Int J Comput Vis (2013) 102:239–255
DOI 10.1007/s11263-012-0572-1

Direct Model-Based Tracking of 3D Object Deformations in Depth and Color Video

Andreas Jordt · Reinhard Koch

Received: 4 November 2011 / Accepted: 7 September 2012 / Published online: 28 September 2012

Abstract The tracking of deformable objects using video data is a demanding research topic due to the inherent ambiguity problems, which can only be solved using additional assumptions about the deformation. Image feature points, commonly used to approach the deformation problem, only provide sparse information about the scene at hand. In this paper a tracking approach for deformable objects in color and depth video is introduced that does not rely on feature points or optical flow data but employs all the input image information available to find a suitable deformation for the data at hand. A versatile NURBS based deformation space is defined for arbitrary complex triangle meshes, decoupling the object surface complexity from the complexity of the deformation. An efficient optimization scheme is introduced that is able to calculate results in real-time (25 Hz). Extensive synthetic and real data tests of the algorithm and its features show the reliability of this approach.

Keywords Deformation · Tracking · Range video

1 Introduction

In the field of computer vision, changing scenes and deforming, flexible objects have always been a very challenging subject to 3D reconstruction. The classic reconstruction from color video based on structure from motion (SfM) (Hartley and Zisserman 2000) encounters intractable ambiguity problems in these environments, since the constancy assumption does not hold.

Time of Flight cameras and the recently released Kinect camera provide a possibility to solve this problem by providing reliable, dense depth images at high frame rates. In this article we will introduce a novel tracking algorithm using depth and color video.

1.1 Related Work

Although there are several algorithms in deformation reconstruction using contour information (Rosenhahn et al. 2007), patchlet based reconstruction (de Aguiar et al. 2007) or optical flow (Hilsmann and Eisert 2009), most of the work on deformation tracking has been done utilizing feature points. Correspondences tracked throughout an image sequence provide a comfortable data foundation to formulate mathematically sound solutions to a tracking problem.

A well known class of feature-based approaches that has been studied for over a decade is Non-Rigid Structure from Motion (NRSfM) (Del Bue et al. 2007). NRSfM is based on the factorization method (Tomasi and Kanade 1992), which allows to simultaneously recover 3D shape and motion from a set of 2D feature correspondences—the Structure from Motion (SfM) problem (Hartley and Zisserman 2000)—which was extended to the non-rigid case over the years. The essential SfM constraint, the assumption of a constant scene, was first relaxed to 'piece-wise' rigidity by Costeira and Kanade (1994) who allowed multiple moving elements by segmenting the features into separate movements.

Bregler et al. (2000) introduced the NRSfM method, extending the pose space by a set of linear deformation dimensions spanned by basis shapes, so the factorization yields

This work was supported by the EU Interreg A4 Project "Intelligent Robots for Handling flexible Objects" (IRFO), 33-1.2-09.

A. Jordt (✉) · R. Koch
Multimedia Information Processing Group, University of Kiel,
Hermann-Rodewald-Str. 3, 24098 Kiel, Germany
e-mail: jordt@mip.informatik.uni-kiel.de

R. Koch
e-mail: rk@mip.informatik.uni-kiel.de