Int J Comput Vis (2013) 102:129–145
DOI 10.1007/s11263-012-0568-x

A Scale Independent Selection Process for 3D Object Recognition in Cluttered Scenes

Emanuele Rodolà · Andrea Albarelli · Filippo Bergamasco · Andrea Torsello

Received: 12 October 2011 / Accepted: 1 September 2012 / Published online: 19 September 2012

Abstract During the last years a wide range of algorithms and devices have been made available to easily acquire range images. The increasing abundance of depth data boosts the need for reliable and unsupervised analysis techniques, spanning from part registration to automated segmentation. In this context, we focus on the recognition of known objects in cluttered and incomplete 3D scans. Locating and fitting a model to a scene are very important tasks in many scenarios such as industrial inspection, scene understanding, medical imaging and even gaming. For this reason, these problems have been addressed extensively in the literature. Several of the proposed methods adopt local descriptor-based approaches, while a number of hurdles still hinder the use of global techniques. In this paper we offer a different perspective on the topic: We adopt an evolutionary selection algorithm that seeks global agreement among surface points, while operating at a local level. The approach effectively extends the scope of local descriptors by actively selecting correspondences that satisfy global consistency constraints, allowing us to attack a more challenging scenario where model and scene have different, unknown scales. This leads to a novel and very effective pipeline for 3D object recognition, which is validated with an extensive set of experiments and comparisons with recent techniques at the state of the art.

Keywords Object recognition · Partial surface registration · Game theory · Object in clutter

E. Rodolà (✉) · A. Albarelli · F. Bergamasco · A. Torsello
Dipartimento di Scienze Ambientali, Informatica e Statistica, Università Ca' Foscari Venezia, Venice, Italy
e-mail: rodola@dsi.unive.it

A. Albarelli
e-mail: albarell@unive.it

F. Bergamasco
e-mail: bergamasco@dsi.unive.it

A. Torsello
e-mail: torsello@dais.unive.it

1 Introduction

In the recent past, the acquisition of 3D data was only viable for research labs or professionals that could afford to invest in expensive and difficult to handle high-end hardware. However, due to both technological advances and increased market demand, this scenario has been altered significantly: Semi-professional range scanners can be found at the same price level of a standard workstation, widely available software stacks can be used to obtain reasonable results even with cheap webcams, and, finally, range imaging capabilities have been introduced even in very low-end devices such as game controllers. Given this trend, it is safe to forecast that range scans will be so easy to acquire that they will complement or even replace traditional intensity-based imaging in many computer vision applications. The added benefit of depth information can indeed enhance the reliability of most inspection and recognition tasks, as well as providing robust cues for scene understanding or pose estimation. Many of these activities include fitting a known model to a scene as a fundamental step. For instance, a setup for inline quality control within a production line could raise the need to locate the manufactured objects that are meant to be measured (Newman and Jain 1995). Range-based SLAM systems can exploit the position of known 3D reference objects to achieve a more precise and robust robot localization (Borrmann et al. 2008). Finally, non-rigid fitting may be used to recognize hand or whole-body gestures in next