Pattern Recognition Letters 78 (2016) 41–47

Contents lists available at ScienceDirect

Pattern Recognition Letters

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

Adaptive descriptor-based robust stereo matching under radiometric changes☆

Yong-Ho Kim, Jamin Koo, Sangkeun Lee*

MultiMedia Computing Lab., Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 156-756, Republic of Korea

ARTICLE INFO

Article history:
Received 13 March 2015
Available online 26 April 2016

Keywords:
Stereo matching
Radiometric changes
Descriptor
Illumination
Camera exposure

ABSTRACT

In a real stereo vision system, the acquired stereo images suffer from varying radiometric changes due to illumination and camera parameter changes. Therefore, we propose an effective matching scheme created by building a content adaptive descriptor. Specifically, the descriptor reflects image contents and its element are adaptively weighted and applied to estimate the correct corresponding pixels based on the entropy energy function even under radiometric changes. For the performance evaluation, the proposed scheme is compared with the state-of-art algorithm using Middlebury and KITTI Vision stereo datasets that have radiometric changes. Specifically, 24 of 71 indoor image pairs in the Middlebury and 3 of 7 outdoor pairs are selected, respectively. Experimental result shows that the proposed method reports 6.23% bad pixel matching on average, but it outperforms state-of-the-art algorithms by reducing around 2% bad pixel matching error, which achieves about 16.5% performance improvement.

1. Introduction

Stereo vision is a field of research for estimating 3D information using the corresponding relationship in the stereo images obtained at different geometric positions. The common procedures in generating 3D information consist of image rectification, pre-processing, stereo matching, post-processing and 3D reconstruction [1,8,17]. In this process, stereo matching plays an important role; its matching cost is computed by considering the assumption that corresponding pixels in the same object should have similar color values for each image [1,3,17]. However, in practice, a stereo vision-based system may have some problems in the image acquisition step because it uses multiple cameras. Among them, one major problem comes from the radiometric changes as follows [3,12,13,16]. First, changes in color and exposure levels occur due to inconsistent parameter settings for cameras. Second, an illumination change is not avoidable in accordance with the different paths of a light source. Third, the geometric inconsistency from the different locations of objects and cameras causes the changes in reflectance and shadows. Conventional stereo matching algorithms, which do not consider the aforementioned problems, cannot find correct corresponding points between given images deformed by radiometric changes because it is assumed that corresponding pixels have similar color values, which is called color consistency [6,7,10]. Therefore, to overcome this problem, much research has been performed over a long period of time. Representative algorithms consist of adaptive normalized cross-correlation (ANCC) [11] using color consistency in low-frequency region, a gradient similarity-based adaptive support-weight (GradAdapWgt) [5], and cumulative distributions of gradients (CDFofGrad) [14] using high-frequency information. The ANCC algorithm models a color space under the assumption of Lambertian reflectance and estimates log-chromaticity robust to the global and local radiometric changes. Final matching cost for the estimated log-chromaticity is computed by normalized cross-correlation (NCC). However, ANCC has trouble with images obtained under multiple illumination conditions or cases of non-Lambertian reflectance of objects. In addition, ANCC requires high computational cost because the matching cost is iteratively computed for global optimization. In contrast, the matching cost for gradient-based GradAdapWgt and CDFofGrad algorithms is calculated from the intensity difference between pixels. The GradAdapWgt approach computes the matching cost from the magnitude and orientation of intensity gradients, and the cost for the CDFofGrad algorithm comes from accumulating the cumulative distribution of the gradient information under the assumption that each channel is independent and stereo pairs share similar distribution in high-frequency attributes (including gradient, edge, and valley) even under radiometric changes. The two algorithms finally estimate the matching cost using the adaptive support-weight (ASW) [18]. However, the GradAdapWgt and CDFofGrad al-

☆ This paper has been recommended for acceptance by D. Coeurjolly

* Corresponding author. Tel./fax: +82 2 820 5839/+82 2 814 5404.
E-mail address: sangkny@cau.ac.kr (S. Lee).

http://dx.doi.org/10.1016/j.patrec.2016.04.015

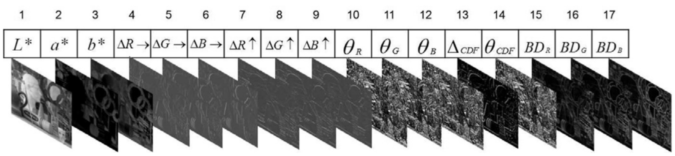

Fig. 1. Construction of local descriptor.

gorithms have two major problems. First, continuous and similar pattern within the search range of the gradient image causes mismatching. Second, wide and low frequency-region also gives incorrect corresponding points. To isolate the above issues, we propose a novel scheme by constructing a multi-dimensional descriptor reflecting sub-band frequency information and by adaptively weighting the local descriptor according to entropy energy-based measurement under various radiometric changes. This paper is organized as follows. In Section 2, local descriptor generation and matching cost computation are presented. Experimental results are given in Section 3. Section 4 then concludes this work with some discussions.

2. Proposed method

2.1. Robust stereo matching under radiometric changes

In general, stereo matching method determines corresponding points based on color consistency, but it is not easy under radiometric changes. To overcome this problem, researches have been performed to find correct corresponding points under radiometric changes. Existing methods [5,11,14] estimated a color formation model or used insensitive information under radiometric changes such as gradient. However, estimating accurate color formation model is difficult and it requires high computations. Furthermore, determining correct corresponding points is hard when the pattern of local gradient is similar or its value is extremely small. In addition, the existing approaches have limitations since they tried to compute the matching costs without considering combination of local features. To isolate the above issue and to obtain robust corresponding points under radiometric changes, local attributes need to be considered. In the proposed scheme, to determine the matching cost efficiently while reflecting the local attributes, we extract multiple local feature information to construct a descriptor and apply an adaptive weighting function according to the corresponding local attributes.

2.2. Local descriptor construction

Many of the existing algorithms only use a single feature such as low- or high-frequency information and cause an erroneous matching cost. In this work, a novel descriptor-based approach is presented for robust stereo matching under various radiometric changes. The descriptor is generated from image content attributes due to the sub-band decomposition. In particular, the attributes consist of pixel value, gradients, and textural properties from the block difference of inverse probabilities (BDIP) [4]. The descriptor vector consists of 17 elements as shown in Fig. 1. The first three represents the CIE Lab color values, which can be suitable for indicating color difference in the low-frequency sub-band. However, the descriptor consisting of the intensity value itself is not enough to represent local image content, and thus it needs to include high-frequency components as well. This fact is reflected in the rest of elements in the descriptor. High-frequency components are extracted from the gradient and texture information. There is a high probability that an intensity value at a pixel location is the same as or similar to its neighbors. In other words, the degree of brightness

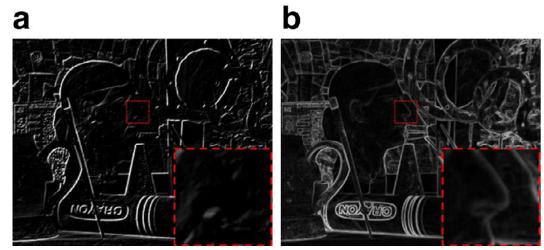

Fig. 2. Local features: (a) gradient, and (b) texture using BDIP operator.

between a reference and its neighboring pixels under radiometric changes is similarly biased. The elements from fourth through ninth in the descriptor come from the horizontal and vertical gradients of an input pair in RGB color space as

$$\Delta c_{\rightarrow_{c\in\{r,g,b\}}} = I_c(x-1,y) - I_c(x+1,y) \quad (1)$$

$$\Delta c_{\uparrow_{c\in\{r,g,b\}}} = I_c(x,y-1) - I_c(x,y+1) \quad (2)$$

Where, $\Delta c_{\rightarrow}$ and $\Delta c_{\uparrow}$ are vertical and horizontal gradients, respectively. I_c is an input image in color channel c. Similarly, the tenth through twelfth elements $\theta_{c\in\{r,g,b\}}$ are edge directions for each channel.

$$\theta_{c\in\{r,g,b\}} = \arctan\left(\frac{\Delta c_{\uparrow}}{\Delta c_{\rightarrow}}\right) \quad (3)$$

In the real world, intensity value in the stereo image pair, which has radiometric change from illumination or color temperature changes, changes independently for each channel. Consequently, stereo images with radiometric change contradict the color consistency because their color is distorted during acquisition. To overcome this phenomenon, more robust means of accumulating the gradient and direction are applied to the thirteenth and fourteenth features, respectively, in the proposed descriptor as

$$\Delta_{CDF} = \sum_{c\in\{r,g,b\}} (\Delta c_{\rightarrow} + \Delta c_{\uparrow}) \quad (4)$$

$$\theta_{CDF} = \sum_{c\in\{r,g,b\}} (\theta_c) \quad (5)$$

Important edges such as shape boundary, in which the intensity difference is not sufficiently large to properly acquire its gradient, need special care for better stereo matching performance.

To handle this case, BDIP (block difference of inverse probabilities) operator [4] is adopted as

$$BD_{c\in\{r,g,b\}} = N_K - \frac{\sum_{q\in K} I_c(q)}{\max_{q\in K} I_c(q)} \quad (6)$$

Where, N_K is the total pixel count in a block K which is centered at pixel q. Fig. 2 shows the results of image gradients without and with BDIP operator. BDIP operator effectively extracts texture information including edges, valleys, and even smoothness from local brightness change that is normalized by the local maximum intensity value. Therefore, it can reveal and boost the texture information (see Fig. 2(b)) from invisible gradient in Fig. 2(a). The improved texture attributes go to the fifteenth through seventeenth elements in the proposed descriptor.

2.3. Entropy-based adaptive weighting scheme

Entropy is a mathematical tool for uncertainty measurement and defined by

$$DE = -\sum_{i=0}^{255} h(i)\log h(i) \quad (7)$$