Int J Comput Vis (2012) 99:1–24
DOI 10.1007/s11263-012-0515-x

TILT: Transform Invariant Low-Rank Textures

Zhengdong Zhang · Arvind Ganesh · Xiao Liang · Yi Ma

Received: 9 December 2010 / Accepted: 2 January 2012 / Published online: 12 January 2012
© Springer Science+Business Media, LLC 2012

Abstract In this paper, we propose a new tool to efficiently extract a class of "low-rank textures" in a 3D scene from user-specified windows in 2D images despite significant corruptions and warping. The low-rank textures capture geometrically meaningful structures in an image, which encompass conventional local features such as edges and corners as well as many kinds of regular, symmetric patterns ubiquitous in urban environments and man-made objects. Our approach to finding these low-rank textures leverages the recent breakthroughs in convex optimization that enable robust recovery of a high-dimensional low-rank matrix despite gross sparse errors. In the case of planar regions with significant affine or projective deformation, our method can accurately recover both the intrinsic low-rank texture and the unknown transformation, and hence both the geometry and appearance of the associated planar region in 3D. Extensive experimental results demonstrate that this new technique works effectively for many regular and near-regular patterns or objects that are approximately low-rank, such as symmetrical patterns, building facades, printed text, and human faces.

Keywords Transform invariant · Low-rank texture · Sparse errors · Robust PCA · Rank minimization · Image rectification · Shape from texture · Symmetry

Z. Zhang
Department of Computer Science and Technology, Tsinghua University, Beijing, China
e-mail: zhangzdfaint@gmail.com

A. Ganesh (✉) · Y. Ma
Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA
e-mail: abalasu2@illinois.edu

Y. Ma
e-mail: yima@illinois.edu

X. Liang
Institute for Advanced Study, Tsinghua University, Beijing, China
e-mail: liangx04@mails.tsinghua.edu.cn

Y. Ma
Visual Computing Group, Microsoft Research Asia, Beijing, China
e-mail: mayi@microsoft.com

1 Introduction

One of the fundamental problems in computer vision is to identify certain feature points or salient regions in images. These points and regions are the basic building blocks for almost all high-level vision applications such as image matching, 3D reconstruction, object recognition, and scene understanding. Through the years, a large number of methods have been proposed in literature for extracting various types of feature points or salient regions. The detected points or regions typically represent parts of the image that have distinctive geometric or statistical properties such as Canny edges (Canny 1986), Harris corners (Harris and Stephens 1988), and textons (Leung and Malik 2001).

One of the important applications of detecting feature points or regions in images is to establish point-wise correspondences or measure similarity between different images of the same object. This problem is especially challenging if the images are taken from different viewpoints under different lighting conditions. Thus, it is desirable that the detected points/regions are somewhat stable or invariant under transformations incurred by changes in viewpoint or illumination. In the past two decades, numerous "invariant" features and descriptors have been proposed, studied, compared, and combined in the literature (see Mikolajczyk