J. Vis. Commun. Image R. 39 (2016) 107–119

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Edge-preserving guided filtering based cost aggregation for stereo matching ☆

Shiqiang Zhu, Zhi Wang *, Xuequn Zhang, Yuehua Li

Zhejiang University, State Key Laboratory of Fluid Power and Mechatronic Systems, No. 38, Zheda Road, Hangzhou 310027, China

ARTICLE INFO

Article history:
Received 7 January 2016
Revised 27 April 2016
Accepted 23 May 2016
Available online 24 May 2016

Keywords:
Stereo matching
Adaptive support window
Cost aggregation
Guided filter

ABSTRACT

Stereo matching has been widely used in various computer applications and it is still a challenging problem. In stereo matching, the filter-based stereo matching methods have achieved outstanding performance. A local stereo matching method based on adaptive edge-preserving guided filter is presented in this paper, which can achieve proper cost-volume filtering and keep edges well. We introduce a gradient vector of the enhanced image generated by the proposed filter into the cost computation and the Census transform is adopted in the cost measurement. This cost computation method is robust against radiometric variations and textureless areas. The edge-preserving guided filter approach is proposed to aggregate the cost volume, which further proves the effectiveness of edge-preserving filter for stereo matching. The experiments conducted on Middlebury benchmark and KITTI benchmark demonstrate that the proposed algorithm produces better results compared with other edge-aware filter-based methods.

1. Introduction

Stereo matching, as a challenging problem in the research of computer vision, has been used in many applications, including 3d reconstruction, navigation of autonomous driving system, panoramic stereo imaging [1] and DoF rendering [2]. The main problem is to search for the corresponding pixels in two images. Various stereo matching methods have been presented, which are categorized as global methods and local methods [3]. The global methods usually make the smoothness assumption in the energy function, and then the disparity map is determined by minimizing the global energy. Popular implementations of global methods include dynamic programming [4], graph cut [5] and belief propagation [6]. These global algorithms usually produce poor results on the discontinuity edges and lead to streaking artifacts [6]. Most of the global optimization methods need complex global model, so they are computationally expensive and not feasible in the real-time applications. Local methods are considered to be more straightforward and simple. Compared with global methods, local methods are less time-consuming. Most of the local stereo matching methods perform the following steps: At first, the raw matching cost is computed. And then the cost aggregation is implemented in the local support window. Finally, the optimal disparity map is selected through optimization [3].

1.1. Related works

1.1.1. Cost computation

A variety of cost computation methods have been studied to compute the initial matching cost, including the sum of absolute difference, the Rank transform and the normalized cross correlation. However, these methods make the assumption that each pixel and its corresponding pixel have equal intensity values. Thus, they are not suitable for outdoor images, in which radiometric differences commonly exist [7]. Census transform encodes the local image information, and then the matching cost values are computed through Hamming distance. This method is robust under radiometric distortions [7] and many improvements have been made on it. Hirschmuller [8] proposed a stereo matching algorithm based on the mutual information (MI), which measures the similarity by utilizing the probability distribution function. But the large kernel size in the MI-based stereo matching method leads to poor performance in the object borders.

1.1.2. Cost aggregation

Cost aggregation is an essential part of local stereo matching, on which both the efficiency and accuracy largely depend. The most straightforward method that can be used is the low pass filters

☆ This paper has been recommended for acceptance by M.T. Sun.

* Corresponding author.
E-mail address: 11325067@zju.edu.cn (Z. Wang).

http://dx.doi.org/10.1016/j.jvcir.2016.05.012
1047-3203/

with fixed kernel size, such as box filter, Gaussian filter. However, these methods produce poor results with fatten edges. To solve this problem, many modified cost aggregation strategy have been proposed, such as the variable support window (VSW) methods [9] and adaptive support weight (ASW) methods [10].

Many VSW methods have been introduced in [9]. These methods try to build a support window with optimal size or shape that fits the region. Veksler [11] proposed a variable window approach that determines the support window by minimizing the window cost. The integral image technique is used in this algorithm to reduce the computational complexity of the cost aggregation. But the window size adjustment step is quiet time-consuming. Zhang et al. [12] presented a novel cross-based support window. This support window only consists of horizontal and vertical slices, and the cost aggregation is implemented in two directions independently.

In recent years, the ASW based methods have drawn a great deal of attention due to its outstanding performance [10]. The ASW based methods compute a weight for each pixel. Kuk-Jin and In first introduced the effectiveness of ASW based methods for stereo matching [13]. In this method, bilateral filter (BF) was used to compute the weights through the spatial distance and color dissimilarity between the corresponding pixels. After this study, various weights adjusting functions based on BF were proposed to improve the performance. However, one shortage of the bilateral filter is the high complexity in computation [14]. He et al. [15] proposed a novel algorithm called guided filter (GF) that produces high quality results and outperforms the BF.

1.2. Motivations

In local methods, those edge-preserving filters, such as BF, GF, can produce better results in stereo matching while keeping fine edges at the same time [16]. GF has shown better performance and efficiency compared with BF. But GF simply averages the pixel values without using any weighted fusion. Thus, it cannot preserve the edge well in some cases. In this paper, we incorporate an edge-preserving constraint into GF to improve the performance. Nevertheless, according to the previous research in [17], the raw BF or GF function cannot sort the ambiguity caused by those pixels with similar colors, but located at different disparity areas. A simple implementation of the GF combined with the edge-preserving factor cannot deal with the ambiguity effectively, which will be further explained in Appendix A.

To solve the problem mentioned above, adaptive support window is utilized in our proposed cost aggregation function. In the literature, several adaptive support window methods combined with GF have been presented, such as the adaptive guided filtering [16], the adaptive shape support window (ASSW) guided filtering [18], cross-based local multi-point filtering [19]. Xu et al. [18] proposed the ASSW based guided filter, which can find the optimal window with arbitrary size and shape. Cross-based local multi-point filtering, as an improved GF, determines the optimal support window based on the image information.

After surveying and analyzing the existing stereo matching algorithms, we propose an adaptive edge-preserving guided filtering (AEGF) based stereo matching algorithm. First, an enhanced image guided gradient vector is introduced into the matching cost computation. And the Census transform is adopted in the cost measurement. This measurement keeps robust against radiometric differences and textureless regions. Second, the edge-preserving guided filter (EGF) is adopted as the cost aggregation method. It can perform proper cost filtering and edge preserving, and the adaptive cross-based support window can resolve the ambiguity effectively. Third, the 'Winner-Take-All' strategy is taken to compute the raw disparity values. Finally, a multi-step post-processing method is applied to refine the disparity map. Except for stereo matching, there are other matching tasks, such as flow matching [20] and motion region matching [21]. In this study, we only consider the application to stereo matching. Comprehensive experiments have been conducted on the Middlebury benchmark [22] and KITTI benchmark [23], and the experimental results have shown the effectiveness of the proposed method.

1.3. Organization of the paper

The rest of the paper is organized as follows. Section 2 introduces the pipeline of the proposed method. The edge-preserving guided filter is introduced in Section 3. The combined cost measurement will be presented in Section 4. In Section 5, the proposed adaptive cross-based support window and overall cost aggregation method is described in detail. The multi-step post-processing is adopted in Section 6. The experiment results are given and analyzed in Section 7, and Section 8 summarizes this paper.

2. AEGF-based stereo matching method

The proposed AEGF-based stereo matching method consists of the following 5 steps: (1) preprocessing; (2) cost measurement; (3) matching cost aggregation; (4) initial disparity computation; (5) disparity refinement.

(1) Preprocessing: Due to noise and small texture, the edges of the input images are usually obscure. This will lead to disparity inconsistence and low accuracy. To improve these disadvantages, the edge-preserving guided filter is applied to preprocess the raw input images.
(2) Cost measurement: using the Census transform and gradient to compute the matching cost is proposed in this paper as the combined matching cost function usually performs better than the single method. Moreover, the enhanced image has rich gradient information for cost computation. Thus, in this paper, the combined cost measurement is adopted, which consists of the Census transform and the gradient vector of the initial image and the enhanced image.
(3) Matching cost aggregation: Many stereo matching algorithms employ adaptive support window to achieve better results. In this study, the cross-based window is adopted to construct the support window. As the cost aggregation needs repetitive computation, the cost aggregation method is proposed based on the orthogonal integral image (OII) technique [11] to accelerate this process.
(4) Initial disparity computation: The winner-take-all method is used to determine the initial disparity maps. That is, the disparity candidate which gives the minimum cost is the optimal disparity.
(5) Disparity refinement: The disparity map generated in the above steps contains many invalid matches and occlusions. Thus, a multi-step post-processing method is proposed in this paper. First, the left-to-right-consistency (LRC) check is adopted to determine the unstable and invalid pixels. Once the outliers are detected, they are corrected with the nearest valid pixel of outliers in the vertical and horizontal directions. Cross-based Occweight filtering [24] is used to correct the unstable pixels. Finally, the slanted plane smoothing [25] is used to post process the results, which is efficient and effective.

3. Edge-preserving guided filter

Inspired by GF, and weighted guided filter (WGF) [26], an edge-preserving guided filter is introduced in this section. First, we