Review of 3D Reconstruction Methods from Single View

Hetal Chauhan, Riya Parmar, Barkha Bhavsar

1 CE Dept. LDRP-ITR, 2 CE Dept. LDRP-ITR, 3 CE Dept. LDRP-ITR

Abstract - The increase in demand for 3D content has inspired researchers to devise techniques which allow 3D models to be acquired directly from the images. Various techniques have been developed to generate 3D model from 2D images. This paper covers various approaches for 3D generation. Model based techniques use image features such as texture, shading for generating 3D model. Constraint-based techniques use geometric properties like co-linearity, coplanar, point, normal. The geometric approach deals with geometric relationships between points, lines, planes, etc under imaging.

Keywords—3D Reconstruction, Shape from silhouette, 3D mesh

1. INTRODUCTION

2D images contain cues about real object. However, their interpretation is ambiguous because depth information is lost when 3D object is projected on 2D image. Multiple images from different viewpoints resolve these ambiguities. Nowadays cameras of good quality are available. But there are situations where we cannot get multiple images, like we have a image of object and we want to visualize it or we have single photograph of heritage and want to document it. Or even when working with multiple images, parts of the scene appear in only one image due to occlusions or lack of features to match between images then need to generate its 3D model from only one image. So, in these kinds of situations, it is necessary to generate 3D surface using only one image

A. Stereo Vision

It is technique of generating 3D model from overlapping images like human eyes. Each eye captures its own view and the two separate images are sent on to the brain for processing. When the two images arrive simultaneously in the back of the brain, they are united into one picture. The mind combines the two images by matching up the similarities and adding in the small differences. The small differences between the two images add up to a big difference in the final picture. The combined image is more than the sum of its parts.

B. Structure from motion

Structure from motion refers to the process of finding the three-dimensional structure by analyzing the motion of an object over time. Finding structure from motion presents a similar problem as finding structure from stereo vision. In both instances, the correspondence between images and the reconstruction of 3D object needs to be found. To find correspondence between images, features such as corner points need to be tracked from one image to the next. The feature tracked over time are then used to reconstruct their 3D positions and the cameras motion.

C. Structure from focus

It is the problem of reconstructing the depth of the scene changing actively the optics of the camera until the point of interest is in focus. The change in the optics is obtained by changing either the lens position or the object position relative to the camera. A focus measure is computed in the small image regions of each of the image frame in the image sequence. The value of the focus measure increases as the image sharpness or contrast increases and it attains the maximum for the sharpest focused image. Thus the sharpest focused image regions can be detected and extracted. This facilitates auto-focusing of small image regions by adjusting the camera parameters (lens position and/or focal length) so that the focus measure attains its maximum value for that image region. Also, such focused
image regions can be synthesized to obtain a large image where all image regions are in focus. Further, the distance or depth of object surface patches that correspond to the small image regions can be obtained from the knowledge of the lens position and the focal length that result in the sharpest focused images of the surface patches.

D. Shape from silhouette
It uses image outline of object for 3D generation. Silhouette is the dominant source of shape of object. A silhouette is a view of an object or scene consisting of the outline and a featureless interior, with the silhouetted object usually being black. It can use more than one images to extract silhouette of object. From that silhouette various constraints are given like point constraints, normal constraints, depth constraints, and optimized 3D model can be generated.

II. APPROACHES FOR 3D FROM SINGLE VIEW
3D reconstruction techniques like stereovision, structure from motion, structure from focus etc. are used to generate 3D from multiple views. When only single view is available, 3D image is reconstructed using vanishing point, image contour and other constraints like linearity, co-planar etc.

A. Using User Defined Vanishing point
First Single View Reconstruction(SVR)system proposed by Horry [2] in 1997. In that system, GUI is provided to user for specifying vanishing point.3D geometry of the scene’s background is defined as model centering user defined vanishing point. System models the 3D background of the scene using polygons. Foreground objects stand on polygons of the background model.

● Background Modeling
In order to approximate the geometry of the background scene, several polygons should be generated to represent the background with the vanishing point being on its base. From vanishing point defined by user five 2D rectangles may be deduced as shown in Figure-2.6 and the rectangles are tentatively called the floor, right wall, left wall, rear wall, and ceiling, respectively. It defines the textures of these 2D rectangles to be taken from the background image.

It then defines the 3D background model in 3D space as being these five 3D rectangles, assuming that the following conditions hold:
- Every adjacent 3D rectangle mentioned above is orthogonal to the others
- The 3D rear wall is parallel to the view plane
- The 3D floor is orthogonal to the view up vector
- The textures of the 3D rectangles are inherited from those of the corresponding 2D rectangles

● Foreground Modeling
Based on the foreground mask information provided by user, 3D polygonal model for foreground objects in the scene is created. To surround the 2D image of a foreground object, 2D quadrangle in the input image is specified as shown in Figure-2.7. 3D position of the quadrangle in the 3D background model is derived under the condition that the quadrangle should be perpendicularly put on one of the five 3D regions: floor, right wall, left wall, rear wall, and ceiling. In the example of figure 2.7, the person is a foreground object to be modeled, and is surrounded by the quadrangle. The quadrangle in the 3D scene is perpendicularly attached to the 3D floor.

![Figure-1: 3D Background Modeling](image1)

![Figure 2.7: Specifying Foreground Object](image2)
Foreground object model is hierarchy among the polygons.
- Each model consists of one or more polygons. In particular a single polygon itself is a foreground object model.
- For any polygon F1 belonging to the model, another polygon F2 can be added to the model, if F2 is orthogonally attached to F1 so that one side of F2 is on F1. Then F2 is called a child of F1 (or F1 is a parent of F2). This constitutes a hierarchy among the polygons belonging to a foreground object model.
- If a polygon of the model is at the highest level in the hierarchy, it is orthogonally attached to one of the five 3D regions of the 3D background. Then only one side of the highest level is only on the region.

Figure 2.8 illustrates how to construct the foreground object models. First, two quadrangles F0 and F1 are defined on the 3D floor. Then F2 is added to F1 and F3 is added to F2.

![Figure 2.8: Hierarchical Positioning of the Foreground Objects](image)

**B. Vanishing point Automation**

Subsequent systems [3],[4] improved the SVR System. It addresses the problem of extracting three-dimensional geometric information from a single image of a scene. But here vanishing point is not defined by user but automatically estimated by system.

Steps for Vanishing point Estimation
- Automatic Canny edge detection and straight line fitting to obtain the set of straight edge segments E
- Randomly select two segments s1, s2 from edge segments E and intersect them to give the point p;
- The support set Sp is the set of straight edges in E going through the point p;
- Set the dominant vanishing point with the largest support Sp as the vanishing point
- Remove all edges in Sp from E and go to 2 for the computation of the next vanishing point

**c. 3D for Curved Surfaces**

Above techniques are restricted to planar surfaces. They do not permit for the object having curvilinear shape.

2.5 Dimensional Reconstruction

Given a set of user-specified constraints on the shape of the scene, a smooth surface that satisfies the constraints is generated [5]. This problem is formulated as a constrained optimization problem. Measure of surface smoothness is thin plate energy function.

Graphics interface is provided for user-specified constraints like surface positions, normals creases and discontinuity. As each constraint is specified, the system recalculates and displays the reconstruction in real time.

- Point constraints specify the position or the surface normal of any point on the surface.
- Surface discontinuity constraints is a curve across which surface depth is not continuous, creating a tear in the surface.
- Crease constraints specify curves across which surface normals are not continuous.
- Planar region constraints determine surface points that lie on the same plane.
- Fairing curve constraints allow users to control the smoothness of the surface along any curve in the image.

Using given constraints and thin plate function as objective function linear system is formed and iterative optimization technique is used to find solution of linear system.

In this approach, subset of surface that is visible from single image is modeled and represented as f(u,v,z(u,v)). So generated surface is 2.5D rather than 3D, because it...
represents just one side of 3D model.

Three Dimensional Reconstruction
Given a single photo of a curved object, recover 3D model with minimum of user input which projects to object silhouette[6]. Like former approach[5], smoothest surface satisfying user defined constraints is generated. But in this approach, generated 3D surface projects to object silhouette. 3D Surface is represented as \( r(u, v) = [x(u, v), y(u, v), z(u, v)]^T \) where \( u, v \) are value between 0 to 1.

Energy Function: Surface smoothness is defined in terms of an energy on the surface, and it is thin plate energy

\[
E(r) = \int_0^1 \int_0^1 \left[ ||r_{uv}||^2 + 2||r_{vu}||^2 + ||r_{uu}||^2 + ||r_{vv}||^2 \right] du \, dv
\]

Constraints: Various kind of constraints are defined at \((u,v)\) points.

a. Position constraints are of the form \( r(uk, vk) = pk \) for known values of \( uk, vk, pk \)

b. Normal constraints require the surface normal at \((uk, vk)\) to equal a supplied normal \( nk \). The normal to \( r \) at a point is the unit vector along \( ru \) _ rv_. they are of the form

1. \( nk \cdot ru(uk, vk) = 0 \)
2. \( nk \cdot rv(uk, vk) = 0 \)

c. Partial position constraints act on just one component of \( r \), for example \( z(uk, vk) = zk \) constrains only the \( z \) coordinate at \((uk, vk)\)

d. Silhouette constraint forces that contour generator of the 3D surface must project to object silhouette in given 2D image of that object. contour generator is the set of 3D points on the surface at which the viewing direction is in the tangent plane.

e. Inflation constraints provides constraints for \( z \) values. Without inflation generated surface is 2D surface and it projects to given silhouette for which \( z(u,v)=0 \) for all \( u, v \)

This approach represents surface as \( r(u, v) = [x(u, v), y(u, v), z(u, v)]^T \) and finds depth at \((x,y)\) pixel using mapping between \((x,y)\) and \((u,v)\) rather than representing surface as \( r(u, v) = [u, v, z(u, v)] \) like 2.5D reconstruction

### III. COMPARISON

<table>
<thead>
<tr>
<th>Approach</th>
<th>Type of image of application</th>
<th>Characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polygon model based approach</td>
<td>Image of scene with multiple objects without curved surface</td>
<td>They do not permit for the object having curvilinear shape</td>
</tr>
<tr>
<td>Using vanishing point</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5 Dimensional Reconstruction</td>
<td>Image of single object with curved surface</td>
<td>Generated surface is 2.5D rather than 3D, because it represents just one side of 3D model</td>
</tr>
<tr>
<td>Three Dimensional Reconstruction</td>
<td>Image of single object with curved surface</td>
<td>3D mesh is generated. Object can be visualized from all sides</td>
</tr>
</tbody>
</table>

### IV. CONCLUSION

Problem of 3D reconstruction from single image has minimal information because only one image is available instead of multiple images from different views. Generating 3D from single image is the problem of optimization and generated 3D model is approximation rather than actual model. But single view reconstruction is the only way when multiple images are not available or even multiple images are available, but part of image is visible only in one image because of occlusion

### REFERENCES


