J. Vis. Commun. Image R. 38 (2016) 487–503

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

# Noise robust image matching using adjacent evaluation census transform and wavelet edge joint bilateral filter in stereo vision ☆

Kechen Song *, Xin Wen, Yongjie Zhao, Zhipeng Dong, Yunhui Yan

*School of Mechanical Engineering & Automation, Northeastern University, Shenyang, Liaoning 110819, China*

A R T I C L E I N F O

*Article history:*
Received 17 February 2016
Revised 28 March 2016
Accepted 29 March 2016
Available online 30 March 2016

*Keywords:*
Stereo matching
Census transform
Random walk
Joint bilateral filter

A B S T R A C T

Automation application systems based on stereo vision require robust image matching methods to achieve available depth image information. This paper presents a novel noise robust stereo matching using adjacent evaluation census transform and wavelet edge joint bilateral filter. The adjacent evaluation census is firstly proposed to improve the robustness against noise of the census transform. Meanwhile, two different and complementary types of metrics are extracted (the adjacent evaluation census mean and the adjacent evaluation census weighted difference). Moreover, the weighted template is composed of four different directions. Then, to improve the robustness of cost aggregation and disparity optimization, the random walk is integrated into the proposed stereo matching method. Additionally, a disparity map post-processing method named wavelet edge joint bilateral filter is employed to eliminate error regions. An obtained wavelet-based edge image is considered as an important weighted coefficient to guide the post-processing. Experimental results demonstrate that the proposed method presents the best performance of the robustness against noise on the Middlebury dataset. Even in the toughest situation with additive Gaussian noise, our method can still achieve the moderate disparity map. In addition, the wider applicability of the proposed method is demonstrated on the KITTI (i.e., Karlsruhe Institute of Technology (KIT) and Toyota Technological Institute at Chicago (TTI-C)) dataset and some typical real-world sequences.

© 2016 Elsevier Inc. All rights reserved.

## 1. Introduction

Three-dimensional (3D) depth image information are required in many engineering applications, such as automatic driving system and robot system. In these systems, a series of stereo images are firstly obtained using a binocular camera, and then these obtained images are analyzed using stereo vision methods including stereo rectification, stereo matching, and 3D reconstruction.

Stereo matching is a key technology in stereo vision, which is one of the most active research topic in computer vision. To achieve available depth image information, a large number of methods for stereo matching have been developed [1–3]. For these methods, there are some fundamental challenges.

The first challenge originates from occlusion. Different from the general definition, i.e., usual occlusion just refers to an object blocked by other objects. However, in stereo vision, the occlusion issue mentioned above is for the case of binocular stereo images, some pixels in an image (e.g., left image) could be seen, while in another image (e.g., right image) could not. The second difficult challenge is from the low textures and repetitive regions, which are typical for structured environments. Since the low textures and repetitive regions have not very rich texture information, a bigger ambiguity will arise in the area for mutual matching points. The third challenge comes from the illumination differences. Due to the different light intensity and environment illumination, the quality of disparity map will be severely affected. Moreover, additional important challenge originates from image noise. Since the image noise inevitable appear in the captured image under environmental interference, pixel continuity of disparity map will be severely affected. In addition, fast calculations are often required because of real-time applications [4].

### 1.1. Motivation

These challenges mentioned above, i.e., occlusion, low textures regions, illumination differences, and image noise, are well known fundamental challenges in stereo matching. As a matter of fact, most of these challenges come from radiometric differences for the input images. Therefore, to solve these challenges, it is

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

* Corresponding author.
*E-mail addresses:* unkechen@gmail.com (K. Song), yanyh@mail.neu.edu.cn (Y. Yan).

http://dx.doi.org/10.1016/j.jvcir.2016.03.026
1047-3203/© 2016 Elsevier Inc. All rights reserved.

necessary to make matching methods robust to radiometric differences. Furthermore, it is safe to say that any real-world stereo application requires radiometric robustness [5].

As we all know, stereo matching methods depend on matching costs for computing the similarity of binocular stereo images. Hence, the performance of matching cost directly affects the final matching results. Hirschmuller and Scharstein [5] evaluated the performance of different costs for passive binocular stereo methods in the presence of simulated and real radiometric differences, including exposure differences, vignetting, varying lighting, and noise. According to the conclusions of the literature [5], census showed the best and most robust overall performance. However, we observe that the census transform is not really very robust to image noise, especially for strong image noise, more detailed description is presented in Section 3. Therefore, improving the robustness against noise of the census transform becomes an important issue for stereo matching. Meanwhile, this issue is also our motivation of this work.

### 1.2. Contributions

In this work, to improve the robustness against noise of the census transform, a novel, simple, yet robust stereo matching method is proposed. The main contributions of this work are summarized as follows:

(1) A novel cost metric against noise named the adjacent evaluation census (AECensus) is proposed for stereo matching. Meanwhile, two different and complementary types of metrics are extracted, i.e., the adjacent evaluation census mean (*AECensus_M*) and the adjacent evaluation census weighted difference (*AECensus_WD*). Moreover, the weighted template is composed of four different directions.
(2) To improve the robustness of cost aggregation and disparity optimization, the random walk is integrated into the proposed stereo matching method. In more detail, entropy rate of random walk segmentation is employed to segment the stereo images in cost aggregation. Then, a modified random walk method is used to accomplish the disparity optimization.
(3) A disparity map post-processing method named wavelet edge joint bilateral filter (WEJBF) [2] is employed to eliminate error regions remain after the cost optimization. A wavelet-based edge detection method is applied to obtain the edge image of the disparity map. Then the obtained edge image is considered as an important weighted coefficient to guide the post-processing.

The rest of this paper is organized as follows: Section 2 summarizes the related works of stereo matching. Section 3 gives the problem definition of this work. Section 4 introduces the proposed adjacent evaluation census in detail. Section 5 presents the random walk for cost aggregation and disparity optimization. Section 6 reveals the disparity map post-processing with WEJBF. Then Section 7 elaborates the experiments and discusses the experimental results. Finally, Section 8 concludes the paper.

## 2. Related works

According to different standards, there are various classification methods for stereo matching. Scharstein and Szeliski [3] presented a taxonomy and categorization scheme to achieve an informed comparison of stereo matching methods. From the perspective of taxonomy, stereo matching methods generally consist of the following four steps: matching cost computation**,** cost aggregation**,** disparity optimization**,** and disparity refinement.

Generally, matching cost computation is the purpose of forming the initial disparity space image (i.e., the initial cost volume). Currently, most matching cost computation methods include parametric costs and nonparametric costs [5]. Common parametric costs use the magnitude of pixel values such as absolute difference (AD), the sum of absolute or squared differences (SAD/SSD), normalized cross correlation (NCC), etc. Common nonparametric costs use the local ordering of intensities and handle all monotonic mappings, such as rank, softrank, census, etc. Besides, there are also other cost metrics including gradient-based metric, mutual information (MI), scale invariant feature transform (SIFT) information, etc. In order to improve the matching accuracy, different cost metrics are combined into one. For instance, combined absolute difference and census (AD-Census) [6], jointed gradient and census [7,8], combined SIFT and mutual information [9], jointed SIFT and census [10], etc.

The goal of the cost aggregation is to update each cost value in initial cost volume based on cost values in its local support regions. Different aggregation approaches differ in the way of selecting the support region, as well as the function used for calculating the new costs [11]. Tombari et al. [12] evaluated the effectiveness and efficiency of the various aggregation methods including shiftable window, reliability, variable windows, etc. As a summary, adaptive weight [13] and segment support [14] presented the best and most robust overall performance. In addition, some new cost aggregation methods are proposed, such as the successive weighted summation (SWS) [15] and guided filter [16,17].

To accomplish the disparity optimization, three different categorization methods are usually employed, i.e., local-based, global-based, and semi-global. The local-based methods mainly focus on cost computation and cost aggregation. Thus, these methods perform a local "winner-take-all" (WTA) optimization at each pixel. In contrast to the local-based methods, global-based methods usually skip the aggregation step. Many global methods are formulated in an energy-minimization framework consisting of a data and a smoothness term. To find an exact optimal solution, several approximation methods have been proposed, such as Graph Cuts (GC) [18], Belief Propagation (BP) [19], color weighted hierarchical BP [20], random walk with restart (RWR) [8], and two-step global optimization (TSGO) [21]. In order to reduce the optimization complexity of the global cost function, semi-global matching (SGM) [4] method has been proposed. In addition to the above methods, there are some other methods including the dynamic programming [22] and inter-regional cooperative optimization [23].

Disparity refinement is an integral part of stereo matching methods and indispensable for good results. Typically, left-to-right disparity map cross-checking is used to detect the occluded areas. Spurious mismatches can be eliminated by a median filter. Moreover, holes due to occlusion can be filled by surface fitting or by distributing neighboring disparity estimates. In addition, sub-pixel accuracy correction is also performed to reduce the errors caused by discrete disparity.

## 3. Problem definition

Since the local image structures being encoded base on the relative ordering of pixel intensities, the census transform [24] has always been considered as the robust metric to radiometric variations and image noise. However, is the census transform really very robust to image noise? Let's do an experiment.

The "Tsukuba" image pairs (i.e., download from the Middlebury dataset) are obtained as experimental object. Firstly, the "Tsukuba" images are corrupted by Gaussian noise with different signal-to-

Download English Version:

https://daneshyari.com/en/article/529687

Download Persian Version:

https://daneshyari.com/article/529687

Daneshyari.com