

*Self-calibration of a stereo rig from unknown camera
motions and point correspondences*

Quang-Tuan LUONG

Olivier FAUGERAS

N° 2014

Juillet 1993

PROGRAMME 4

Robotique,
image
et vision

 *Rapport
de recherche*

1993

Self-calibration of a stereo rig from unknown camera motions and point correspondences

Quang-Tuan LUONG
Olivier FAUGERAS

Programme 4 — Robotique, image et vision
Projet Robotvis

Rapport de recherche n ° 2014 — Juillet 1993 — 35 pages

Abstract: The problem of calibrating a stereo rig is extremely important for practical applications. Existing work is based on the use of a calibration pattern whose 3D model is a priori known. We show theoretically and with experiments on real images, how it is possible to completely calibrate a stereo rig, that is to determine each camera's intrinsic parameters and the relative displacement between the two or three cameras, using only point matches obtained during unknown motions, without any a priori knowledge of the scenes.

The first part of the paper is devoted to the computation of the intrinsic parameters of the cameras by a method based upon the estimation of the so-called fundamental matrix associated with camera displacement. Three different displacements are sufficient to solve the Kruppa equations which yield these parameters.

The second part of the paper is devoted to the computation of the extrinsic parameters. We first explain how to recover the unknown motions previously used, once we have an estimate of the intrinsic parameters and the fundamental matrices. The computation is quite robust to the inaccuracy of the determination of the camera parameters. We then present the equations which allow us, from two displacements of the stereo rig, for which the camera motions are computed independently, to compute the relative displacement between the cameras. This technique allows us to compute the relative displacement between two or three cameras and complete the full calibration of the rig.

(Résumé : tsvp)

Ce travail a été financé en partie par le contrat Esprit 5390, RTGC.

Auto-calibration d'un système de stéréovision à partir de ses déplacements inconnus et de correspondances ponctuelles

Résumé : Le problème de la calibration d'un système de stéréovision est très important en pratique. Les méthodes existantes ont l'inconvénient d'être fondées sur l'utilisation d'une mire dont un modèle 3D est supposé connu. Dans ce papier, nous présentons une méthodologie complète, illustrée par des expériences conduites avec des images réelles, en vue d'*auto-calibrer* entièrement un système de stéréovision, c'est-à-dire de déterminer l'ensemble des paramètres intrinsèques de chaque caméra et le déplacement relatif entre les deux ou trois caméras, en partant uniquement d'appariements ponctuels obtenus au cours de déplacements inconnus, sans utiliser aucune connaissance a priori sur les scènes observées.

La première partie du papier est consacrée au calcul des paramètres intrinsèques d'une caméra par une méthode fondée sur l'estimation de la matrice dite fondamentale associée au déplacement de la caméra. Trois différents déplacements permettent de trouver une solution aux équations de Kruppa qui conduit à ces paramètres.

La seconde partie du papier est consacrée au calcul des paramètres extrinsèques. Nous expliquons d'abord comment obtenir les valeurs des mouvements inconnus précédemment utilisés, à partir des paramètres intrinsèques et des matrices fondamentales. Le calcul est relativement insensible à l'imprécision d'estimation des paramètres des caméras. Nous présentons ensuite les équations qui nous permettent, à partir de deux déplacements d'un système de stéréovision pour lesquels le mouvement de chaque caméra est déterminé indépendamment, de calculer le déplacement relatif entre les deux ou trois caméras et d'achever ainsi la calibration complète du système.

1 Introduction

1.1 The stereo calibration problem

Camera calibration is an important task in computer vision. The purpose of camera calibration is to establish the relationship between the 3D world coordinates and their corresponding 2D image coordinates. Once this relationship is established, 3D information can be inferred from 2D information, and vice versa. Thus camera calibration is a prerequisite for any application where this relation between 2D images and the 3D world is needed. In the case of a stereo rig, calibration is needed for at least two reasons. First to establish the epipolar geometry of the system and cut down the complexity of the stereo correspondence process. Second, to reconstruct the 3-D data after matching between the retinas has been completed.

The model which we consider is the most widely used. It is the pinhole model. The basic assumption behind this model is that *the relationship between the world coordinates and the pixel coordinates is linear projective*. Thus no camera distortion is considered which allows us to use the powerful tools of projective geometry, which is emerging as an attractive framework for computer vision [32]. In this chapter, we assume that the reader is familiar with the elementary projective geometry described in [39] for example. The equation of the model for one camera is:

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = \mathbf{A}_1 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{D}_1 \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{P}_1 \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (1)$$

where u, v are retinal coordinates, X, Y, Z , world coordinates, \mathbf{A}_1 a 3×3 transformation matrix accounting for camera sampling and optical characteristics, \mathbf{D}_1 a 4×4 displacement matrix accounting for camera position and orientation: \mathbf{D}_1 is the displacement from the world coordinate system (identified to the object that has been used for the calibration) to the camera coordinate system. The 3×4 matrix \mathbf{P}_1 is the perspective projection matrix, which relates 3D world coordinates and 2D retinal coordinates. \mathbf{A}_1 depends on a variable number of parameters, depending on the sophistication of the camera model: these parameters are called intrinsic. We will consider here a five-parameter model, represented in figure 1, that we explain later. \mathbf{D}_1 depends on 6 parameters, called extrinsic: 3 defining a rotation, 3 a translation, and has the form:

$$\mathbf{D}_1 = \begin{pmatrix} \mathbf{R}_1 & \mathbf{t}_1 \\ \mathbf{0}_3^T & 1 \end{pmatrix} \quad (2)$$

Using a second camera, introduces another set of intrinsic parameters through matrix \mathbf{A}_2 , and another set of extrinsic parameters through matrix \mathbf{D}_2 . However, in general the choice of a particular world coordinate system versus another one is arbitrary, the significant thing being the relative position and orientation of the two cameras, thus we can consider that there are only 6 significant extrinsic parameters for a binocular stereo rig. They are given by the 4×4 displacement matrix \mathbf{D} , representing the displacement from the first camera to the second camera, which can be easily computed from \mathbf{D}_1 and \mathbf{D}_2 . If we don't consider a coordinate system linked to an object of the world, we must express this displacement in a