Computer Vision and Image Understanding 117 (2013) 688–704

Contents lists available at SciVerse ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Occlusion filling in stereo: Theory and experiments

Shafik Huq, Andreas Koschan *, Mongi Abidi

Min Kao Department of Electrical Engineering and Computer Science, The University of Tennessee at Knoxville, TN 37996, USA

ARTICLE INFO

Article history:
Received 17 February 2011
Accepted 17 January 2013
Available online 1 February 2013

Keywords:
Stereo matching
Occlusion
Probability
Image noise
Clustering

ABSTRACT

A number of stereo matching algorithms have been developed in the last few years, which also have successfully detected occlusions in stereo images. These algorithms typically fall short of a systematic study of occlusions; they predominantly emphasize matching and regard occlusion filling as a secondary operation. Filling occlusions, however, is useful in many applications such as image-based rendering where 3D models are desired to be as complete as possible. In this paper, we study occlusions in a systematic way and propose two algorithms to fill occlusions reliably by applying statistical modeling, visibility constraints, and scene constraints. We introduce a probabilistic, model-based filling order of the occluded points to maintain consistency in filling. Furthermore, we show how an ambiguity in the interpolation of the disparity value of an occluded point can safely be avoided using color homogeneity when the point's neighborhood consists of multiple scene surfaces. We perform a comparative study and show that statistically, the new algorithms deliver good quality results compared to existing algorithms.

Published by Elsevier Inc.

1. Introduction

Stereo vision has been a subject of research for many years [1]; in stereo vision, disparities between points in a scene are estimated from two or more images of the scene. In many recent works, occlusions are detected in addition to estimating a disparity map [2–9]; in much of the work recently reported the detected occlusions are filled by assigning a derived or estimated disparity value to the occluded points. These works regard the occlusion filling as a secondary problem. Therefore, a lack of systematic study of occlusion filling is noted in the literature on stereo vision. Many algorithms consider occluded regions as noise during matching and thus avoid detection or filling of occlusions [10–13]; if occlusions are detected, detection is performed implicitly [14,15] or explicitly [1,3,5]. In implicit detection, occlusions are handled while matching is established. In the explicit version of occlusion detection, occlusions are first detected by performing matching both ways, i.e., left to right and right to left, followed by comparing disparity values of corresponding points from both disparity maps [3,5]. Inequality of magnitudes of the two-way disparity values for a matching point indicates that the point is occluded. In a subsequent step the detected occluded points are filled. Although algorithms for temporal occlusion filling exist, where the background model is learned from disocclusions of previous frames [16], in this paper, our focus is occlusion filling in static stereo images.

The disparity value of an occluded point is usually in agreement with the slope of the plane that fits disparities of the point's non-occluded neighbors (this plane is called the disparity plane). Occluded regions, filled directly with neighbors' disparities, appear inconsistently as fronto-parallel shapes unless the disparity plane is fronto-parallel. In the papers of Yang et al. [5], Sun et al. [15], and Wang et al. [17], the disparity estimation of an occluded point did not consider the slope of the disparity plane. In [5] and [15], the disparity was directly assigned with the disparity of the horizontally closest left non-occluded neighbor. Although disparity information was used in occlusion filling in [17], the usage was limited to depth segmentation of widely separated objects, which is not always the case in stereo images of a scene. Hosni et al. [9] assign a disparity value to an occluded point based on the minimum among disparities of its horizontally closest left and right non-occluded points; this method is essentially the same as considering disparities of only the left non-occluded points as in [5,15], since the left points always have smaller disparities than those of the right points. Occlusion filling in these ways often introduces horizontal streaks in the disparity maps. In order to smooth out the streaks, Hosni et al. [9] applied a median filter based on weights computed from minimum geodesic distance in color. The geodesic distance between two pixels was the lowest aggregated cost along a path between the points, where a cost is Euclidean distance between three color channels of two neighboring points. The filter, which is difficult to design appropriately because of its manually chosen parameters, cannot completely correct errors in the filling, but can only mitigate them up to a limit. In the paper of Yang et al. [18], disparities of occluded points

* Corresponding author.
E-mail address: akoschan@utk.edu (A. Koschan).

1077-3142/$ - see front matter Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.cviu.2013.01.008

were obtained using Graph-cuts [19,20]. Graph-cuts assign a number of disparity values ranging from 0 to a scene maximum for the occluded points and compute an energy cost for each assignment. The disparity value that requires the least energy is assigned to the occluded points. The algorithm continues assigning different disparity values in several consecutive iterations until no further decrement in energy is observed. Graph-cuts are known to have stair-like effects in disparity maps, which can worsen while filling occlusions, since the cost function of Graph-cuts does not have a valid*data cost* term inside occluded regions. Besides, the cut-off function for the *smoothness cost* is not designed to carry surface slope information further out into the occluded region.

Furthermore, algorithms exist that detect and fill occlusions during the stereo matching process by using a data term that is occlusion-aware [20–22], i.e., they give a constant penalty for occluded pixels and return the matching costs otherwise. By using such data terms, occlusions are filled automatically via disparity extrapolation using the smoothness term of the energy function. Among these algorithms, Kolmogorov and Zabih [20] do not consider the slope of the disparity plane in occlusion filling, while Woodford et al. [21] and Bleyer et al. [22] do consider the slope. In another implicit algorithm, Klaus et al. [14] segment disparity planes of similar color intensities iteratively using a mean shift algorithm and analysis of matching costs. In their algorithm, disparity of an occluded point is implicitly extrapolated from the disparity plane. In this iterative scheme of the stereo-matching algorithm, current matching cost is used for segmentation of the disparity plane. Although the slope of the disparity plane is included in the disparity estimation in [21,22,14], the occlusion filling algorithm is not independent of the matching algorithm, and therefore cannot be applied to algorithms that detect occlusions explicitly in a separate processing step after matching is performed. Min and Sohn [3] detect occlusions explicitly from cross-checking of disparity maps obtained from both-way matching. Later, occlusions are filled by diffusing energies of neighboring non-occluded points. The energies are obtained from the matching process and therefore occlusion filling is still not independent. In addition, the disparity plane of the occluded points found through this process is not guaranteed to be in agreement with the slope of the disparity plane that contains the surrounding non-occluded points. Oh and Kuo [1] fill occlusions by applying intensity cues and interpolation to avoid slope disagreement. However, in their work ambiguity in occlusion filling is not resolved and occlusion filling order is not maintained; we will see later that ambiguity and order are two aspects of an occlusion filling algorithm which should not be neglected. Table 1 summarizes working principles, pros, and cons of the occlusion filling algorithms mentioned above.

In this paper, we describe two new algorithms for occlusion filling. One is based on absolute color difference and weighted least squares and the other one is based on least squares with segmented points. We compare the results obtained from these two new algorithms with results obtained applying two of the algorithms mentioned above: one is the neighbor's disparity assignment [4,5,9,15] and the other one is an extension of the diffusion in intensity method [3]. All algorithms assume that occlusions are already detected. The new algorithm with the best performance takes surface slope into account when filling the occlusion using linear interpolation. The contributions of this paper include:

(1) A comprehensive study of different occlusion types and their origins in stereo images.
(2) Discussion of a number of ambiguities in occlusion filling and methods for their removal via application of various scene constraints and cues.
(3) Introduction of a specific filling order for the occluded points to achieve higher accuracy in occlusion filling; the order is determined by applying color homogeneity of an occluded point, where homogeneity is defined in a probabilistic framework to avoid manual thresholds/parameters.

We demonstrate that the accuracy of occlusion filling in stereo vision can be improved when applying a filling order in addition to ambiguity removal. These occlusion filling algorithms are independent of the stereo matching methods. In order to be usable with the widest possible number of stereo matching methods, these newly developed algorithms can also be applied as a post-processing step with any stereo matching algorithms that also deliver occlusion maps. In Section 3 we introduce a detailed theory of occlusions, where different kinds of occlusions are studied and the stereo scene surface arrangements that are responsible for the occurrence of each of these kinds are pictorially elucidated. Section 4 introduces a theorem of several pre-existing and our newly developed occlusion filling algorithms. Here, we focus on exclusively filling partial occlusions since partial occlusions are by far the dominant occlusion type. Section 5 presents experimental results including comparisons between previous methods and our new algorithms; in our experiments, we use ground truth disparity maps and disparity maps generated by 12 different stereo matching algorithms listed in the Middlebury College stereo algorithm evaluation site. Middlebury images have ground truth disparity maps with occlusions labeled and they are available to the computer vision community through the worldwide web [23]. In Section 6, we draw conclusions and suggest some performance improvements.

Table 1
Summary of occlusion filling algorithms.

Papers	Occlusion filling strategies	Working principles, pros, and cons
Oh and Kuo [1]	Intensity cue	Assigns disparity to the occluded point from non-occluded points that have similar color intensities. Applying interpolation, occlusion filling considers slope of the planes belonging to the occluded point. Ambiguity in occlusion filling is not resolved and occlusion filling order is not addressed
Min and Sohn [3]	Diffusion in intensity	Assigns disparity to the occluded point from non-occluded points with similar color intensities. Occlusion filling does not consider slope of the planes belonging to the occluded point
Klaus et al. [14]	Implicit filling during the matching process	Occlusion filling is done implicitly during the matching process by segmenting planes with similar color intensities. Occlusion filling of occluded planes remains undefined. The occlusion filling algorithm is not independent to matching
Yang et al. [4,5], Sun et al. [15], Hosni et al. [9], and Wang et al. [17]	Neighbor's disparity assignment	Occlusion filling does not account for the slope of the plane that belongs to the occluded point. With this filling, disparity maps may contain visible artifacts in occluded regions (such as horizontal streaks)
Yang et al. [18]	Graph-cuts	Since occluded regions usually do not have matched points, the algorithm only has a smoothness term. The cut-off function for smoothness term fails to carry surface slope information further out into the occluded region