Pattern Recognition 41 (2008) 3295 – 3301

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

Shape recovery from turntable sequence using rim reconstruction

H. Zhong*, W.S. Lau, W.F. Sze, Y.S. Hung

Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, China

ARTICLE INFO

Article history:
Received 26 May 2006
Received in revised form 29 March 2008
Accepted 12 May 2008

Keywords:
Silhouette
Rim reconstruction
Surface extraction
Circular motion

ABSTRACT

This paper makes use of both feature points and silhouettes to deliver fast 3D shape recovery. The algorithm exploits object silhouettes in two views to establish a 3D rim curve, which is defined with respect to the two frontier points arising from two views. The images of this 3D rim curve in the two views are matched using cross-correlation technique. A 3D planar rim curve is then reconstructed using point-based reconstruction method. A set of 3D rim curves enclosing the object can be obtained from an image sequence captured under circular motion. Silhouettes are further utilized to check for mismatched rim points. The proposed method solves the problem of reconstruction of concave object surface, which is usually left unresolved in general silhouette-based reconstruction methods. In addition, the property of the reconstructed 3D rim curves allows fast surface extraction. Experimental results with real data are presented.

1. Introduction

The 3D modelling of real world objects is an important problem in computer vision and has many practical applications, such as in virtual reality, video games, motion tracking systems, etc. The point-based approach is the oldest technique for 3D reconstruction. Once feature points across different views are matched, a cloud of feature points lying on the surface of object can be recovered by triangulation methods. However, these points are not ordered, which means that the topological relationship among them is unknown. As a result, they are not readily usable for reconstructing the surface of an object. Besides feature points, silhouettes are prominent features of an image that can be extracted reliably. They provide useful information about both the shape and motion of an object, and indeed are the only features that can be extracted from a textureless object. But they do not provide point correspondences between images due to the viewpoint dependency of silhouettes. There is a need for the development of methods combining both point-based and silhouettes approaches for modelling a 3D object.

The contribution of this paper is that it takes advantage of both point-based and silhouettes approaches to provide fast shape recovery. The method reconstructs a set of 3D rim curves on the object surface from a calibrated circular motion image sequence, with concavities on the object surface retained. These 3D rim curves are reconstructed in a known spatial order, leading to low computation complexity in the subsequent surface triangulation.

This paper is organized as follows. Section 2 briefly reviews the literature on model reconstruction. Section 3 introduces the theoretical principles which are used in this paper. Section 4 describes how the rim curves are computed. The extraction of surface from the reconstructed rim curves is given in Section 5. Experimental results are given in Section 6 and a summary follows in Section 7.

2. Previous works

There are two main streams of methods for model reconstruction, namely point-based and silhouette-based methods. In a point-based method, a 3D model is obtained by computing a dense depth map [1,2] of the 3D object from the calibrated images using stereo matching techniques [3–5]. In this kind of model acquisition methods, a metric reconstruction is first computed using matched correspondences. As only a limited number of 3D points are initially reconstructed from matched feature points, dense correspondence maps between images must be obtained in order to recover the depths of all the object points. For instance, this can be done by stereo matching algorithms such as using dynamic programming to determine the global 3D maximum-surface [6] and graph cuts to deal with occlusions [7]. It is then necessary to fuse the entire dense depth map into a common 3D model [8] and extract coherently connected surface by interpolation [9]. Although a good and realistic 3D model may be obtained by carrying out these processes, both the estimation of dense depth maps and the fusion of them are of high computation complexity and reported to be formidably time consuming [10].

* Corresponding author. Tel.: +852 2859 2728; fax: +852 2559 8738.
E-mail addresses: hzhong@eee.hku.hk (H. Zhong), sunny@eee.hku.hk, wfsze@eee.hku.hk, yshung@eee.hku.hk.

0031-3203/$30.00 © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.patcog.2008.05.017

Another model reconstruction approach utilizes object silhouettes instead of feature points. There are two groups of methods in this approach, namely surface and volumetric methods. The surface method was pioneered by Giblin and Weiss [11] who demonstrated the possibility of reconstructing surfaces from apparent contours for the case of orthogonal projection under planar viewer motion. This was later extended to perspective projection in Ref. [12] where a more direct approach making use of the epipolar parameterization was introduced. In Ref. [13], volume segment representation was first introduced for constructing the bounding volume which approximated the actual 3D structure of the rigid object generating the contours in multiple views. This is known as the volume intersection technique. In this technique, octree representation is adopted to provide a volumetric description of an object in terms of regular grids or voxels [14–17]. Given the volumetric data in the form of an octree, the marching cubes algorithm [18] can be used to extract a triangulated surface mesh for visualization. Since only silhouettes are used, no feature points are needed. However, such techniques produce only the visual hull [19,20] of the object with respect to the set of viewpoints from which the image sequence is captured. Hence the main drawback is that concave surfaces of the object cannot be reconstructed.

More recently, many methods combining photometric and silhouette information to achieve higher-quality reconstruction results were developed, e.g. Refs. [21–25]. Broadly speaking, these methods all use a two-step approach. In the first step, a surface mesh is initialized by computing a visual hull satisfying the silhouette consistency constraint. In the second step, the surface mesh is refined using the photo-consistency constraint. The key difference of these methods lies in how the surface mesh is refined. Specifically, in Ref. [21] the mesh was deformed by using iterative method to minimize reprojection errors evaluated at randomly sampled points in the texture space error image; in Ref. [22] the surface mesh was deformed using texture and silhouette forces; and in Refs. [23–25] graph cuts formulation was used to optimize the surface reconstruction. A detailed evaluation and comparison of these multi-view reconstruction algorithms can be found in Ref. [26].

In this paper, we propose to use silhouettes and feature points to develop a practical solution for shape recovery from a calibrated circular motion image sequence. Unlike the methods mentioned above which aimed to reconstruct the whole photo-consistent surface based on a visual hull, the proposed method computes discrete surface curves which are circumnavigating the object to represent the shape with no surface initialization required. The discrete surface curves are computed to satisfy both the photo-consistency and silhouette constraints, and hence readily represent the photo-consistent surface of the object. As a result, surface triangle patches can be formed by a simple strategy of extracting a mesh from the intersections of the set of reconstructed discrete surface curves and the parallel slicing planes. This is the major difference of the proposed method as compared to the aforementioned methods in surface formation. The advantages of using discrete surface curves to represent object shape and to form surface are two-fold. First, no visual hull and time-consuming surface optimization are needed. Applying the silhouette and photo-consistency constraints to individual rim curves does not require iteration and is computationally simpler than the surface optimization process of [21,22]. Secondly, surface concavities can be recovered. This is particularly suitable for fast modelling with more accurate shape description than visual hull. Experimental results show that through combining the photo-consistency and silhouette constraints, the surface curves can be computed accurately.

3. Theoretical principles

3.1. Geometry of the surface

Consider an object with a smooth surface viewed by a pin-hole camera. The *contour generator* on the object surface is the set of points at which the rays cast from the camera center is orthogonal to the surface normal [27]. The projection of a contour generator on the image plane forms an *apparent contour*. A *silhouette* is a subset of the apparent contour where the viewing rays of the contour generator are tangent to the object surface [28].

Due to the viewpoint dependency of the contour generators, silhouettes from two distinct viewpoints will be the projections of two different contour generators. As a result, there will be no point correspondence in the two silhouettes except for the *frontier point*(*s*) [27]. A frontier point is the intersection of the two contour generators in space and is visible in both silhouettes. For two distinct views, there are in general two frontier points at which the two outer epipolar planes are tangent to the object surface. Hence, the projections of them in one view are two outer epipolar tangent points on the object silhouette. The outer epipolar tangent points in the two associated views are corresponding points. Note that as pointed out in Ref. [29], due to noisy calibrations and discretization of silhouettes in image plane, frontier points in practice cannot be determined from viewing cones since perfect cone tangency do not occur. In our case, the quantization and calibration noise will also affect the accuracy of matching of the outer epipolar tangents points, i.e. they may not be perfectly matched points. Nonetheless, as other points are matched independently, this effect is negligible in the final reconstruction.

3.2. Definition of a 3D rim curve

A 3D rim curve on an object surface is defined with respect to two distinct views as follows. The two frontier points associated with the two views, together with one of the two camera centers, define a plane. This plane intersects the object surface in a 3D planar curve on the object surface. The two frontier points are on this curve and they cut the curve into segments. The one closer to the chosen camera center is defined a 3D rim curve. The projection of this rim curve in the view associated with the chosen camera is a straight line, and the image of it in the other view is generally a 2D rim curve. Fig. 1 illustrates the 3D rim curve and 2D rim curve so defined. Note that a "3D rim curve" is different from a "rim" defined in many other papers in that a 3D rim curve is not necessarily a contour generator, but a "rim" is. To avoid confusion, we will use "3D rim curve" and

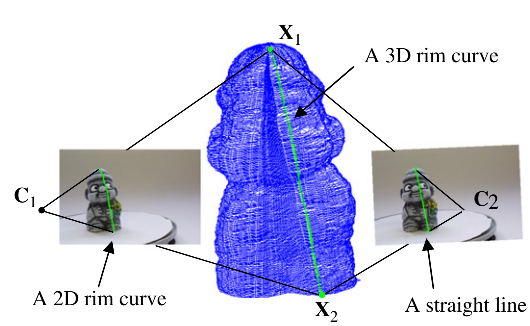

Fig. 1. A 3D rim curve associated with frontier points for two views. The two cameras' centers C_1 and C_2 define two epipolar planes tangent to the objects at two frontier points X_1 and X_2 whose images are outer epipolar tangent points. The plane containing C_2, X_1, and X_2 cuts the object at a 3D planar rim curve. The curve segment visible to cameras C_1 and C_2 projects onto them as a 2D rim curve and a staright line, respectively.