Machine Vision and Applications (2005) 16: 122–127
Digital Object Identifier (DOI) 10.1007/s00138-004-0164-3

Machine Vision and Applications

Tracking-based depth recovery for virtual reality applications

Dan Zetu[1], **Ali Akgunduz**[2]

[1] Power Information Network, LLC, An Affiliate of J.D. Power and Associates, 5435 Corporate Dr., Suite 300, Troy, MI 48098, USA (e-mail: Dan.Zetu@jdpa.com)

[2] Department of Mechanical and Industrial Engineering, Concordia University, 1455 de Maisonneuve Blvd, W, Montreal, QC H3G 1M8, Canada (e-mail: akgunduz@me.concordia.ca)

Received: 16 August 2003 / Revised version: 13 August 2004
Published online: 20 December 2004 –

Abstract. This paper describes a technique for tracking-based reconstruction from camera images. The position and orientation data are provided by a commercial laser tracker rigidly attached to the camera. The tracker enables the determination of the extrinsic camera parameters. Focal length is automatically calibrated in parallel with the depth computation process by means of a Kalman filter. In addition to its simplicity, our technique enables the fast generation of three-dimensional models from two-dimensional images to be used in virtual reality applications.

Keywords: Camera-tracker unit – Laser tracking – Camera calibration – Kalman filtering – 3D reconstruction

1 Introduction

Reconstructing three-dimensional (3D) objects from camera images has been a challenging task. It is well known that if all camera parameters (i.e., intrinsic and extrinsic) [7,20] are known, and provided a set of features has been extracted from images and feature correspondences have been established, the reconstruction is straightforward [18]. Typically, camera parameters are determined through camera calibration [7,5]. The disadvantage of standard calibration algorithms is that they require the presence, within the camera's field of view, of a calibration pattern each time an image is captured. Such a calibration pattern is not always available throughout the image-acquisition process. If we rely only on images, with no knowledge about camera parameters, unambiguous reconstruction cannot be achieved. We can still extract important information from images [4], but that information is not suitable for virtual reality (VR) applications (where full-scale, metric reconstruction is needed). If no information besides images is available, reconstruction can be achieved only up to an unknown projective transformation and therefore is not useful for VR applications. If only the intrinsic parameters are known, reconstruction can be achieved only up to an unknown scale factor [9]. The scale factor can be manually determined by physically measuring one dimension of the real scene and comparing it with the same dimension as given by the algorithm.

The reconstruction possibilities in various cases are summarized in Table 1. What happens when one needs to reconstruct a model from a long sequence of images, such as a portion of a factory floor? In this case, it is not practical to repeatedly measure physical dimensions to be compared to the ones given by the algorithm. The user might not know beforehand what dimensions to measure because it is not guaranteed that they will be visible in any of the captured images. It is obvious that we need both intrinsic and extrinsic camera parameters in order to achieve accurate metric reconstruction. Consequently, a method for recovering these parameters without using a calibration pattern is needed.

We describe a methodology for recovering extrinsic and intrinsic camera parameters by collecting camera pose data from a tracking system rigidly attached to the camera. As the camera changes its pose, a new measurement from the tracking system is collected and the camera pose is updated accordingly. By using the tracking system, only the extrinsic parameters can be recovered. To recover the intrinsic parameters (i.e., focal length), we design a Kalman filter that enables simultaneous depth computation and focal length autocalibration.

The remainder of this paper is organized as follows. Related work is discussed in Sect. 2. The depth extraction methodology is described in detail in Sect. 3. Experimental results are shown in Sect. 4, and conclusions are drawn in Sect. 5.

2 Related work

To achieve calibration-free 3D reconstruction, most researchers have assumed that images are the only information available. A methodology for recovering structure from two images via the fundamental matrix is described in [23,26]. The fundamental matrix is used to recover the epipolar geometry relating two images, and then, provided at least eight matches are available in both images, the 3D structure is recovered up to an unknown scale factor (intrinsic parameters are assumed to be available). The problem with this technique is that the scale factor differs for each pair of images. Techniques sim-

Correspondence to: A. Akgunduz