Int J Comput Vis
DOI 10.1007/s11263-013-0654-8

A Super-Resolution Framework for High-Accuracy Multiview Reconstruction

Bastian Goldlücke · Mathieu Aubry · Kalin Kolev · Daniel Cremers

Received: 23 May 2012 / Accepted: 12 August 2013

Abstract We present a variational framework to estimate super-resolved texture maps on a 3D geometry model of a surface from multiple images. Given the calibrated images and the reconstructed geometry, the proposed functional is convex in the super-resolution texture. Using a conformal atlas of the surface, we transform the model from the curved geometry to the flat charts and solve it using state-of-the-art and provably convergent primal–dual algorithms. In order to improve image alignment and quality of the texture, we extend the functional to also optimize for a normal displacement map on the surface as well as the camera calibration parameters. Since the sub-problems for displacement and camera parameters are non-convex, we revert to relaxation schemes in order to robustly estimate a minimizer via sequential convex programming. Experimental results confirm that the proposed super-resolution framework allows to recover textured models with significantly higher level-of-detail than the individual input images.

B. Goldlücke (✉)
Heidelberg Collaboratory for Image Processing,
University of Heidelberg, Heidelberg, Germany
e-mail: bastian.goldluecke@iwr.uni-heidelberg.de

M. Aubry · D. Cremers
Computer Science Department, Technical University of Munich,
Munich, Germany
e-mail: mathieu.aubry@in.tum.de

D. Cremers
e-mail: cremers@tum.de

K. Kolev
Department of Computer Science, ETH Zurich, Zurich, Switzerland
e-mail: kalin.kolev@inf.ethz.ch

Keywords Multi-view 3D reconstruction · Texture reconstruction · Super-resolution · Camera calibration · Variational methods

1 Introduction

Modern image-based 3D reconstruction algorithms achieve high levels of geometric accuracy. However, due to intrinsic limitations like voxel volume resolution and mesh size, the geometric resolution of the model is usually well below the pixel resolution in a rendering. This leads to a number of problems if one wants to estimate a high-resolution texture for the model from the camera images. Most importantly, since geometry is never perfectly accurate on the level of individual texels, the image registration cannot be exactly correct, which leads to a blurry estimated texture, see Fig. 1. Consequently, previous methods on texture generation usually employ some form of additional registration before estimating texel color (Bernardini et al. 2001; Lensch et al. 2001; Theobalt et al. 2007).

In methods fitting a local appearance model on a per-texel basis, it is generally true that the fewer source cameras influence the result for a single texel, the sharper the resulting texture will be. However, if only the contributions of few cameras are blended for a given texture patch, it is likely that seams and discontinuities arise at visibility boundaries, so some form of stitching has to take place to smoothen the result (Allne et al. 2008; Lempitsky and Ivanov 2007). Furthermore, not using all available source images implies discarding a lot of potentially useful information.

In particular, in multi-view settings, usually every patch of the surface is captured from several cameras. Therefore, using a suitable super-resolution model, one should be able to recover the texture map in higher resolution than provided by

Published online: 25 August 2013