Available online at www.sciencedirect.com
SCIENCE DIRECT®
Image and Vision Computing 22 (2004) 795–806
image AND vision COMPUTING
www.elsevier.com/locate/imavis

Dense stereo using pivoted dynamic programming

P.H.S. Torr[a,*], A. Criminisi[b]

[a]Department of Computing, Oxford Brookes University, Oxford OX33 1HX, UK
[b]Microsoft Research Ltd, 7 JJ Thompson Avenue, Cambridge CB3 0FB, UK

Received 12 March 2003; received in revised form 13 January 2004; accepted 12 February 2004

Abstract

This paper describes an improvement to the dynamic programming approach for dense stereo. Traditionally dense stereo algorithms proceed independently for each pair of epipolar lines, and then a further step is used to smooth the estimated disparities between the epipolar lines. This typically results in a streaky disparity map along depth discontinuities. In order to overcome this problem the information from corner and edge matching algorithms are exploited. Indeed we present a unified dynamic programming/statistical framework that allows the incorporation of any partial knowledge about disparities, such as matched features and known surfaces within the scene. The result is a fully automatic dense stereo system with a faster run time and greater accuracy than the standard dynamic programming method.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Dynamic programming; Dense stereo; Epipolar lines

1. Introduction

Automatically generating 3D models from images is an on going topic of research. A successful image to model system has been developed by the research groups at Oxford [2] and Leuven [11]. The method proceeds as a set of independent modules: first features are extracted and matched, second projection matrices and calibration recovered, third a dense stereo algorithm based on dynamic programming is used to extract depths and finally a 3D model is constructed. It can be seen that this process involves two representations, one sparse and feature based, the other a dense depth map yielded by dynamic programming. Within this paper the dynamic programming algorithm for recovering the dense depth map is discussed and various improvements to its speed and accuracy are advocated, which utilize the results already obtained by the sparse feature matcher. It will be seen that this unique synthesis of the sparse and dense matching techniques leads to improved results and run times.

The problem of obtaining dense correspondence along pairs of corresponding epipolar lines may be solved using dynamic programming as an optimal path finding problem on a 2D plane. However, there is no corresponding efficient method to impose a smoothness constraint between the epipolar lines. The solution along consecutive epipolar lines can vary significantly, creating artifacts across depth discontinuities and in homogeneous patches of intensity. An example of this is shown in Fig. 1. It can be seen that indeed the estimated disparity along the depth discontinuities is poor, despite being estimated by a method that attempts to enforce inter epipolar line consistency [5] (the epipolar lines in this case run roughly horizontally).

There have been a variety of proposals to solve the 'streaky' artifact. Henderson [9] was the first to use dynamic programming to solve the structure from motion problem. He proposed a sequential approach to impose inter epipolar line constraints, starting at the top of the image and proceeding down through the epipolar lines using the result of the previous as a guide for the next. This method, however, suffers from an avalanche effect in the errors, in that small errors made early on can be magnified as the algorithm progresses.

Baker and Binford [1] popularized dynamic programming for stereo, explaining it in terms of the Viterbi algorithm [6]. First using Viterbi to match edges, and then a second round of Viterbi to match pixels between edges. However, their method does not adequately deal with occlusion and cannot recover if the edges are mismatched.

* Corresponding author.
E-mail addresses: philiptorr@brookes.ac.uk (P.H.S. Torr), antcrimr@microsoft.com (A. Criminisi).

0262-8856/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2004.02.012

Fig. 1. (a,b) The left and right images of the Pentagon standard test stereo pair obtained from the CMU VASC database (http://www.vasc.ri.cmu.edu/idb/). (c) the disparity map generated by the Cox MLMH + V algorithm, typical of the dynamic programming approach [5] (lighter shades indicate larger disparity). What this disparity map does not show is those pixels that are unmatched (occlusion), this will be discussed in more detail in Section 5. Note the lack of consistency between epipolar lines, which run horizontally.

A slightly different idea is put forward for edgel matching by Ohta and Kanade [13] who propose performing dynamic programming to solve a path planning problem in the product space of the epipolar lines. To impose consistency between epipolar lines they extend the dynamic programming to 3D. However, this furnished only a sparse representation in terms of edges.

A problem with the above methods is that they do not deal with occlusion very well, and have no mechanism for detecting occluded pixels. A series of methods [3,5,7] that were developed roughly concurrently are all characterized by the modelling of an occlusion process together with a shift to estimate a per pixel disparity (as opposed to matching edges first as in Ref. [1,13]). Cox et al. [5] propose a purely maximum likelihood approach, whereas Belhumeur [3] and Geiger et al. [7] prescribe a Bayesian philosophy. The other difference is algorithmic, all the methods can be cast as one of finding the best path through a graph, but the structure of the graphs are different. However, the methods all operate on each pair of epipolar lines independently and rely on further iterative steps to enforce smoothness constraints between epipolar lines.

Recently a new class of methods based on maximum flow/minimum cut algorithms has promised to generalize the dynamic programming approach to incorporate inter epipolar line constraints. Roy and Cox [14] introduced a maximum flow algorithm on an undirected graph for stereo, however, as pointed out in Ref. [10] their method does not really generalize dynamic programming and does not model discontinuities and occlusions as all pixels are forced to have a match. Ishikawa and Geiger [10] propose a max flow algorithm on a directed graph with the aim of imposing constraints between epipolar lines, this approach seems interesting but there is still evidence of the streaky effect at depth discontinuities suggesting that the algorithm does not adequately enforce inter epipolar line constraints. These algorithms are also very computationally intensive when compared with the pair-wise epipolar line solutions.

None of these papers exploit prior information that might already have been gathered. Within this paper we explore a class of methods that combines the sparse but accurate representations yielded by feature detectors and matchers with the dense representation yielded by dynamic programming.

This paper is laid out as follows: Section 2 explains the dynamic programming approach to dense stereo. Section 3 briefly describes the feature extraction and matching algorithms. It also suggests an improvement on the edge matching by making use of the results of the corner matching. Section 4 shows how the corner and edge information can be readily incorporated into the dynamic programming framework. A comparison of the disparity map computations for our methods and for that of Cox et al. [5] is given in Section 5.

2. Dynamic programming to solve stereo correspondence

Within this section the dynamic programming approach is described in detail, and it is observed that the difference between the various methods that have been previously proposed lies in the structure of the graph on which the path planning is performed. We then lay out a statistical framework so that additional constraints are easily incorporated (in Section 4).

All of the dynamic programming approaches to dense stereo described in the introduction can be thought of in terms of a path planning problem on a graph. There are two types of graph that are typically constructed differing in their connectivity. The first (espoused in Refs. [5,7,13]) is a graph formed on the product space of the two epipolar lines as shown in Fig. 2, the second (advocated in Refs. [1,3,9]) is a graph formed on the product space of one of the epipolar lines and the set of putative disparities. The two approaches lead to similar results and the differences are small, in this paper we explore the first in some detail and show how it can be improved, but the improvements are generic to both methods. The path defines a mapping between the two epipolar lines for each corresponding pairs of points

دریافت فوری ← متن کامل مقاله

ISIArticles
مرجع مقالات تخصصی ایران

✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
✓ امکان دانلود نسخه ترجمه شده مقالات
✓ پذیرش سفارش ترجمه تخصصی
✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
✓ دانلود فوری مقاله پس از پرداخت آنلاین
✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات