Available online at www.sciencedirect.com
ScienceDirect
Image and Vision Computing 26 (2008) 1371–1383
www.elsevier.com/locate/imavis

Iterated dynamic programming and quadtree subregioning for fast stereo matching

Carlos Leung [a,*,1,2], Ben Appleton [a,1,3], Changming Sun [b]

[a] *Electromagnetics and Imaging, ITEE, The University of Queensland, Brisbane, Qld 4072, Australia*
[b] *CSIRO Mathematical and Information Sciences, Locked Bag 17, North Ryde, NSW 1670, Australia*

Received 18 February 2005; received in revised form 12 November 2007; accepted 23 November 2007

Abstract

The application of energy minimisation methods for stereo matching has been demonstrated to produce high quality disparity maps. However, the majority of these methods are known to be computationally expensive requiring minutes of computation. In this paper, we propose a fast minimisation scheme that produces high quality stereo reconstructions for significantly reduced running time, requiring only a few seconds of computation. The minimisation scheme is carried out using our iterated dynamic programming algorithm, which iterates over entire rows and columns for fast stereo matching. A quadtree subregioning process is also used for efficient computation of a matching cost volume where iterated dynamic programming operates on.

Keywords: Stereo matching; Energy minimisation; Iterated dynamic programming; Quadtree subregioning

1. Introduction

The study of computational stereo has undergone intensive research since its inception in the 1970s. Stereo matching is the main step for the recovery of the 3D structure of the scene given a pair of images. By matching primitives such as points, curves and regions between the stereo pair, such that the matched primitives are projections of the same 3D identity in the scene, a disparity map of the scene can be computed.

While simple local correspondence methods have the advantage of low computational complexity they suffer from high sensitivity to matching ambiguity and the choice of matching metric. Reconstructions based on simple local correspondence can be improved by incorporating global constraints and structural information into the stereo matching process. One such class of global correspondence methods are those based on dynamic programming (DP). Since the dynamic programming framework allows efficient optimal solutions, it has been applied to locate the path of minimum matching cost for each scanline of the image [1–3]. However, since DP is typically applied independently to each scanline, methods that employ this technique suffer from interscanline inconsistencies. Several studies have addressed this issue by applying postprocessing to iteratively improve the reconstruction, enforcing interscanline constraints. These techniques include minimising the number of horizontal and vertical discontinuities [4], estimating vertical slopes [5], and using edge maps [6]. These methods attempt to retain the computational benefits of a dynamic programming formulation while avoiding the problem of horizontal streaking. However, while these heuristics improve interscanline consistencies they do not entirely solve the problem.

* Corresponding author. Tel.: +61 7 3836 1606.
E-mail addresses: carlos.leung@gmail.com (C. Leung), appleton@google.com (B. Appleton), changming.sun@csiro.au (C. Sun).
[1] Carlos Leung and Ben Appleton are supported by the Australian Postgraduate Award and CSIRO Mathematical and Information Sciences.
[2] Present address: Suncorp, P.O. Box 1453, Brisbane, Qld 4001, Australia.
[3] Present address: Google Inc., 201 Sussex Street, Sydney, NSW 2000, Australia.

0262-8856/$ - see front matter © 2008 Published by Elsevier B.V.
doi:10.1016/j.imavis.2007.11.013

Another class of global correspondence techniques are those that formulate the stereo matching problem into a two-dimensional energy minimisation framework. By designing an energy functional whose minima will correspond to good stereo reconstructions, the aim is to compute a disparity function that minimises this energy. Geman and Geman [7] applied simulated annealing to minimise the energy function. Sun [8] proposed a two-stage dynamic programming technique to compute disparity surfaces of maximum total correlation. In recent years algorithms based on graph cuts and iterated graph cuts have been proposed to solve the optimisation problem [9–14]. Graph cut methods produce excellent results at the cost of orders of magnitude greater computation than dynamic programming techniques.

In many stereo matching algorithms, there is a need to evaluate the similarity or other metric values for matching points. Metric values evaluated for all overlapping regions of interest between the stereo pair defined by the window size and disparity range can be constructed into a matching cost volume. For fast stereo matching, efficient computation of a matching cost volume is essential. Faugeras et al. [15] developed a recursive technique which is invariant to the size of the correlation window to calculate correlation coefficients. Sun [16] applied the box-filtering technique to achieve fast cross correlation computations. Efficient algorithms have previously been proposed to compute the full matching cost volume by processing the entire image simultaneously.

Sun [8] proposed a rectangular subregioning algorithm in order to reduce the computation cost when constructing the matching cost volume. By subregioning the images into rectangular regions optimised for minimal computational load, the reevaluation of the banded cost volume at each finer scale can be computed efficiently and quickly. However, there are situations where the rectangular subregioning process is not optimal.

In this paper, we present two new techniques for fast stereo matching. We propose an iterated dynamic programming (IDP) algorithm which minimises an energy function that incorporates both intrascanline and interscanline regularity. Although our energy minimisation method is also based on DP, the use of a multi-directional smoothing energy prevents streaking. This framework overcomes the interscanline inconsistency problem inherent to DP techniques and produces comparable results to existing energy minimisation algorithms. Hence, we propose IDP as a new alternative for minimising general energy functions of the form to be described in Section 2. We explain theoretically and demonstrate with results that the proposed algorithm is efficient and a competitive energy minimisation scheme. We also propose a quadtree subregioning (QSR) algorithm that segments stereo images into subregions for the fast computation of the matching cost volume in a multiscale framework. We improve on previous techniques and present a quadtree partitioning scheme that efficiently evaluates a banded cost volume. In Section 4, we present timings and results to demonstrate the quality and speed of the proposed techniques. Initial work by the same authors was presented in [17].

2. Energy function

The goal of dense two-view 3D reconstruction is to recover the depth of each pixel from a stereo image pair. The stereo pair in this paper is assumed to be rectified such that corresponding horizontal scanlines in the two images lie in the same epipolar plane. A *disparity function* $d(\vec{x})$ represents the horizontal displacement for each point $\vec{x}$ of the reference image and is related to the depth of that point in the scene. We formulate the stereo problem into an energy minimisation framework, such that good reconstructions correspond to minima of the energy function. To each disparity function $d(\cdot)$ we associate an energy $E[d]$ quantifying the matching quality. We minimise an energy function that includes terms for data fidelity and regularisation:

$$E[d] = \sum_{\vec{x}} c(\vec{x}, d(\vec{x})) + \sum_{\vec{x}_1 \sim \vec{x}_2} e(|\, d(\vec{x}_1) - d(\vec{x}_2)\,|). \tag{1}$$

The first term in Eq. (1) accounts for the matching cost $c(\cdot)$ of pixel correspondences. Many matching metrics have been proposed in the literature, either measuring similarities or dissimilarities between corresponding primitives [18]. Stereo reconstruction based solely on matching criteria however is an ill-posed problem which has many solutions. The second term of Eq. (1) imposes the assumption of regularity onto the disparity function to obtain solutions which are considered likely from prior knowledge. Here the edge function $e(\cdot)$ is selected to penalise discontinuities in $d(\cdot)$ and $\sim$ is the neighbourhood relation between points. A variety of edge functions have been proposed in the literature, including quadratic functions, discontinuity-preserving functions, and terms dependent on intensity differences [19]. Discontinuity-preserving edge functions have the property of giving bounded penalties to very large disparity jumps. Minimising such functions therefore allows large jumps in the disparity estimate at object boundaries, thus avoiding the oversmoothing which is common to other energy functions.

Although discontinuity-preserving energy functions are desirable in stereo reconstruction, Boykov et al. showed that the minimisation of such an energy function is NP-hard by analogy to the Potts model [10]. They proposed a multiway graph cut framework which can compute a strong local minimum of such energy functions. Kolmogorov and Zabih [12] extended their graph cut framework and minimised an energy function with an extra visibility term in order to model occlusions. In this paper we do not include a visibility term and just minimise the energy function described in Eq. (1). While these optimisation schemes rely on an iterated application of minimum cuts, we propose a fast alternative using IDP.