RECONSTRUCTION OF THE 3D OBJECT MODEL: A REVIEW

S.C. Ng¹, N. Ismail²
¹School of Information Technology, SEGi University College, Kota Damansara, Malaysia; ²Faculty of Engineering, Universiti Putra Malaysia, Selangor, Malaysia.

ABSTRACT

The three-dimensional (3D) reconstruction model of a real object is useful in many applications ranging from medical imaging, product design, parts inspection, reverse engineering to rapid prototyping. In the medical field, imaging techniques such as computed tomography (CT), magnetic resonance imaging (MRI) and single positron emission tomography (SPECT) are applied to create 3D images from emanation measurements for disease diagnoses and organ study. On the other hand, reconstruction is widely utilized to redesign manufacturing parts in order to save production cost and time. A typical reconstruction application consists of three major steps, which are data acquisition, registration and integration as well as surface fitting. Based on the nature of data captured, the 3D reconstruction model can be categorized into two groups: methods working on (i) two-dimensional (2D) images and (ii) sets of 3D points. This paper reviews different methods of 3D object model reconstruction and techniques subjected to each method.

1.0 Introduction

Reconstruction is a process of rebuilding by new construction the exact form and detail of a vanished structure, object or a part thereof as it appeared in as a specific period of time. In early time, it has been widely applied in medical imaging. CT, MRI, SPECT, 3D Ultrasound (US) and positron emission tomography (PET) provide a wide range of 3D medical imaging modalities. In industry various models for the automatic reconstruction of objects are proposed [Ghassan et al., 2007; Ng et al., 2009]. The need for the 3D reconstruction model is motivated by several factors. The manual creation of the 3D model is time consuming and therefore expensive (Lim & Shamsuddin, 2004). Besides, the increasing availability of 3D scanning devices has responded to this need.

Generally, the reconstruction procedure encompasses four (4) steps [Lim & Shamsuddin, 2004] (Figure 1). Data acquisition is the process of sampling physical objects to be reconstructed. Sampling devices can be coordinated measuring machines (CMM) that generate sets of points, cameras that produce intensity images or a stereoscopic image, acoustic devices such as sonar that generate height maps and medical scanning devices such as CT and MRI scanner that generate volumetric information. The most common sampling device today is the laser scanner that generates range data. Registration refers to the process where the range data are aligned to transform all the measurements into a common Euclidean space [Montani et al., 2004]. In the integration stage, all the data are merged into a single surface model. The aim of surface fitting is to construct a concise and accurate approximation of the physical surface. The display stage denotes the visualization of the final model representation by using a rendering technique. The objective of this
paper is to review different 3D object model reconstruction methods and various techniques subjected to these methods.

Figure 1: 3D Object Model Reconstruction Procedure

The remainder of this paper is organized as follows. Section 2 describes the methods of reconstruction of a 3D object model from a set of 2D images. Two main issues of reconstruction, registration and integration, are discussed in this section. Section 3 describes the methods of reconstruction of a 3D object model from sets of 3D range data. In Section 4, some concluding remarks are drawn.

2.0 Reconstruction from Two-Dimensional Images

There are numerous techniques of reconstruction of a 3D object model from 2D images. Tomography is imaging by section. Examples of 3D image reconstruction by tomography are CT, MRI, PET and 3D US. CT, MRI and PET are imaging methods that utilize backprojection to reconstruct 3D images but they differ in the methods of acquiring data. CT [Radon, 1917] is a medical imaging method employing tomography in which digital geometry processing is used to generate a 3D image of the inside of an object from a large series of 2D x-ray images taken around a single rotation over the object. Physically, x-rays can traverse a cross-section of an object along straight lines, be attenuated and detected outside the object. During CT scanning, the cross-section is probed with x-rays from various directions. The attenuated signals are recorded and converted to projections of the linear attenuation coefficient distribution of the cross-section. These x-ray shadows are then taken to the Fourier transform of the cross-section and can be processed to reconstruct the 3D image. This method is known as backprojection (Figure 2). For filtered backprojection, Bracewell & Riddle, 1967 used for the fan beam data, each of the 1D views is convolved with a 1D filter kernel to create a set of filtered views. These filtered views were then backprojected to provide the reconstructed 3D image.

Figure 2: Illustration of Backprojection. (a) The x-ray source and the detector move in a series of steps to project an object’s image from multiple angles around a rotation over the object. (b) Reconstructed image of the inside of the object by slicing the 2D X-ray images.
MRI [Lauterbur, 1973], formerly referred to as Magnetic Resonance Tomography (MRT), is a method used to visualize the inside of living organisms and to detect the composition of geological structure. It is primarily used to demonstrate pathological or other physiological alteration of living tissue and is commonly used in medical imaging. Unlike CT that uses ionizing radiation, x-rays, to acquire images, MRI, on the other hand, uses radio frequency (RF) signals to acquire images. The shift in the RF resulting from magnetic gradients localizes the excited tissue from which the signals originate. The signals attributed to individual volume elements of tissue can be reconstructed into an image by using the technique of Fourier transform to produce a frequency spectrum as an intensity function. The frequency domain spectrum is then backprojected to produce the 3D image (Figure 3). Another imaging method, such as PET, is a radiotracer imaging technique, in which tracer compounds labeled with positron-emitting radionuclide are injected into the subject of the study. The subject of the study is placed within the filed of view (FOV) of a number of detectors capable of registering gamma rays. These tracer compounds can then be used to track biochemical and physiological processes in vivo by an annihilation\(^1\) process of positrons resulting from the decay of radionuclide in the radiotracer. The FOV encompasses the entire volume enclosed by detector modules capable of measuring depth of interaction (DOI). To reconstruct 3D images, the filtered backprojection-based algorithm incorporates DOI to minimize interpolation in the irregular sampling data by defining lines of response between the measured interaction points [Virador \textit{et al.}, 2001]. A 3D mesh model is then created by segmenting the 3D images, known as voxel data, acquired from CT, MRI or PET (Figure 4).

\begin{figure}[h]
\centering
\begin{subfigure}{0.3\textwidth}
\centering
\includegraphics[width=\textwidth]{figure3a}
\caption{(a) When excited by an RF transmitter, the hydrogen atoms return to their lower state in a process called "Relaxation" and re-emit RF radiation at the Larmor Frequency. This signal is detected as a function of time and then is then converted to signal strength as a function of frequency by means of Fourier Transformation. (b) Linear positioning information along a number of different directions in a rotating gradient is combined to produce a 2D map of the proton densities. (c) The MRI Image.}
\end{subfigure}
\end{figure}

3D US imaging [Baba \textit{et al.}, 1989] provides an imaging technique that allows the diagnostician to view the anatomy in three-dimension. In comparison with CT and MRI, it provides tomographic images at higher rates (10-16 images per second) and the orientation of the images is flexible because they are not necessarily acquired as a stack of planes as in CT and MRI imaging. 3D US systems use two basic approaches to reconstruct 3D images. In the first approach, the conventional one-dimensional (1D) arrays are used to produce 2D images, which are later reconstructed into 3D images using knowledge of their relative orientation.

\(^1\) The process in which an electron and a positron collide and result in the creation of gamma ray photons or other particles.
positions. The second approach generates real-time 3D images directly from 2D arrays. Three solutions are proposed to acquire 2D images by 3D US system: (i) freehand acquisitions, (ii) mechanical localizers, and (iii) 3D probes. The reconstruction process of 3D US image has been implemented in two distinct ways (Figure 5). In the first, the series of 2D images are segmented to extract the desired features before the 3D image is reconstructed. The second approach uses the acquired series of 2D images to build a 3D voxel-based Cartesian volume (i.e., 3D grid) by placing each acquired 2D image in its correct location in the volume according to the Nyquist Sampling Theorem [Fenster & Downey, 1996].

While CT, MRI, PET and 3D US are radiological imaging modalities, photogrammetric reconstruction [Kolmogorov & Zabih, 2002] provides another means of reconstructing an object’s 3D shape by capturing several 2D stereoscopic images or pictures. So far, the main practice of photogrammetric is the reconstruction of a scene or outer body of objects by optical sensors. To capture the full geometry of an object, several images are taken from different positions by a camera on the principle of triangulation. Sometimes, several cameras are used to capture a scene. Reconstruction of an object’s model from 2D images needs the information of the relative motion and calibration of the camera and the corresponding image points. Recovering relative motion between consecutive images is accomplished by finding the corresponding image features between these images. Calibration is used to remove the projective skew, yielding a reconstruction equivalent to the original up to a global scale factor [Pollefeys & Gool, 2002]. By facilitating the
knowledge of the camera parameters, all pixels of an image are matched with pixels with neighboring images. By fusing the results of all the images together, a complete dense 3D surface model is obtained. Several volumetric algorithms are used to reconstruct surfaces and obtain dense correspondence [Martin & Aggarwal, 1983; Szeliski, 1993]. They work directly on discretized space based, voxel, on their image consistency and visibility.

3.0 Reconstruction from Three-Dimensional Points
Surface model reconstruction from 3D points has been the most prominent method for reconstructing 3D object model for several decades. This method obtains dense range data in the form of 3D point cloud by range scanning. In the data acquisition stage, the most common form of range sensing is optical triangulation [Curless, 1999]. A laser stripe triangulation scanner first spreads the laser beam into a sheet of light with a cylindrical lens. A CCD camera observes the reflected stripe from which a depth profile is computed. The object sweeps through the FOV and yields a range image. Other scanner configurations then rotate the object to obtain a cylindrical scan. A range image obtained is a collection of points with regular spacing. CMM provides an alternative to obtain range points. CMM are mechanical systems designed to allow for the evaluation of the dimensions of geometrically sophisticated work pieces. It is widely used in manufacturing component inspection and reverse engineering of complex parts and freedom surface. CMM comprises four main components, which are the machine, the measuring probe, the control or computing system and the measuring software. A sensor, which provides the connection between the object surface and a 3D length measuring system of the CMM is called the “probe”. A CMM consists of a mount, which can hold different probes moving in any of the three coordinate directions. During the coordinate measuring process, an object to be measured is normally probed point-by-point using a stylus with a spherical (most commonly) ruby ball tip [Wozniak & Dobosz, 2005]. At the moment of each probing contact, the XYZ coordinates of the ball tip are measured and stored in a computer memory. A 3D vision sensor, OPL-3D proposed by Sansoni, Carmignato & Savio, 2004 uses a CMM for the digitization of the object and the reconstruction of calibrated CAD model of known accuracy. CMM contact-measurement method is not suitable for modeling complex parts in reverse engineering owing to its inherent slow speed. Thus, several non-contact sensors have been developed such as laser beam probes, structured-light sensors and Moire topography projectors [Fan, 1997; Wang et al., 2002; Zhang & Wei, 2002; Zhang & Djordjevich, 1999; Wang et al., 1999; Chang et al., 2003; Kim, Choi & Oh, 1999]. These sensors have been mounted on CMM machines to realize non-contact scanning to improve measurement speeds greatly.

Broadly speaking, there are two strategies to reconstruct surface model, namely reconstruction from unorganized set of 3D points and reconstruction that exploits the underlying structure of the acquired data. The first algorithms are often considered as a generalization of the latter ones [Montani et al., 2004].

3.1 Creating Three-Dimensional Object Model from Range Images
There are two main issues in creating a single model from multiple range images: registration and integration [Turk & Levoy, 1994]. Registration refers to computing a rigid transformation that brings the points of one range image into alignment with the portion of a surface that is shared with another range image [Turk & Levoy, 1994]. The aim of
registration is to remove the variations of images due to the misalignment of images, lighting and atmospheric conditions and object movements, growth or other scene changes during data acquisition. Integration is the process where the separate registered range maps are integrated into a single surface representation (often a polygon mesh) [Pulli et al., 1997].

3.1.1 Registration
After obtaining 3D points from range scanning, these points are converted into a triangular mesh that represents the object surface shape. Each of these point triples is made into a triangle if the edge lengths fall below a distance threshold [Hilton et al., 1996]. Initially, all of the range images are measured in the coordinate system corresponding to the range finder system and they are not aligned to each other. The triangular surface meshes from the range images need to be transformed into a unique object coordinate system and the proper alignment between these meshes must be found so that the points in one image can be related to their corresponding points in the other image. This process is to compute the optimal alignment in the sense of minimizing the distances between all points on the surface represented by the two images. Kanatani performed a 3D-rotation fitting [Kanatani, 1994] that maps a set of 3D points to another set by referring to the centroid of each set of points that are translated in space to become the coordinate origin. The optimal estimation of 3D-rotation is used to map the first set of orientation of data to the second set. Chen & Medioni [Chen & Medioni, 1995] proposed a refinement technique that computes a rigid transformation matrix, which minimizes the sum of the distance between a list of points in first view to their corresponding tangent plane in the second view by a linear least-square technique. Bergevin et al. [Bergevin et al., 1996] arranged multiple views in the form of network topology. In this method, each range view is seen as a node in a network of view. The link between each pair of nodes denotes the inter-frame transformation between the two associated views. In order to obtain a well-balanced network of view, a reference node is defined as the world reference frame and its transformation matrix remains unchanged throughout the refinement process [Bergevin et al., 1996]. Besl & McKay [Besl & McKay, 1992] proposed an approach known as an iterated closest-point algorithm (ICP). This method computes the closest point on a geometric shape to a given point. The set of closest points is then transformed so as to minimize the collective distance between the correspondences. This process iterates until the change in mean square error falls below a preset threshold.

3.1.2 Integration
The registered set of surface measurements is combined to form a single 3D surface model through integration. Three main approaches that have been proposed for the integration of range images are mesh integration, volumetric fusion and deformable models [Montani et al., 2004].

Mesh integration techniques aim to merge multiple range surfaces into a single mesh by removing the redundant parts of the overlapping area between the merged surfaces. Turk & Levoy [Turk & Levoy, 1994] proposed a zipperung approach to integrate two triangulated meshes (Figure 6). Zippering two triangle meshes consists of three steps: (i) Removing overlapping portions of the meshes, (ii) Clipping one mesh against another and (iii) Removing the small triangle introduced in clipping. The overlapping portion of the meshes
is removed by discarding the redundant triangle in a mesh. Clipping is used to smoothly join two meshes that slightly overlap by adding new vertices to form a common boundary that the triangle from both meshes will share. Ratishauser, Stricker & Tribina [Ratishauser, Stricker & Tribina, 1994] performed the surface integration by retriangulating the overlapping areas of the triangular meshes. The retriangulation starts with a seed edge. The \( k \)-nearest neighbour-points are selected as candidates to build new triangles. Then the contour of the resulting new surface is updated and the process proceeds with the next edge in the contour. Häusler & Karbacher [Häusler & Karbacher, 1997] proposed a new method to accurately merge range images into a single triangular mesh with curvature dependent density by the use of local topological mesh operations. Operations such as vertex insertion, gap bridging, and surface growth are performed to merge the input meshes by pairs. Mesh thinning process is carried out continuously throughout the integration process to reduce the memory overhead due to the overlapping area of two meshes. After the merging, edge swap operation is performed to produce a stable triangulation.

![Figure 6: Mesh A is clipped against the boundary of Mesh B [Turk & Levoy, 1994] (a) Circles show intersections between edges of A’s and B’s boundary. (b) Portions of triangles from A are discarded (c) Both meshes incorporate the points of intersection.](image)

In the volumetric fusion approach, each range image is projected onto a volumetric grid that is divided into voxels. A scalar value is associated with each node of the grid, generally known as a signed (possibly averaged) distance between the node and the 3D surfaces [Montani et al., 2004]. The surface representation is an isosurface of a spatial field function, \( f(x,y,z) = \text{constant} \). A surface fitting algorithm such as marching cubes [Lorensen & Cline, 1987] is used for the reconstruction of the zero crossing isosurface from the volumetric grid (Figure 7). In the volumetric method proposed by Curless & Levoy, 1996, the volumetric representation of each range image consists of a cumulative weighted signed distance function. Weight is necessary to represent variations in certainty across the range surfaces. Octrees provide a hierarchical volumetric data structure for efficient representation of field functions [Pulli et al., 1997; Wheeler et al., 1998]. For each leaf of the octree, they checked whether the cube was inside, outside or on the boundary of the object. If the cube is neither inside nor outside the object, it is subdivided into eight smaller cubes and added to the octree as the children of the cube. This process is performed recursively by subdividing the parent cube. Triangular meshes are then obtained by generating triangles between the external nodes and those on the boundary. In Wheeler et
al. [Wheeler et al., 1998], the signed distance to the object surface is estimated by finding a consensus of locally coherent observation of the surface. The consensus-surface algorithm proposed eliminates the problems due to the presence of noise in the input data.

Figure 7: Volumetric fusion approach. (a) 15 Combinations of triangulated cubes to create predefined polygon sets for making the appropriate surface approximation [Lorensen & Cline, 1987] (b) Isosurfaces from Marching Cubes.

In the deformable approach, 3D object shapes are obtained from the dynamic deformation of a simple, topologically equivalent 3D surface [Da Silva & Wu, 1998]. By applying this approach, the initial object shape is deformed until it reaches the equilibrium state. Chen & Medioni [Chen & Medioni, 1995] presented an inflating balloon model to construct a complete surface model of an object from a set of range images. The model starts by representing triangulated patches as a shell. The initial triangulated is an isocohedron. The shell grows in size until it reaches the shape of the object to be reconstructed. Other method using similar approach is radial flow model proposed by Da Silva & Wu, 1998.

3.2 Creating Three-Dimensional Object Model from Scattered Points

To reconstruct a 3D object model from a set of scattered points, it is assumed that [Montani et al., 2004]: (a) given a point \( p \), its \( n \) nearest surface neighbours can be located by finding its \( n \)-nearest 3D neighbours, (b) the density of data points are quite uniform over the surface and (c) the points are measured with the same accuracy. This approach does not make any prior assumptions about the connectivity of the points. Boissonnat [Boissonnat, 1984] proposed a method to represent 3D shape by using Delaunay triangulation on a set of points in 3D space. Prior to the triangulation process, the neighborhood of each point is defined as the set of its \( k \)-nearest neighbours. At each steps of the triangulation process, three entities are updated: (i) the triangulated domain, (ii) the contour of the triangulated domain and (iii) the set of the points that are not yet inside the triangular domain. Hoppe et al. [Hoppe et al., 1992] reconstructed implicit surface of an object by using a signed distance function followed by the extraction of isosurface from a scalar function. Point distribution model (PDM), proposed by Cootes et al, 1995 has been shown to be very successful in applications modeling 2D projections of an anatomical object. It is statistical model of shape which can be constructed from training sets of correctly labeled images. PDM relies upon each object or image structure being represented by a set of points from 2D x-ray projections, giving their mean positions and a small set of modes of variations.
which describe how the object’s shape can change [Tang & Ellis, 2005]. The points can represent the boundary, internal features or even external ones, such as the centre of the concave section of boundary. Applying limits to the parameters of the model enforces global shape constraints ensuring that any new examples generated are similar to those of the training set. Note that the shape only provides a surface mesh. To simulate an image volume using the surface mesh, one can intersect the mesh with a set of parallel planes. The result would be a contour of the shape, represented by line segments. Guided by the local surface normal, these segments were “grown” inward to simulate the thickness of the cortical bone in a real CT slice. Hill, Thornham & Taylor, 1993 had extended the PDM technique to 3D images (Figure 8). In this approach, a PDM model is constructed by generating points along the length of each contour of an object. Curves in 3D space that connect these points on the contours are thus generated. These curves are divided into a number of sub-sections. The generated points can now be used to train a PDM as each of their equivalent location on the surface of each example.

![Figure 8: Generating Points for A 3D PDM from Contour Data [Hill, Thornham & Taylor, 1993]](image)

4.0 Conclusion
This paper reviews the methods of reconstructing 3D object models. Based on the nature of the data captured, the reconstruction of 3D object model can be categorized into two broad ranges. In the first approach, a 3D object is reconstructed from 2D images. In terms of radiological modalities, the reconstruction method, known as tomography is applied. This method reconstructs a 3D image by additive layering of 2D images taken from several different positions. A well-known tomography reconstruction technique is filtered via back-projection where the shadow of a 2D x-ray in the fan beam manner is captured and backprojected in order to obtain the 3D view of the image. The 3D images obtained are then used to create a 3D object model by segmentation. Besides, photogrammetric reconstruction provides a means of 3D object model or scene reconstruction from stereoscopic images. Unlike tomography, this approach is restricted to reconstructing the outer body of an object. The main concern of this method is the issue of calibration of the camera. For reconstruction 3D object model from 3D data, there are two strategies: (i) reconstruction from 2.5D range images and (ii) reconstruction from scattered 3D points. The latter is the generalization of the former strategies. In comparison with volumetric fusion and deformable approach, mesh integration is time consuming as the integration between each pair of 2.5D meshes is performed sequentially. On the other hand, although no prior assumption is needed about the connectivity of the points for reconstruction from
scattered 3D points, this method has discarded the information of the object’s topology. Thus, it is not always robust in regions of high curvature. As future advances, the reconstruction of object model should involve as little human intervention as possible. Besides, the reconstruction method has to be robust enough to deal with the problem of incomplete input dataset. The automated reconstruction process can be accomplished by integration of the machine learning approach such as Evolutionary Algorithm (EA) to obtain the approximate object model. To overcome the redundancy or noise data due to the deviation of acquisition, the data mining approach can be applied to extract the control points for reconstruction from the scattered data. These advancements will lead to the increasingly productivity and efficiency in applications that utilise the reconstruction of object model.

REFERENCES


Fenster, A. 2004. 3-Dimensional Ultrasound Imaging. Imaging Economics.


