Vision Interface '99, Trois-Rivières, Canada, 19-21 May

Dynamic Programming Stereo Vision Algorithm for Robotic Applications

Rafael C. González, Jose A. Cancelas, Juan C. Alvarez, Jose A. Fernández, Ignacio Alvarez

Area de Ingeniería de Sistemas y Automática

Universidad de Oviedo

Campus de Viesques, s/n, 33204 Gijón

corsino@hecate.etsiig.uniovi.es

Abstract

Autonomous navigation applications demand sensors with a low sample time to be able to increase speed. Stereo vision algorithms that produce a dense disparity map present a slow time response (half a minute per frame at high resolution). We have developed a new cost function for dynamic programming stereo algorithms, capable to deliver dense disparity maps for single, high-resolution scanlines at high speed (40 ms/line), even for wide disparity rages (>100). We have tested the algorithm with both synthetic and real image, and we have compared its practical performance with other dynamic programming algorithms. Our cost function is based on a weighted sum of the squared intensity errors. Weight factors are based on gradient values. The occlusion cost is not constant for the whole image, instead we modify it depending on gradient values of matched points. In this paper we present this cost function, and compare its performance for autonomous navigation applications with other dynamic programming solutions.

1 Introduction

Autonomous navigation applications demand low sample time sensors to be able to increase speed. Traditionally, stereo vision algorithms presented high computation times due to the complexity of the algorithms. Some real-time applications have been reported, which provides depth calculation at real-time, but they are usually based on highly specialized hardware, meaning a high price. We have tried to find an algorithm capable to provide reasonably good depth estimations, in real-time, and for large disparity ranges.

The key problem in stereo vision is to establish the correct set of correspondences among selected features in two or more images taken at the same time. Correctness is measured, in general, by a cost function that incorporates terms for similarity and smoothness constraints. The matching problem is therefore stated as a constrained minimization problem. Many different solutions have been proposed. They can be classified according to the characteristics of the optimization algorithm, the kind of image features selected, the way they are described and whether they are multiresolution or not. In addition, some algorithms integrate different kind of image descriptions and even they integrate different depth cues (vergence, depth from focus).

Among this variety of algorithms, those based in dynamic programming present nice characteristics for real-time applications. Dynamic programming is an efficient way to minimize a function of many discrete variables. Epipolar geometry also contributes to speed up the search for an optimum set of matches because it reduces the search space to one dimension. Joining both tools, and assuming ordering constraints, search for correct matches becomes a minimum cost path-finding problem. Even in this case, algorithm performance will depend on the set of image features selected, the way to establish the search and the shape of the cost function.

Cox et al. have proposed a maximum likelihood algorithm independent of selected matching primitives. In their experiments, they use individual pixel intensities for the matching process. Resulting algorithm provides a dense disparity map without requiring feature extraction, and avoids the adaptive windowing problem of area-based correlation methods. Resulting cost function is very simple but presents multiple global minima. To deal with this problem, they introduce cohesivity constraints by searching for the solution that minimizes the number of intra- and inter-scanlines disparity discontinuities.

Intille and Bobick proposed a different approach. They based their algorithm on a data structure called Disparity-Space Image, and used it to look for matches and