branch, National Institute of Arthritis, Metabolism, and
Digestive Diseases, National Institutes of Health, Bethesda, Maryland

AND

GABOR T. HERMAN

Department of Computer Science, State University of New York at
Buffalo, Amherst, New York

I. Statement of the Problem

The problem of reconstructing three-dimensional objects from a set of two-dimensional projected images has arisen and been solved independently in fields ranging from medicine and electron microscopy to holographic interferometry. By using a source of radiation external to the object, we obtain a transmission picture or projection of the three-dimensional object onto a two-dimensional surface such as the film of an ordinary electron micrograph or x-ray. The reconstruction problem is: *Given a subset of all possible projections of an object, estimate its internal density distribution.*

All algorithms for reconstruction take as input the projection data, and all produce as output an *estimate* of the original structure based on the available data. The estimate varies from method to method. The relative performance of the various methods depends on the object and how the data are collected. It is therefore important that qualitative judgments be made only after a careful and exhaustive study.

In order to state the problem clearly, let us consider the particular case of projected images formed by the transmission of x-rays through a patient.

From the intensity at each point on the film or detector, the total or integrated density of the object along the *path* of the radiation between the source and that point can be estimated. In practical applications we can measure the intensity only over small but finite regions of the projection. Therefore we define a *ray* as a bundle of paths from the source to a small region of a given shape on the projection. We define a *ray sum* as the estimate of the total density of the object contained in the bundle of paths defining the ray.

Suppose we want to detect the presence of some abnormality (e.g., a tumor) in someone's brain. Since such an abnormality usually has a density different from healthy tissue, a density map of the brain would be of great help. In fact, a series of two-dimensional sections of this density distribution would be sufficient (provided they are spaced close enough not to miss the abnormality).

For each two-dimensional section, the data are collected as follows. An x-ray source emits a collimated pencil beam in the plane of interest in the direction of a detector (Fig. 1). From the output of the detector, we can estimate the total density of the part of the section of the object that is between a pair of parallel lines of known location. The section of interest can be enclosed in a square region outside which the density may be assumed to be zero. The region in which the parallel lines intersect the square is a *ray* (cross-hatched in Fig. 1). The experimentally obtained total density within the ray is its *ray sum*.

For the present example we assume that the data collection is such that we choose a number of directions, and for each direction we collect the ray sums for a series of nonoverlapping rays of equal width which between them cover the square. The total information for one particular direction is a *projection* (one-dimensional, for one section of the object).

In electron microscopy the same effect is obtained by tilting the stage about a single axis. Corresponding lines on each plate perpendicular to

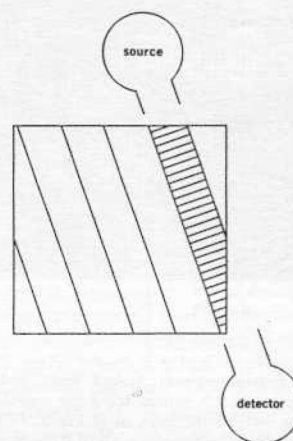


FIG. 1. All the rays of one projection at one angle are shown. The ray whose ray sum is being collected is cross-hatched. The object is assumed to be entirely within the square.

this axis are projections of the same plane. Thus each plane is reconstructed independently of other planes. Such planar reconstructions can be stacked to recover the complete three-dimensional structure.

It is also possible to make reconstructions from projections at spatial angles not lying in a plane. This is in fact a common procedure in tomography (a reconstruction method widely used in medicine). However, when reconstructions are to be made on computers, considerations of computing costs (both time and storage), as well as display ability, make three-dimensional reconstruction via a series of two-dimensional reconstructions more attractive. Unless otherwise stated, we shall assume that we are reconstructing a two-dimensional nonnegative density distribution or picture from its one-dimensional projections. (For the sake of precision, we define a *picture* as a square region of the plane together with a well-defined grayness at every point in the region. We assume that the grayness is measured by a nonnegative real number.)

Figure 2B illustrates a reconstruction of the test pattern in Fig. 2A. The projection data were collected mathematically in the manner of Fig. 1. Also, a reconstruction of a section of a canine heart is shown in Fig. 2C.

In radiology it is usually possible to obtain a large number of projections, but in other fields (electron microscopy, interferometric holography, and radio astronomy) both the number and the range of projections can be rather restricted. Even in radiology the number of projections may have to be small, either to reduce the x-ray dose to the patient, or because the object reconstructed is a rapidly moving one, such as a living heart.

The mathematical and computational difficulty of the reconstruction problem is increased by the fact that the projection data are noisy, or may even contain systematic errors, so that the ray sum is not exactly

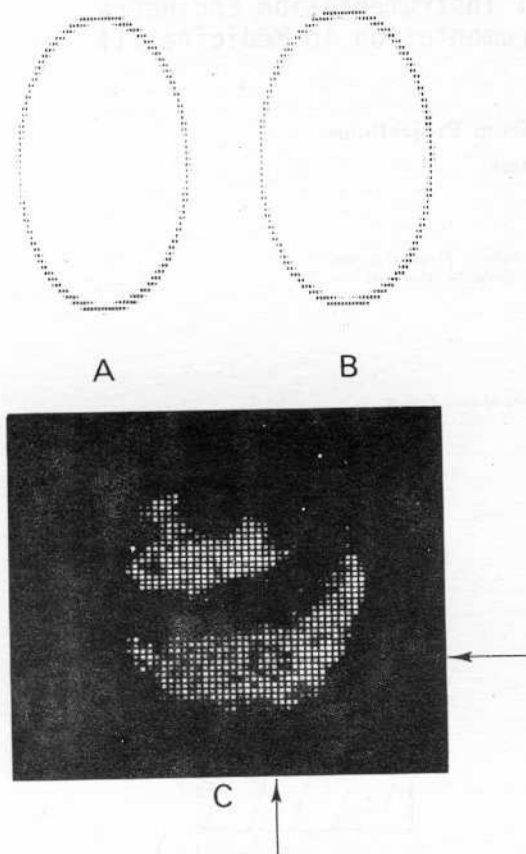


FIG. 2. (A) A test pattern representing skull, brain, and tumor on a 64×64 raster. (B) A reconstruction of this pattern using the convolution method with 25 evenly spaced projections with parallel rays, as in Fig. 1. (C) A section of a dog's heart on a 64×64 raster reconstructed using ART3 from 36 equally spaced slightly divergent x-ray projections from a point source. The ring indicated by the arrows is an inserted catheter.

the total density in the ray. For example, errors may be made in adjusting the directions of the projections and/or in positioning the rays.

We review the algorithms that have been proposed to solve the reconstruction problem, and indicate frequent research duplication which has occurred, limiting ourselves to information that has been published in regular journals and doctoral dissertations.

We have classified the known reconstruction algorithms into four categories:

1. *Summation*: The ray sums of the rays through each point are simply added to obtain an estimate of the density at the point.
2. *Use of Fourier transforms*: The projections are transformed into Fourier space to obtain some of the values of the Fourier transform of the whole picture. Other values are estimated by interpolation, and the reconstruction is obtained by taking the inverse Fourier transform.
3. *Analytic solution of the integral equations*: The relation between the picture and its projections is expressed by a set of integral equations which are then solved analytically. The picture elements are estimated based on the analytic solution.
4. *Series expansion approaches*: It is assumed that any picture we may be interested in can be sufficiently approximated by a linear combination of some predetermined basis pictures. The unknown coefficients in this linear combination are estimated from the equations obtained by expressing the projections of the unknown picture as a linear combination of the projections of the basis pictures.

II. Reconstruction Methods

For each class of algorithms we give: (1) a general intuitive description, (2) a precise mathematical description of a typical reconstruction method of the class, and (3) a brief description of other methods in the class.

A. SUMMATION

The simplest algorithm for reconstruction is to estimate the density at a point by adding all the ray sums of the rays through that point. We call this the summation method.

Tomography is a medical reconstruction technique dating back to French patents in the 1920s, which had some independent starts (see Kieffer, 1938). We will demonstrate that tomography is a summation method.

If we are interested in the density distribution in a cross section C of a patient (Fig. 3), we can obtain a fairly good estimate by the following tomographic method. We place a photographic plate P parallel to the cross section C on one side of the patient, and an x-ray source on the other side. By moving the x-ray source at a fixed speed parallel to C in one direction, and moving P at an appropriate speed in the opposite direction, we can insure that a point in C always projects onto the same point in P, but a point in the patient above or below C is projected onto different points in P. (See Fig. 3.) Thus on the photographic plate the density distribution at the section C will stand out, while the rest of the body will be blurred out.

The density at point A in the body is estimated by summing up (integrating) the total density along the path from X_i to A, as time t varies. Note that A, is always at the same point on the moving photographic plate P, and that A is the only point that two paths from X_i to A, have in common at different times t . Thus all forms of tomography involving such coordinated movements of x-ray source and film are precisely three-dimensional versions of the summation method.

Tomographic devices (Edholm, 1960; Reichmann, 1972) are designed to yield one plane in focus, or possibly a whole stack of planes photographed on different plates. More recent three-dimensional tomographic devices store the individual projections at a finite number of angles, so that by appropriate registration any plane can later be reconstructed ("longitudinal section scanning," Kuhl and Edwards, 1963; Dümmling, 1969; Freedman, 1970, 1972; "photolaminagraphy," Miller *et al.*, 1971). In some cases the projection data are actually optically projected back into a volume in space, which can be sliced in any desired plane by inserting a viewing screen ("tomosynthesis," Garrison *et al.*, 1969; Grant, 1972; Harper, 1968). Hart's (1968) "polytropic montage" in electron microscopy is equivalent to circular tomography.

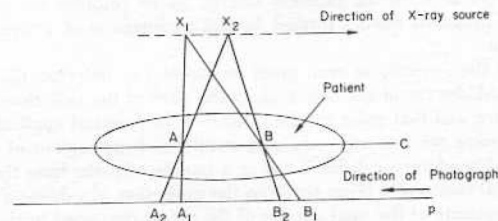


FIG. 3. Linear tomography. C, Patient cross section; A and B, two points in the cross section C; X_1 and X_2 , positions of the x-ray source at times t_1 and t_2 ; p, the photographic plate; A_1 and A_2 , positions of a fixed point on p at times t_1 and t_2 ; B_1 and B_2 , positions of another fixed point on p at times t_1 and t_2 .

In these techniques the quality of the output is often degraded by the presence of sharp density differences outside the cross section of interest (Reichmann, 1972). To avoid this, Kuhl and Edwards (1963) introduced transverse section scanning in which only densities in the plane of interest contribute to the final result (Fig. 4). The detector views the "... structure from many directions, but all views are made in a single transverse plane. The partial images which result from the various scans can be superimposed in the recording system so that their spatial relationships are preserved. To accomplish this, the detector axis is represented on an oscilloscope screen during scanning by a slender line having similar direction, motion, and angle of inclination. The brightness of this line is modulated by the [ray] sum from the detector. During scanning the oscilloscope patterns are integrated on photographic film. Thus, image fragments corresponding to any single structure will coincide on the recording. As a result, an image of the distribution of [density] in the cross section examined will be displayed finally on the film."

The principle underlying this technique is the same as in ordinary tomography: the density at each point is estimated by the sum of the total densities (ray sums) of all the rays through the point.

This planar version of the summation method was independently discovered in electron microscopy by Vainshtein ("method of projecting

