〔非公開〕

TR－C－0020

3－D OBJECT REPRESENTATION IN 3－D VISUAL COMMUNICATION SYSTEMS

李　大力
DANIEL LEE

1988．11．15．

ATR通信システム研究所

3-D OBJECT REPRESENTATION IN 3-D VISUAL COMMUNICATION SYSTEMS

Version 1.0: 3/15/88

Daniel Lee

Artificial Intelligence Department

ATR Communication Systems Research Laboratories

Twin 21 Bldg. MID Tower

2-1-61 Shiromi Higashi-ku

Osaka 540, Japan

e-mail: lee@atr-sw.atr.junet

Abstract

In a conceptual visual communication system that makes use of three-dimensional display technology, applications such as 3-D information system, tele-shopping and 3-D distributed design will require the system to have the capabilities for manipulation, rendering, compositions and animations of objects from a database of 3-D objects captured by 3-D input methods.

An object representation scheme that facilitates a systematic construction of such a database is therefore needed. Such development involves methodologies from CAD systems, computer graphics and computer vision techniques. The scope of such a development and the starting point of the first problem, that of optimal triangulation for surface approximation, are considered in this memo.

I. Introduction

Progress in communications technologies has led to many new developments in visual communication systems. Today, communications of pictorial information are readily available in systems such as high definition television (HDTV), picture telephone and video-conferencing. A higher dimension in visual communication, one where maximum *display realism* can be achieved so that a *virtual presence effect* can be created for the receiver, becomes one of the most exciting challenges for current research.

A key theme of research in the Visual Communication System Project of the Artificial Intelligence Department is to consider the technologies that are necessary and problems that must be overcome in order to build systems that strive to achieve display realism, [Akiyama, et al, 1988]. Such systems will make a great impact to current telecommunication systems enabling new modes of communications (e.g., virtual experience and remote presence) and creating new commercial applications (e.g., tele-shopping, 3-D information system and 3-D distributed design).

To realize such systems, the following technologies are necessary:

(1) Technologies on intelligent, user-friendly human interface for human oriented communications.

(2) Technologies on 3-D image database (i.e., database of 3-D objects).

(3) Technologies on stereoscopic display systems (systems that are adapted to the characteristics of the human vision system).

The problem of 3-D object representation in a 3-D image database is considered in this memo.

II. 3-D Objects

From the scenarios of potential 3-D applications that have been considered (e.g., 3-D information system, tele-shopping, architectural and landscape navigation, etc.), we can extract a set of common characteristics and constraints on the paradigm of the 3-D communication problem that can help us to put a focus on the scope of the problem.

Problem Statement

The main problem can be addressed in two parts:

(1) Object representation for 3-D image database.

(How do we represent 3-D shapes and models? What kinds of representations and computational paradigms, i.e., the data structure in memory and algorithms that operate on it, are useful for the applications?)

(2) Object reconstruction for 3-D visual communication.

(How well can we recreate the object from its representation? How well can we render the object to give the most desirable visual effect?)

We will concentrate on the representation of real objects, but before defining the class of 3-D objects we first consider the general characteristics of the representation that we seek in our problem paradigm.

II.1 Characteristics of Representation Schemes

(i) **Manipulation.** The representation must be good for spatial manipulation, e.g., a homogeneous representation that can be manipulated by the basic geometric operations of computer graphics algorithms interactively would be preferred.

(ii) **Rendering.** The representation must be usable by computer graphics algorithms to render realistic images. Therefore a surface representation is preferred.

(iii) **Composition.** Provisions for the efficient composition of 3-D objects (e.g., merging of different objects) to form new scenes, with stereoscopic viewing capability, in particular, must be considered in the representation.

(iv) **Animation.** Animation of the scenes with respect to stereoscopic presentation, e.g., provisions of parallax computation, should also be considered in the representation.

Remarks: As can be seen from these characteristics, the methodologies for the development of the problem include techniques from computer graphics (for rendering), computer vision (for 3-D input and representation) and CAD systems (for 3-D modeling and manipulation).

II.2 Object Classification

Objects in a 3-D image database are classified into two classes: *synthetic objects* and *real objects.*

Synthetic objects are the general class of images synthesized by computer graphics techniques. They are also obtained readily through the use of solid-modelers in modern CAD/CAM systems. Advances in the broad field of computer graphics have made the goal of synthesizing realistic images more feasible in terms of algorithms and hardware systems. The class of synthetic objects, however, is still limited. For achieving the objectives of our 3-D visual communication paradigms, we will concentrate on the problems dealing with the class of real objects.

Real objects are the physical real-world objects whose forms and shapes we seek to communicate in our 3-D communication paradigm. With respect to stereoscopic viewing, we can further divide the class into objects for *near-field view* and *far-field view*.

Far-field view objects, as the name implies, are typically the background scenes (so-called natural scenes) or objects that are imaged at a distance. Since the depth perception ability of the human vision system decreases substantially at distance greater than about 10 meters and, moreover, the need for manipulation and animation of background scenes are not imminent, we will be less concerned with this class of objects. Also from the recent advances in 3-D imaging technologies (e.g., 3-D video) we can assume that stereoscopic pictorial database of far-field objects can be readily constructed using the available equipments.

Near-field objects are the physical objects that are imaged and viewed at a viewing condition closely matched to the those of current display technologies, i.e., frame and object magnification that fit within typical CRT display systems. As the next section will show, we will be concerned with the acquisition, representation and reconstruction of these objects.

Among this class of objects, we can further decompose into *simple objects* and *complex objects*. Simple objects are fundamental geometric shapes such as cubes, tetrahedrals, spheres and cones that form a set of building block for the complex objects. The rules of the composition of complex objects can be based on *hierarchical representation* or *syntactic representations*, further considerations of which will be given in another place.

In the next section we will outline some of the problems concerning the representation of the class of near-field, real objects for the 3-D communication paradigm.

III. Problems in 3-D Object Representation and Reconstruction

We begin by mentioning the shape acquisition methods that have been developed in the project and consider a problem arising in 3-D range image processing, the surface representations and reconstruction using triangular patches.

III.1 3-D Input Techniques

Image input techniques generally falls into two classes: *intensity-image* and *range image.* At each point in an intensity image, the brightness or luminance value is given by a combination of the surface reflectance and geometry, and scene illumination. The group of methods that deduce the shape of an object using intensity image is referred to as the shape-from-(xxx) techniques in the computer vision literature.

In range-image, the value at each pixel is given by the distance from the sensor to the object surface, i.e., depth information. Therefore the shape information of the object is obtained directly. In the current project, two range-image based shape acquisition methods have been developed: moire topography [Koezuka, et al 1988] and slit-ray projection [Nishino, et al 1988].

High resolution depth maps of objects can be obtained using these methods, but as active triangulation techniques, their utilities are limited by the shadowing (or "missing parts") problem due to the necessary separation of the light source and detectors. On the other hand, the time of flight techniques, although not subject to this problem, require more complicated equipment for the same resolution and offer different kinds of problems. The two range-image input methods are being refined and will be our choice of input methods.

III-2 Surface Representations

We first limit the class of 3-D objects to those whose surfaces can be specified by a small number of parameters. Simple representations in terms of *linear primitives*, such as points, lines, and planes will be used first (*polyhedral approximation*). Extensions to the next level in terms of curved surface patches (*sculptured surfaces*) will be made later.

Parametrized surfaces, e.g., cones and cylinders, are useful in modeling complex objects. from the CAD/CAM discipline it is believed that over 85 percent of manufactured parts, e.g., a telephone hand-set, can be modeled approximately well using only planes, cylinders and spheres. Although our objectives are not the same as those of object recognition or solid-modeling, we would like to seek reliable surface parameter extraction that can provide powerful information for the communication paradigm.

III.3 Surface Triangulation Problem

The starting point of constructing a polyhedral approximation of the object shape is the problem of *optimal triangulation* to form the surface of a 3-D object from a set of range-image data. We note that approximating the surface of the object by triangular patches is only one of many methods, although the most intuitive, for surface representation. An alternative is a so-called top-down approach (often called "subdivision method") where the surface is first fitted by a rough approximating surface which is then progressively refined in successive steps in regions where the data is poorly approximated. For example, rectangular patches are used for bicubic curved surface patches [Schmitt, et al, 1986].

In the triangulation problem, we seek the following:

(1) to define the conditions for optimality in the surface approximation by triangular patches, and

(2) to develop efficient algorithms for their construction.

Under the general condition where the range image data are arbitrarily distributed, one form of triangulation known as the *Delaunay Triangulation* has been considered as optimal from computational geometry consideration. It has several geometrically optimal properties like no long or thin triangles (the so-called "equiangular property"). Much work has been done on this problem including its optimal construction in 2-D but not in 3-D (see [Preparata and Shamos, 1985]).

In the 3-D input methods of moire topography and slit-ray projection, the range data are confined to lie on parallel planar contours. This additional information reduces the complexity of the general problem and leads to the formulation of the problem in a graph theoretic approach [Fuchs, et al, 1977]. A rigorous solution to this problem requires full search of the graph and heuristic techniques have been developed for particular instances of the problem, e.g., [Nishino, et al 1988].

Our first objective is to develop algorithms for efficient extraction of planar regions, i.e., *planar region segmentation*, from the range data following the graph theoretic approach (both rigorous and heuristic approaches will be pursued). We will pursue a subdivision approach to obtain the optimal triangulation. The next objective will be to extend the algorithms to obtain parametrizations of the remaining regions which are modeled as quadric surfaces. The final goal is to obtain an efficient surface representation of 3-D objects that can be used by graphics algorithms to render realistic images.

III.4 Next Step

After planar region segmentation, the next step is to pursue *quadric-surface-region segmentation*. As principal references, we will pursue the methods along the lines of computer vision research from the following groups: INRIA [Faugera and Hebert, 1986], Michigan [Besl and Jain, 1985] and IBM [Bolle, et al, 1987]. We also note that *surface primal sketch* approach from MIT [Brady, et al 1985] will also be pursued.

REFERENCES

[Akiyama, et al, 1988] "Consideration on Three-Dimensional Visual Communication Systems", by K. Akiyama, T. Koezuka and Y. Kobayashi.

[Besl and Jain, 1985] "Three-Dimensional Object Recognition," by P. Besl and R. Jain, Computing Surveys, Vol. 17, No. 1, March 1985, pp. 75-145.

[Bolle, et al, 1987] "Primitive Shape Extraction from Depth Maps," by R. Boole, D. Sabbah and R. Kjeldsen, IBM Technical Report RC-12392.

[Brady, et al 1985] "Describing Surfaces", by M. Brady, J. Ponce, A. Yuille and H. Asada, Computer Vision, Graphics, and Image Processing, Vol. 32, 1985, pp. 1-28.

[Faugeras and Hebert, 1986] "The Representation, Recognition, and Locating of 3-D Objects", by O. Faugeras and M. Herbert, Int. J. Robotics Research, Vol. 5, No. 3, Fall 1986, pp. 27-52.

[Fuchs, et al, 1977] "Optimal Surface Reconstruction from Planar Contours," by H. Fuchs, Z. Kedem and S. Uselton, Comm. ACM, Vol. 20, No. 10, Oct. 1977, pp. 693-702.

[Koezuka, et al, 1988] "A Consideration on Automatic Acquisition of the Shape of A Three Dimensional Object Using Moire Topography", by T. Koezuka, J.-B. Thomine, K. Akiyama and Y. Kobayashi.

[Nishino, et al 1988] "Consideration on Automatic Acquisition and Reconstruction of an Object Shape", by H. Nishino, K. Akiyama and Y. Kobayashi.

[Preparata and Shamos, 1985] *Computational Geometry*, by F. Preparata and M. Shamos, Springer-Verlag, New York, 1985.

[Schmitt, et al 1986] "An Adaptive Subdivision Method for Surface-Fitting from Samples Data," Computer Graphics, Vol. 20, No. 4, Aug. 1986, pp. 179-188.