Optik 160 (2018) 1–12

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

Original research article

Dynamic programming-based dense stereo matching improvement using an efficient search space reduction technique

Behzad Salehian[a], Ali M. Fotouhi[b,*], Abolghasem A. Raie[a]

[a] *Mobile Robots Research Lab, Electrical Engineering Department, Amirkabir University of Technology, Tehran, Iran*
[b] *Electrical Engineering Department, Tafresh University, Tafresh 39518 79611, Iran*

ARTICLE INFO

Article history:
Received 18 November 2016
Received in revised form 7 January 2018
Accepted 10 January 2018

Keywords:
Stereo matching
Dynamic programming
Disparity map
Radial windows
Reduced search space

ABSTRACT

Two most important evaluation parameters of any stereo matching algorithm are the execution speed and the accuracy of generated results. Commonly, reported real-time works provide fast depth calculation based on highly specialized hardware. Thus, algorithmic techniques for increasing the speed which can be employed on every general hardware, are more valuable. In this paper an efficient technique has been proposed in order to improve dynamic programming (dp) based dense stereo matching. The main idea is using a fast and efficient search space reduction algorithm and successively forcing dp based matching process to operate on this reduced search space (RSS). Effective ideas have been proposed to correct RSS and emphasize on more qualified candidates in RSS to facilitate and improve matching process in this space. Experimental results on standard stereo images show significant increasing of execution speed as well as error reduction for two traditional and basic versions of dp based matching algorithms. These simultaneous improvements in both speed and accuracy, considering their usual inverse relationship in other algorithms, seems to be very valuable.

1. Introduction

Stereovision is an important and growing research field in computer vision with a wide range of applications like vision-based robot navigation and object tracking using depth information. The goal of stereo vision is to determine the depth of the scene or extracting its 3D information using two or more input images captured from different viewpoints. In more usual form of stereo vision algorithms called two-frame ones (binocular stereo), only two rectified left and right images are needed [1]. These algorithms usually suppose the left image as the reference one and compute the depth of the scene by finding matching pixels between the reference and target images. The most important part of stereo vision is this matching step. Usually, algorithms consider the epipolar constraint means corresponding (matching) pixels in two images are in the same row. Studying more precisely, if pixel P in the left image matches to pixel Q in the right image, then the disparity of pixel P will be defined as $d_P = x_P - x_Q$ (see Fig. 1). For each pixel in the left image there are some pixels in the right image as candidates of correspondence. It is more convenient to use the equivalent disparities instead of direct referring to those

* Corresponding author.
E-mail address: fotouhi@tafreshu.ac.ir (A.M. Fotouhi).

https://doi.org/10.1016/j.ijleo.2018.01.021

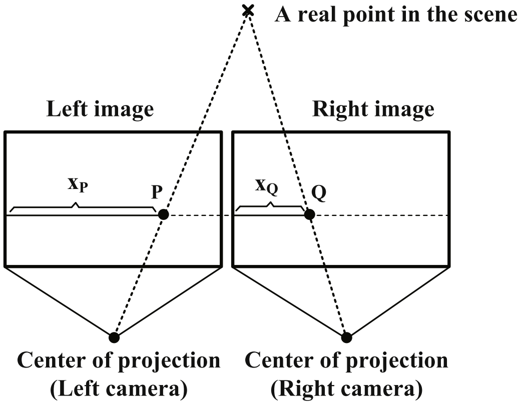

Fig. 1. Corresponding pixels in the left and right images.

candidate pixels. Dense stereo matching algorithms search between the candidate disparities of each pixel and select one of them as its final disparity. According to [1], stereo matching algorithms generally perform a subset of these four steps:

1. Matching cost computation: calculation of the cost of assigning a special disparity to each pixel.
2. Cost aggregation: using information of the neighboring pixels.
3. Disparity computation: assigning the best disparity to each pixel.
4. Disparity map refinement: making some corrections on the estimated disparity map.

The proposed algorithms to find the matching pixels are divided into two main categories according to the strategy of search for correspondence (step 3): local and global [1]. Local algorithms [2,3] search for the most similar pixel in the right image with the aim of intensity information of the neighboring pixels, but global ones define an energy function and minimize it by assigning the best disparity to each pixel using some optimization methods like graph cut [4,5], and belief propagation [6,7].

Two most important evaluation parameters of any stereo matching algorithm are the execution speed and the accuracy of generated results. All stereo algorithms are preferred to have a high accurate and real-time performance. Local approaches have high execution speed but suffer from low accurate results; on the other hand, global ones can provide better accuracy but low execution speed makes them unsuitable in fast applications. A different class of global optimization algorithms are those based on dynamic programming [8–11]. These methods which are also known as semi global algorithms, solve matching optimization problem across corresponding scanlines and have more acceptable performance. They generally are faster than global methods and more accurate than local ones. Commonly, reported real-time accurate works provide fast depth calculation based on highly specialized hardware [12–14]. Thus, algorithmic techniques for increasing the speed (by decreasing complexity) which can be employed on every general hardware, are more valuable.

In this paper, the key parameter which we have focused on, is the speed of dynamic programming based stereo matching algorithms. The idea is utilizing an early efficient search space reduction technique to decrease complexity of these algorithms. This search space reduction method has been introduced in [15]. Considerable decrease in the average number of candidates per pixel, low computational cost and high assurance of retaining the correct answer are its three major and important characteristics. Such an output from the first stage, while speeding up the final selection of disparity in the second stage due to search space reduction, is also promising a more accurate result due to having more reliable candidates. Applying dynamic programming on this reduced disparity candidate space and obtaining improvement in the accuracy of matching process, need to correct and improve this space, which are also done in this paper.

To provide a baseline for comparisons, two basic, standard and well-known dynamic programming based algorithms called dynamic programming (DP) and scanline optimization (SO) which are the best representative of two separable classes of dense stereo matching algorithms based on dynamic programming, are selected. DP and SO have been introduced and implemented in [1]. In SO, global minimum across each scanline is computed using dynamic programming; however, unlike DP algorithm, the ordering constraint does not need to be enforced, and no occlusion cost parameter is necessary. These selections can also give a generality to our proposed technique and its evaluation results. In other words, this paper can propose a complete and effective algorithmic technique to improve two important parameters, speed and accuracy, of other more complicated dynamic programming based algorithms which have a base similar to DP or SO.

The rest of the paper is organized as follows. Section 2 presents a brief review of the related literature. Sections 3 and 4 provide general overview on DP algorithm and the search space reduction technique of [15], respectively. The method of applying DP to the reduced search space and the proposed algorithm for its correction and refinement are explained in Section 5. Description of SO and its improvement are presented in Section 6. Section 7 compares the results before and after applying algorithms to the reduced search space and finally, Section 8 makes conclusion and final discussion.