ARTICLE IN PRESS

Neurocomputing ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Neurocomputing

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

Stereo matching by using the global edge constraint

Feiyang Cheng [a], Hong Zhang [a,*], Ding Yuan [a], Mingui Sun [b]

[a] Image Processing Center, Beihang University, Beijing 100191, China
[b] Department of Neurosurgery, University of Pittsburgh, PA 15213, USA

ARTICLE INFO

Article history:
Received 26 March 2013
Received in revised form
11 October 2013
Accepted 12 October 2013
Communicated by Huiyu Zhou

Keywords:
Stereo matching
Occlusion handling
Edge constraint
Graph cuts

ABSTRACT

Stereo matching, the key problem in the field of computer vision has long been researched for decades. However, constructing an accurate dense disparity map is still very challenging for both local and global algorithms, especially when dealing with the occlusions and disparity discontinuities. In this paper, by exploring the characteristics of the color edges, a novel constraint named the global edge constraint (*GEC*) is proposed to discriminate the locations of potential occlusions and disparity discontinuities. The initial disparity map is estimated by using a local algorithm, in which the *GEC* could guarantee that the optimal support windows would not cross the occlusions. Then a global optimization framework is adopted to improve the accuracy of the disparity map. The data term of the energy function is constructed by using the reliable correspondences selected from the initial disparity map; and the smooth term incorporates the *GEC* as a soft constraint to handle the disparity discontinuities. Optimal solution can be approximated via existing energy minimization approaches such as Graph cuts used in this paper. Experimental results using the Middlebury Stereo test bed demonstrate the superior performance of the proposed approach.

1. Introduction

Stereo matching, the problem of establishing the correspondences between two stereo pair of images, is one of the central research topics in the field of computer vision. Recently more and more modern applications, such as structure-from-motion and image-based rendering etc., require dense stereo matching result of high accuracy. However, after having been heavily investigated for decades, how to achieve a reliable dense stereo matching result is still a very challenging task.

Traditional techniques on stereo matching could be categorized into two groups: namely local and global algorithms [1]. Local algorithms (also mentioned as window-based algorithms) determine whether two pixels are the projections of the same position in the 3D scene or not by measuring the texture similarities of their local support windows [2,3]. The two candidates will be considered as a match if the matching cost is larger than a specific threshold. Such approaches usually have an acceptable performance when all pixels within a specific local support window are of approximately equal disparities, while they often fail when the support windows cover the disparity discontinuities or occlusions. Some approaches tried to handle the disparity discontinuities and occlusions with emphasis on seeking optimal support windows from several predetermined window models. For instance, *shiftable-window* [4,5] and *adaptive-window* methods [6–8] are two typical strategies among them. However, the predetermined window models have their limitations on adjusting their shape and size arbitrarily, thus the support windows selected from those models are usually not the real locally optimal ones. Color segmentation is also applied to locate the disparity discontinuities [9,10]; nevertheless the color segmentation results would not provide reliable information for disparity discontinuities locating if the captured scene was highly textured. Yoon and Kweon assigned each pixel within the support window an adaptive support-weight (ASW) [11]. These support-weights were determined by checking the color differences and spatial distances from the central pixel. They defined that larger color differences or spatial distances would give lower support-weights, which mean less contribution to the matching cost. Obviously this approach could not tackle the problem of highly textured scenes as well. Hosni et al. proposed an improved algorithm in which the geodesic support-weight was applied instead [12]. In recent years, the ASW strategy has been further studied by using filter-based methods for either quality enhancement or speedup [13–15]. For detailed classification and evaluation of different cost aggregation methods, the readers are referred to four recent reviews [16–19].

Employing the support window of an arbitrary shape and size seemed to be a good solution to handle the disparity discontinuities, however depending on the texture complexity. Actually,

* Correspondence author. Tel.: +86 10 82338991.
E-mail addresses: chfybuaa@gmail.com (F. Cheng), dmrzhang@buaa.edu.cn, dmrzhang@sohu.com (H. Zhang), dyuan@buaa.edu.cn (D. Yuan), drsun@pitt.edu (M. Sun).

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2013.10.022

compared to adjusting the shape of the support window, the size of the window is more convenient to control. Moreover, if we could force the support window not to cover the disparity discontinuities, square windows are eligible to fulfill the task on local support window searching even if when dealing with the highly textured or textureless regions in the image domain. Then the critical problem is how to locate the disparity discontinuities and occlusions in the preprocessing procedure for the task of searching local support windows.

Global algorithms usually treat stereo matching as a pixel-labeling problem. The labels are the representation of disparities and the optimal solution is achieved by minimizing an energy function. In general the energy function consists of two items: the data term penalizes solutions that are inconsistent with the observed data (i.e., pixel dissimilarity); and the smooth term enforces the piecewise smoothing assumption. The optimal solutions could be approximated via the optimization algorithms, such as Graph cuts [20,21] and Believe Propagation (BP) [22–24]. However, traditional global algorithms usually encounter the following problems: the data term has higher costs at the areas where occlusions happen, and the smooth term is assigned to high penalty at disparity discontinuities because of the explicit smoothness assumptions. The Middlebury benchmark [1] shows that almost all the top-rank global algorithms have to incorporate their priors into their global framework to leverage the ill-posed stereo matching problem. Among all the priors involved in the global algorithms, segment-based prior has been widely explored due to its performance of first-rank [25,26]. The color-based segmentation usually performs well at the disparity discontinuities, but often fails if the images are highly textured. Yuri Boykov et al. suggested that the color edges indicate the locations of potential disparity boundaries, and should be involved in the energy function in terms of static cue [20].

Color information, sometimes in terms of edge features, has been employed in enormous stereo matching algorithms. In adaptive support-weight (ASW) methods [11,12] and filtering-based methods [13–15], they often employ the color similarity as a constraint to aggregate the matching costs locally. And in global methods [20,25,26], they employ the color difference as a smoothness constraint based on the assumption that color edges are potential disparity discontinuities. Kun Wang employed the detected edges to search local support windows but treated all the edges as disparity discontinuities [27]. As a matter of fact, only a small portion of the color edges could be related to disparity boundaries. The idea of incorporating all the color edges as a prior into the stereo matching framework would unavoidably introduce numerous wrong cues to locate the disparity discontinuities. Moreover, either the segment-based or the edge-based prior can hardly handle highly textured images. Hence, selecting the reliable edge candidates related to the disparity boundaries from all the color edges is a solution to guarantee that the smoothness constraint could be applied reasonably and effectively for both local support window searching task and global optimization framework. Motivated by the problems of traditional local and global algorithms discussed above, we put forward the idea of categorizing the color edges into different groups and selecting the particular color edges to locate disparity discontinuities. In detail, some particular color edges are selected to define a new constraint named global edge constraint (*GEC*). Compared to the methods we have discussed, we employ the edge features explicitly and selectively.

Our proposed method integrates both local and global algorithms. The *GEC* is employed to estimate the optimal support windows in our local algorithm to establish the initial disparity maps, and then is incorporated as a soft constraint into our global optimization framework. After obtaining the initial matches by using the local algorithm, the reliable matching results which are consistent with the observed data are selected to construct the data term of the energy function. The optimal solution will be approximated by using the α–β-swap algorithm [20] eventually. The main contributions of this paper are as follows: (1) to propose a feasible scheme to categorize the color edges into different groups according to whether they are potentially related to the disparity discontinuities or not; (2) to define and employ the *GEC* in both local and global stereo matching algorithms to generate the disparity maps of high accuracy. Superior performance of our algorithm on the Middlebury Stereo test bed demonstrates that the *GEC* is effective in establishing the accurate and dense stereo matching results.

This paper is an extension of our previous work [28–30] with more discussion of parameter setting and detailed analysis of the experiments. The remainder of this paper is organized as follows. In Section 2, the definitions of different groups of edges and the *GEC* are introduced briefly. And then our edge categorization scheme is presented in Section 3. Section 4 describes the strategy on employing the *GEC* in both local and global algorithms to establish the disparity maps. The experimental results are shown in Section 5 with related discussion. Finally in Section 6 we draw conclusion and put forward the possible future work.

2. Related definitions

Actually, not all of the color edges detected in the stereo pair of images have their correspondences because of occlusions or different viewpoints of the cameras. Hence, we have to discriminate the color edges having their correspondents in the second image from those who have not, as the edges who have no correspondent provide important cues for locating occlusions. In our algorithm, we categorize the detected edges of both the stereo pair images into two groups: the *match-edges* represent edges on which the pixels do have correspondences in the second image, and the *unmatched-edges* represent edges on which the pixels have no correspondence in the second image. Obviously, the color edges which are the potential disparity boundaries are categorized into the group of *match-edges* in the last step. Therefore we further categorize the *match-edges* into two groups: *disc-edges* and *internal-edges*. The *disc-edges* are actually possible candidates of disparity boundaries. The *internal-edges* (i.e., normal color edges) are the edges inside some specific regions, and approximately have constant disparities. It is reasonable to assume that those specific regions only cover small portions of the smooth and continuous surfaces in the scene.

By examining the stereo pair of images thoroughly, it is reasonable to assume that the regions surrounding the *unmatched-edges* could not be matched and a part of these regions would possibly be occluded. Moreover, occlusions would also happen in the areas neighboring the *disc-edges*, as the *disc-edges* indicate where the potential disparity boundaries locate. Thus our *GEC* are proposed to identify the *disc-edges* and the *unmatched-edges*. In other words, the *GEC* are composed of the *disc-edges* and the *unmatched-edges*.

The accuracy of the matching results would be improved by means of selecting the optimal support windows, that is, it must be ensured that the local support windows would not cover disparity boundaries and occluded regions. Therefore we adopt the *GEC* to estimate the optimal support windows, since the *GEC* implies the possible locations of the occlusions and disparity boundaries. The *internal-edges*, defined as the edges which have sufficient intensity variation and lie within specific regions of approximately equal disparities, are more likely to receive their accurate matches by using some existing approaches. Hence the

Please cite this article as: F. Cheng, et al., Stereo matching by using the global edge constraint, Neurocomputing (2013), http://dx.doi.org/10.1016/j.neucom.2013.10.022