J. Vis. Commun. Image R. 33 (2015) 1–9

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Boundary-preserving stereo matching with certain region detection and adaptive disparity adjustment ☆

Cheolkon Jung [a,*], Xufeng Chen [a], Jiji Cai [a], Haozhen Lei [a], Inyong Yun [b], Joongkyu Kim [b]

[a] School of Electronic Engineering, Xidian University, Xi'an 710071, China
[b] College of Information and Communication Engineering, Sungkyunkwan University, 440-746 Suwon, Republic of Korea

ARTICLE INFO

Article history:
Received 17 December 2014
Accepted 18 August 2015
Available online 28 August 2015

Keywords:
Disparity estimation
Image segmentation
Stereo matching
Disparity adjustment
Disparity classification
Boundary-preserving
Certain region detection
Stereoscopic 3D

ABSTRACT

In this paper, we propose boundary-preserving stereo matching using certain region detection and adaptive disparity adjustment. The main contribution of the proposed method lies in the detection and adjustment of incorrect disparities. In the detection process, we detect certain regions in initial disparity maps to distinguish errors in disparity caused by the disparity discontinuity. In the adjustment process, we adjust the disparity map to reduce the adverse effect of errors in disparity using the certain region detection results. Based on the color similarity, spatial distance, and the reliability of the certain and uncertain regions, we set an adaptive support-weight to each pixel for adjusting the accuracy. Experimental results demonstrate that the proposed method produces high-quality disparity maps by successfully preserving disparity discontinuities along the object boundaries.

1. Introduction

With the development of the stereoscopic 3D displays, the disparity map plays an important role in recent years. The disparity map has been widely used in various applications such as free viewpoint television (FTV), 3DTV and depth-image-based rendering (DIBR) [1–3]. The disparity map provides spatial cues to accurately produce the depth information for different objects from different views. Thus, high-quality disparity maps are very important to improve the accuracy of multi-view video coding, multi-view video-plus-depth representation, and layered depth video [4–6]. How to improve the efficiency and accuracy of the disparity map becomes a key research issue in computer vision and multimedia processing.

1.1. Relate work

The purpose of the disparity estimation is to find the correspondence between the two reference stereo images. The stereo matching for disparity estimation generally perform in four steps: matching cost computation, cost aggregation, disparity computation, and disparity refinement [7]. They are mainly categorized in two classes: local stereo matching and global stereo matching. The local methods calculate the disparity value using the correlation between the two reference images with the color information in the support-window. Because it only considers the local information, it achieves high speeds with a low computation cost. However, because of the lack of enough information, it can't ensure the accuracy of the obtained disparity map and preserve the discontinuities of the object boundaries. Global methods estimate the disparity map through minimizing a global energy function consisting of the data term and smooth term. The global methods can achieve a high quality disparity map; however the computation cost is very large. These methods are often quite slow, which are not suitable for interactive applications or dealing with large amounts of high resolution data. Furthermore, the complexity makes the global methods inflexible and difficulty to implement.

A large number of researchers paid much attention in finding a method with the high speed of local method at the same time achieve high accuracy of global method. Kanade and Okutomi presented an adaptive-window method [8]. They evaluated the local variation of intensity to select an appropriate window. However, the windows were limited to rectangles so that they caused depth discontinuities. Veksler [9] developed a new method based on a range of window sizes and shapes. It needed some specific parameters and the shapes of the windows were constrained. Prazdny [10] and Xu et al. [11] presented a stereo matching method based on support-weights. Prazdny assigned weights to neighboring

☆ This paper has been recommended for acceptance by M.T. Sun.
* Corresponding author.
E-mail address: zhengzk@xidian.edu.cn (C. Jung).

http://dx.doi.org/10.1016/j.jvcir.2015.08.010
1047-3203/

pixels, and Xu et al. set weights using radial computation. However, both of the two methods depended on the initial disparity estimation, which should be accurate enough. With the advent of segmentation-based methods [12–19], the accuracy and efficiency of the disparity map have been improved. The segmentation-based methods firstly used the color information to segment the input stereo images, and then estimated the disparity map using a known stereo matching algorithm based on the segmentation result. However, due to the effect of the disparity inconsistency problem, which led more frequently to overly high disparity values near the object boundary. To deal with the unexpected effects of the disparity inconsistency problem, Gerrits and Bekaert [21] presented a method based on the combination of the disparity maps. They warped one of the disparity maps to the other view, and then combined the two disparity maps into the final disparity map.

1.2. Contributions

In this paper, we propose segmentation-based stereo matching using modified disparity adjustment. Fig. 1 illustrates the entire framework of the proposed method. As shown in the figure, the proposed method consists of three main steps: segmentation-based disparity estimation, certain region detection, and adaptive disparity adjustment based on certain regions. Firstly, we estimate two initial disparity maps using segmentation-based disparity estimation. In this first step, we adopt the Winners-Take-All (WTA) strategy to estimate the disparity of each pixel from multiple candidates [20]. Although WTA is computationally fast and easy, there exists the disparity inconsistency problem along the object boundaries which leads to excessively high disparity values in disparity maps and enormously reduces the accuracy of the disparity estimation. Secondly, to address the disparity inconsistency problem, we set the obtained disparity maps as the input images and make pixel-based matching to detect and extract certain regions of the disparity map based on the correspondence between left and right views. Finally, we refine the disparity map using a modified adjustment method through setting an adaptive support-weight for every pixel to adjust the accuracy of the disparity map. The adaptive support-weight is calculated by the photometric and geometric properties of the input stereo images and the reliability factor which reduces negative effects of the incorrect disparities. The proposed method is similar to [12–19] in using segmentation to estimate disparity maps from stereo images. However, our method is different from [12–19] in obtaining two initial disparity maps based on segmentation and then adjusting the disparity map based on certain region detection. Whereas, conventional segmentation-based methods [12–19] have focused on improving cost aggregation based on segmentation to obtain high-quality disparity maps.

The rest of this paper is organized as follow. In Section 2, we explain segmentation-based disparity estimation. In Section 3, we describe how to detect and classify the disparities into uncertain and certain regions. In Section 4, we address a modified disparity adjustment method based on the certain regions. In Section 5, we provide experimental results and draw conclusions in Section 6.

2. Segmentation-based disparity estimation

First, we employ the mean-shift segmentation algorithm [20] to segment the input stereo images. After that, based on the segmentation result, we calculate two initial disparity maps using segmentation-based stereo matching [21]. To avoid over-segmentation effect, we use the segmentation result to decide the pixels in the matching window which are regarded as the outliers. We set a small weighting factor to the outliers instead of ignoring them in calculating the matching cost. Fig. 2 shows the disparity map in *Tsukuba* produced by the high-speed stereo matching algorithm. However, a large number of matching errors still exists in the disparity map. These matching errors will be corrected in the next section.

Denote the matching cost function as follows:

$$E(p,d) = \sum_{q\in W} w_q * |r(x,y) - l(x+d,y)| \qquad (1)$$

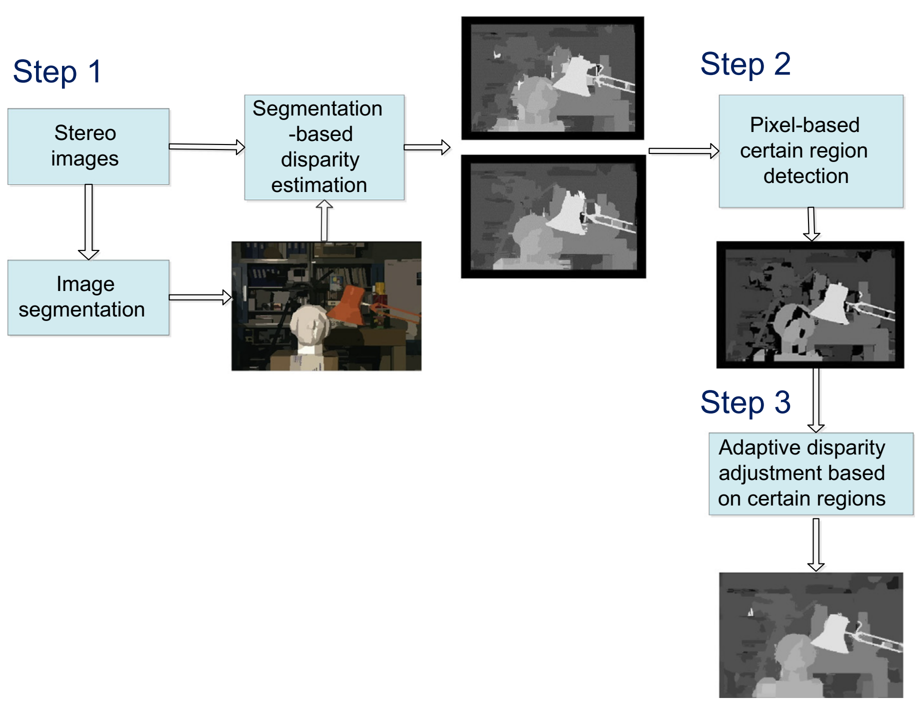

Fig. 1. Entire framework of the proposed method. Step 1: Segmentation-based disparity estimation. Step 2: Pixel-based certain region detection. Step 3: Disparity adjustment based on certain regions.