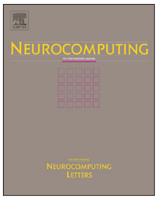
Neurocomputing 194 (2016) 24–33

Contents lists available at ScienceDirect

Neurocomputing

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



Fast spatio-temporal stereo matching for advanced driver assistance systems

Ilyas El Jaafari *, Mohamed El Ansari, Lahcen Koutti, Abdenbi Mazoul, Ayoub Ellahyani

[a] *LabSIV, Department of Computer Science, Faculty of Science, Ibn Zohr University, BP 8106, 80000 Agadir, Morocco*

ARTICLE INFO

Article history:
Received 25 June 2015
Received in revised form
20 December 2015
Accepted 12 February 2016
Communicated by Liang Lin
Available online 23 February 2016

Keywords:
Stereo vision
Edge matching
Disparity map
Intelligent vehicle
Dynamic scene

ABSTRACT

This paper presents a new fast spatio-temporal stereo matching approach devoted to Advanced Driver Assistance Systems (ADAS). The proposed method uses dynamic programming (DP) to match edge points of stereo images. The main idea of the paper consists in involving the matching results obtained at a frame in the computation of the disparity map at its following one. At current frame, we propose to compute local and global disparity ranges based on the disparity map of its preceding one. A global disparity range is computed for each image line based on the v-disparity of the preceding frame. If an edge point in the left image at the current frame can be associated with another one in the left image at the preceding frame, a local disparity range is searched for this edge point. Using these disparity ranges in the DP algorithm allows the reduction of the search space as well as the mismatches and speeds up the matching process. The proposed approach has been tested on both real and virtual sequences and the results are satisfactory.

1. Introduction

Advanced Driver Assistance Systems aim at enhancing car safety and driving comfort. The objectives of ADAS include object detection and tracking [1–5], traffic sign detection and recognition [6,7], pedestrian detection [8] and so on. One of the most important difficulties that ADAS face is the perception of the environment of the vehicles in real outdoor scenes. Stereo vision is a well known method that allows having accurate and detailed 3D representation of the environment around the IV. The key problem in stereo vision consists in finding correspondence between pixels of stereo images taken from different viewpoints [9]. Exhaustive surveys on the methods tackling the correspondence problem are available in [10–12]. An updated taxonomy of dense stereo correspondence algorithms together with a testbed for quantitative evaluation of stereo algorithms is provided by Scharstein and Szeliski [13]. It is demonstrated from [13] that graph cuts' methods [14–17] produce good results. However, they are time consuming which make them not suitable for real-time applications, e.g. IV applications.

The present work is devoted to ADAS, where stereo sensor is embedded aboard an IV. Both the IV and objects of the scene, e.g. cars and pedestrians, are moving. Therefore, the stereo approach we propose should deal with dynamic scenes. The stereo sensor provides stereo images at each time depending on its fps. The matching algorithm will be applied to each stereo images acquired at each time. Incorporating temporal information in stereo approaches can improve the results of the matching as mentioned in [3,18–20]. Nevertheless, only a small amount of research has been devoted to the reconstruction of the dynamic scenes from stereo sequences. We believe that by considering the consistency between consecutive frames the results of the stereo matching will be considerably improved [21–25].

This paper presents a new fast stereo matching approach based on both spatial and temporal information. The proposed method is an improvement of that early presented in [24], where global disparity ranges were computed at the image lines. Instead of using global disparity range for the whole image line, we propose a new disparity range computation depending on the nature of the edge points. We keep using global disparity ranges for edge points of the current frame that could not be associated with any edge point of the preceding frame. However, local disparity ranges are computed for the edge points in the current frame that can be associated with others in the preceding frame.

The rest of the paper is organized as follows. Section 2 highlights state-of-the-art stereo methods that use temporal information. The new proposed method is detailed in Section 3. Experimental results and comparisons are presented in Section 4. Finally, Section 5 concludes the paper.

---

* Corresponding author.
E-mail addresses: eljaafari.ilyas@gmail.com, ilyas.eljaafari@edu.uiz.ac.ma (I. El Jaafari), m.elansari@uiz.ac.ma, melansari@gmail.com (M. El Ansari), l.koutti@uiz.ac.ma (L. Koutti), a.Mazoul@uiz.ac.ma (A. Mazoul), ayoub.ellahyani@gmail.com (A. Ellahyani).

http://dx.doi.org/10.1016/j.neucom.2016.02.010

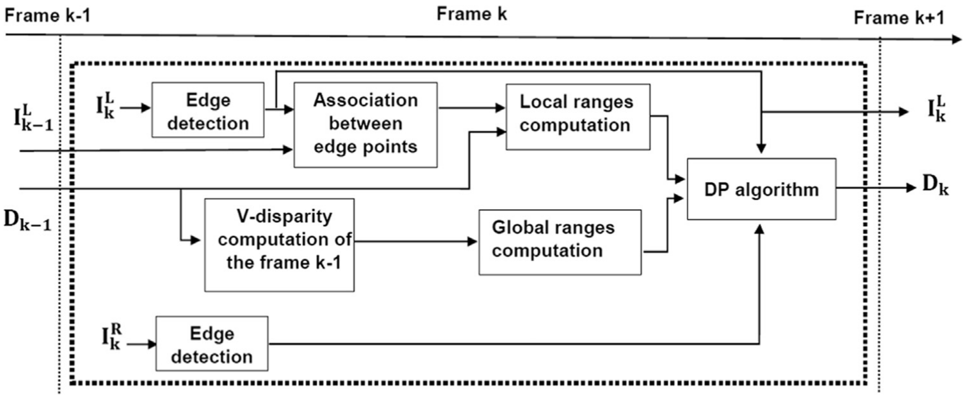


Fig. 1. Scheme of the proposed algorithm.

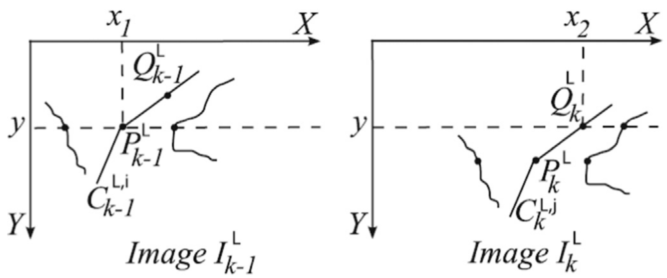


Fig. 2. I^L_{k-1} and $I_k{}^L$ two consecutive left images. The point P^L_{k-1} in image I^L_{k-1} is the associate of the point $Q_k{}^L$ in the image $I_k{}^L$.

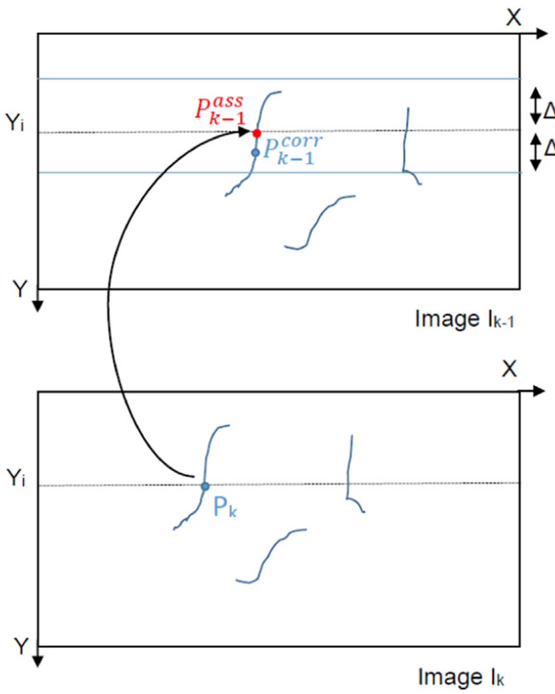


Fig. 3. Relationship between the edge point P_k of the image I_k and the ones P^{ass}_{k-1} (associate) and P^{corr}_{k-1} (corresponding) in the image I_{k-1}.

2. Related work

In order to improve the results of the stereo matching, temporal consistency has been used where different types of methods have been developed, e.g. optical flow, spatio-temporal and disparity prediction [18–20,26].

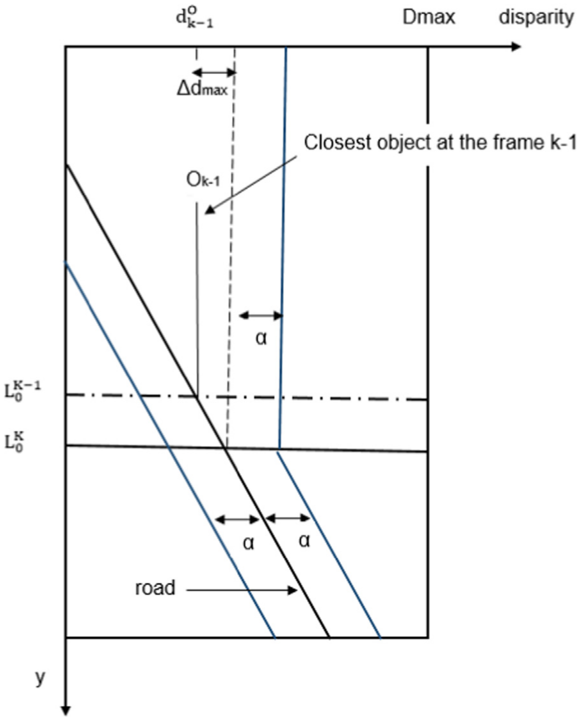


Fig. 4. The global range computed from the v-disparity. The vertical axis refers to the image lines and the horizontal one refers to the disparity values. The possible disparities are in the region between the blue lines. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this paper.)

Toa et al. [18] proposed a method for extracting depth information for non-rigid dynamic 3D scenes, in which the scene is represented by a collection of 3D piecewise planar surface patches based on color segmentation of input images. This representation is estimated by an incremental formulation. The spatial match measurement and the scene flow constraint [27,28] are employed in the matching process. The algorithm execution time and the accuracy of the results are limited by the image segmentation algorithm used. Hung et al. [29] proposed a depth and image scene flow estimation method to preserve motion-depth temporal consistency. Zhang et al. [28] proposed a 3D scene flow computation method, in which the 3D motion model is fit to each local image region and adaptive global smoothness regularization is applied to the whole image.

Zhang et al. [19] use spatial and temporal information by extending the spatial window to a spatiotemporal window, spatial window is used to compute the sum of squared difference (SSD)