Stable Recovery of Shape and Motion from Partially Tracked Feature Points with Fast Nonlinear Optimization

Akira Amano, Tsuyoshi Migita, Naoki Asada
Faculty of Information Sciences, Hiroshima City University
{a-amano,migita,asada}@cv.its.hiroshima-cu.ac.jp

Abstract

The linearized approach to the shape from motion problem, e.g. the factorization method, is robust to find a unique solution with fast computation, but the occlusion and perspective distortion are out of scope in the linear formulation. In contrast, the nonlinear approach is free from such limitations, yet it involves two problems; one is to find the globally optimal solution and the other to reduce the computation time.

In this paper, we present an effective nonlinear optimization method to recover 3D shape and motion. To overcome the shortcomings of the nonlinear scheme, we propose "double search (DS) procedure" and "preconditioned conjugate gradient (PCG) algorithm". The DS procedure enables us to find two major solutions that correspond to the true and false shapes and then we select the globally optimal solution by evaluating the error of them. The PCG algorithm is an improved CG one whose computational performance is several times faster than the conventional Levenberg-Marquardt (LM) algorithm.

We carried out experiments with simulation and real data, and the results have demonstrated that the proposed method allows us to obtain the correct shape and motion with 3-9 times faster computation than the LM algorithm. Finally we have shown the applicability of the method to a large building reconstruction from a set of partially tracked feature points.

1 Introduction

The common interest in computer vision is to obtain 3D information from 2D images. Various approaches have been studied so far, and one practical method among them is "Shape from Motion"; that is, 3D shape of an object and relative motion between the object and camera are recovered from an image sequence.

3D reconstruction is the inverse problem of imaging process by perspective projection, and thus the relationship between 3D coordinates and the corresponding 2D ones is represented by nonlinear expressions in terms of the depth parameter. There are two different mathematical schemes to solve the nonlinear problem; one is linear approximation and the other nonlinear optimization.

The linearized approach to the shape from motion, e.g. the factorization method, is indeed robust to find a unique solution with fast computation, but some problems such as occlusion and perspective distortion are out of scope due to the linear formulation. In contrast, the nonlinear approach is free from such limitations, yet it involves two problems; one is to find the globally optimal solution and the other to reduce the computation time of the iterative search process.

In this paper, we propose a nonlinear optimization method to recover 3D shape and motion with fast and stable computation. To overcome the shortcomings of the nonlinear scheme described above, we have developed "double search (DS) procedure" and "preconditioned conjugate gradient (PCG) algorithm". The DS procedure enables us to find two major solutions that correspond to the true and false shapes and then we can select the globally optimal solution by evaluating the error of them. The PCG algorithm is an improved CG one whose computation performance is several times faster than the conventional Levenberg-Marquardt algorithm.

In what follows, we first give a brief survey on the shape from motion problem, then formalize it in the scheme of nonlinear optimization. Next, the DS procedure and PCG method are presented with technical details. The performance of the proposed method is evaluated quantitatively by using simulation and real data. Finally, we will show some experimental results of reconstructing buildings from a set of partially observed image data.

2 Shape from Motion

In shape from motion studies, "eight point algorithm[1]" established the theoretical framework of 3D information recovery from two images. The method however hardly works in actual situations because it takes no account of the observation error. Therefore the method is not robust to feature