画像解析と３次元モデル再構成

ダニエル　エイ　テフェラ

広島大学大学院工学研究科

Scene understanding and 3D model reconstruction

Daniel A Tefera

Abstract of Ph.D thesis (accepted February, 2002)

This thesis addresses the problem of 3D reconstruction from sets of images of a particular scene and the associated methodologies used to uncover the relationship among the views under consideration. The problem of 3D reconstruction is one of the extensively studied fields of computer vision. The thesis sets out with a brief introduction on the matters inspired the author to embark on this field and a description of the different applications generated as a by product to the solution of the correspondence problem. The thesis is designed to be self contained not only in its integrity but also among chapters (only a faint back and fro reference among chapters). It is my wish that each chapter stands as a prerequisite to the next one, however, readers with a knowledge of a chapter may consider to skip to the next.

Any computer vision application or algorithm uses a bunch of feature point, edge and corner, detectors as a preprocessor to assist in speeding up the process of analysis of a scenario. Of the few feature detector types edge detectors and corner detectors are popular. Edge detection algorithms are found in their tens in the literature of computer vision and image processing. Many produce similar results. Instead of taking you through a plethora of algorithms, the design, theoretical background and evaluation of Canny edge detector, probably the most widely used edge detector in today's computer vision community, is directly introduced in chapter two. Corner detectors also have been in literature with different guises quite for a long time, though not as popular as edge detector but equally important. Chapter five presents yet another novel corner detection approach.

Much in the same way that you need sport gears to get yourself tuned in a particular sport. We need some purely mathematical tools to understand this thesis. If there is one such topic which deserves a chapter it must be Grobner bases, other less significant mathematical tools are in appendix while the latter is in chapter three. However, we will not make use of Grobner bases anywhere before chapter six. As a result, readers who are interested only on the first half of the thesis may not need to read on Grobner bases.

An edge-pixel matching algorithm based on dynamic programming is presented in chapter four. Unlike

広島大学総合科学部紀要Ⅳ理系編、第28巻 (2002)
*広島大学審査学位論文
口頭発表日：2002年1月31日、学位取得日：2002年2月14日

classical dynamic programming methods, where a path is searched for across a graph of the whole space of edge pixels, this approach allows the optimal path to be within a constrained search space, and thus tremendously cutting off the searching time. The knowledge of the camera parameters, more specifically the input images being rectified, lends itself to reduce the general problem of image to image correspondence problem into a scanline to scanline matching problem.

Images of a single rigid object or scene are related by the so called epipolar geometry, which can be described by a 3 X 3 singular matrix. If the intrinsic parameters of the imaging system are known, one can work with the normalized image coordinates, and the matrix is known as the essential matrix; otherwise, we have to work with the pixel image coordinates, and the matrix is known as the fundamental matrix. It contains all geometric information that is necessary for establishing correspondences between images, from which three dimensional structure of the perceived scene can be inferred. Derivation and approximation of the fundamental matrix is the topic of chapter five.

Chapter six presents an algorithm which generates an initial set of matching corner points between two uncalibrated images of the same scene. Matching different images of a single scene remains one of the bottlenecks in computer vision. And, the core part of matching algorithms is recovering the epipolar geometry, which heavily depends on the accuracy of the set of eight corresponding points that are used to calculate the fundamental matrix between the sets of images under consideration. And, chapter six presents one such solution to alleviate this problem.

In chapter seven a through discuss is given on the relationship between the number of images considered simultaneously in an algorithm and the additional information found from the images about the 3D world therein. In particular it is shown that trilinearity is an optimum solution (arrangement among image sequences). The trifocal, aka trilinear, tensor arises as a relationship among corresponding points and lines in three views. This tensor may be computed linearly from a set of correspondences in three images, and leads to an algorithm for projective reconstruction from few line (point) correspondences in three views.

Finally, some full fledged algorithms on the reconstruction of the 3D world, provided that the fundamental matrix and/or few corresponding features among images (bilinear or trilinear arrangement) are known, are defined by the images is presented. Depending on the amount of a priori information available the method of reconstruction varies and we have covered different admissible arrangements without sacrificing the integrity of the thesis.