Int J Comput Vis (2012) 97:322–338
DOI 10.1007/s11263-011-0490-7

Sparse Occlusion Detection with Optical Flow

Alper Ayvaci · Michalis Raptis · Stefano Soatto

Received: 12 July 2010 / Accepted: 5 August 2011 / Published online: 6 October 2011

Abstract We tackle the problem of detecting occluded regions in a video stream. Under assumptions of Lambertian reflection and static illumination, the task can be posed as a variational optimization problem, and its solution approximated using convex minimization. We describe efficient numerical schemes that reach the global optimum of the relaxed cost functional, for any number of independently moving objects, and any number of occlusion layers. We test the proposed algorithm on benchmark datasets, expanded to enable evaluation of occlusion detection performance, in addition to optical flow.

Keywords Occlusion detection · Optical flow · Convex optimization · Sparse optimization · Nesterov's algorithm · Split-Bregman method

1 Introduction

Occlusion phenomena are a critical component of the image formation process, shaping the statistics of natural images. Occlusion detection plays an important role in priming visual recognition of detached objects (Ayvaci and Soatto 2011), in navigation and interaction with natural environments, and more in general in the visual information-gathering process: The complexity of the image in the "discovered (disoccluded) region" is related to the *Actionable Information Increment* in visual exploration (Soatto 2011).

Occlusions arise when a portion of the scene is visible in one image, but not another. In Da Vinci Steropsis, portions of the scene that are visible from the left eye are not visible from the right eye, and vice-versa. In a video-stream, occlusions typically occur at depth discontinuities. We are interested in determining the *occluded regions*, that is the subset of an image domain that back-projects onto portions of the scene that are not *co-visible* from a temporally adjacent image.[1] The occluded region is, in general, multiply connected, and can be quite complex, as the example of a barren tree illustrates.

Portions of the scene that are *co-visible* can be mapped onto one another by a domain deformation (Sundaramoorthi et al. 2009), called *optical flow*. It is, in general, different from the *motion field*, that is the projection onto the image plane of the spatial velocity of the scene (Verri and Poggio 1989), unless three conditions are satisfied: (a) Lambertian reflection, (b) constant illumination, and (c) constant visibility properties of the scene. Most surfaces with benign reflectance properties (diffuse/specular) can be approximated as Lambertian almost everywhere under sparse illuminants (e.g., the sun). In any case, widespread violation of Lambertian reflection does not enable correspondence, so most optical flow methods embrace (a), either implicitly or explicitly. Similarly, constant illumination (b) is a reasonable

A. Ayvaci and M. Raptis contributed equally to this work.

A. Ayvaci (✉) · M. Raptis · S. Soatto
University of California, Los Angeles, 405 Hilgard Ave, Boelter Hall #3811, Los Angeles, 90095 CA, USA
e-mail: ayvaci@cs.ucla.edu

M. Raptis
e-mail: mraptis@cs.ucla.edu

S. Soatto
e-mail: soatto@ucla.edu

[1] This process could be generalized to global co-visibility, resulting in a model of the world with topologically distinct "layers" (Wang and Adelson 1994). This is beyond the scope of this paper and has already been addressed in a variational setting (Jackson et al. 2005, 2008).