Object Pose: The Link between Weak Perspective, Para Perspective, and Full Perspective

Radu Horaud, Stéphane Christy, Fadi Dornaika

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Object Pose: The Link between Weak Perspective, Para Perspective, and Full Perspective

Radu Horaud, Stéphane Christy, and Fadi Dornaika

N° 2356

September 94

PROGRAMME 4

Robotique,

image

et vision

Rapport de recherche

1994

ISSN 0249-6399

Object Pose: The Link between Weak Perspective, Para Perspective, and Full Perspective *

Radu Horaud, Stéphane Christy, and Fadi Dornaika

Programme 4 — Robotique, image et vision
Projet MOVI

Rapport de recherche n ° 2356 — September 94 — 25 pages

Unité de recherche INRIA Rhône-Alpes
46 avenue Félix Viallet, 38031 GRENOBLE Cedex 1 (France)
Téléphone : (33) 76 57 47 77 – Télécopie : (33) 76 57 47 54

Abstract: Recently, Dementhon & Davis [4] proposed a method for determining the pose of a 3-D object with respect to a camera from 3-D to 2-D point correspondences. The method consists of iteratively improving the pose computed with a weak perspective camera model to converge, at the limit, to a pose estimation computed with a perspective camera model. In this paper we give a simple derivation of Dementhon and Davis' method and we show that it belongs to a larger class of methods where the perspective camera model is approximated either at zero order (weak perspective) or first order (para perspective). We analyse the convergence of these methods and we conclude that the iterative para perspective method has better convergence properties than the iterative weak perspective method. We introduce a simple way of taking into account the orthogonality constraint associated with the rotation matrix. We analyse the sensitivity to camera calibration errors and we define the optimal experimental setup with respect to imprecise camera calibration. We compare the results obtained with this method and with a non-linear optimization method.

Key-words: Object pose, weak perspective, para perspective, camera calibration, extrinsic camera parameters.

(Résumé : tsvp)

*This work has been supported by the Esprit programme through the Basic Research Project "SECOND" (Esprit-BRA No. 6769). R. Horaud acknowledges partial support from the University of Bern, Institut fur Informatik und angewandte Mathematik, where he was a visiting scientist during July 1994.

Calcul de pose : le lien entre perspective faible, para perspective et perspective

Résumé : Récemment, Dementhon & Davis [4] ont proposé une méthode pour déterminer la pose d'un objet 3-D par rapport à une caméra à partir de correspondances de points 3-D/2-D. La méthode améliore itérativement la pose calculée avec une perspective faible pour converger, à la limite, vers une estimation de la pose calculée avec une vraie perspective. Dans ce papier on présente une fa con très simple d'obtenir le résultat de Dementhon & Davis et on montre que leur méthode appartient à une classe de méthodes plus générales lorsque le modèle de caméra perspective est approximé à l'ordre 0 (perspective faible) ou à l'ordre 1 (para perspective). On analyse la convergence de ces méthodes et on conclut que que la methode para perspective itérative a une meilleure convergence que la methode perspective faible itérative. On introduit une fac,on très simple de prendre en compte la contrainte d'orthogonalité associé avec une rotation. On analyse ensuite la sensibilité de la méthode par rapport aux erreurs de calibration de la caméra et on définit les conditions expérimentales optimales par rapport à une calibration imprécise. On compare les résultats obtenus avec cette méthode avec une méthode de minimisation non-linéaire.

1 Introduction

Recently, Dementhon & Davis [4] proposed a method for determining the pose of a 3-D object with respect to a camera from 3-D to 2-D point correspondences. The method consists of iteratively improving the pose computed with a weak perspective camera model to converge, at the limit, to a pose estimation computed with a perspective camera model. At our knowledge, Dementhon & Davis method is among one of the first attempts to link linear techniques associated with the weak perspective camera model with non linear techniques associated with a perspective camera model. On one side, linear resolution methods can be used but the solution thus obtained is just an approximation and, on the other side, non linear resolution methods lead to a very accurate solution but proper initialization is required.

One possibility could be to use a robust numerical method to perform the non linear minimization that is initialized with the solution obtained by a linear method. However, there are several drawbacks. First, such an approach does not take into account the simple mathematical link that exists between the perspective model and its linear approximations. Second, the linear approximation may be quite faraway from the true solution and a large number of iterations would be required before the non-linear minimization process converges to a stable solution.

The perspective projection is modelled by a projective transformation from the 3-D projective space to the 2-D projective plane. Weak perspective is just an affine approximation of full perspective. More precisely, it may well be viewed as a zero-order approximation: $1/(1+\epsilon) \approx 1$. Para perspective [1] is a first order approximation of full perspective: $1/(1+\epsilon) \approx 1-\epsilon$. The method proposed by Dementhon & Davis starts with computing the pose of an object using weak perspective and after a few iterations converges towards a pose estimated under perspective. The method is very elegant, very fast, and quite accurate. It is however limited to situations where the weak perspective approximation is valid. If the object is closed to the camera and/or at some distance away from the optical axis then the pose algorithm of Dementhon & Davis either converges very slowly (100 iterations rather than 5 to 10) or it doesn't converge at all.

In this paper we show how the initial method as proposed in [3] and [4] may be extended to para perspective. More precisely, we describe a method for computing object pose with a para perspective model and we establish the link between para perspective pose and perspective pose. We show both theoretically and experimentally that our method has better convergence properties than the method proposed by Dementhon & Davis. Moreover we introduce a simple computational way of taking into account the orthogonality constraint associated with the 3×3 rotation matrix describing the orientation between the 3-D object and the camera. Indeed, the linear pose algorithms using weak perspective and para perspective do not guarantee that this rotation matrix is orthogonal. The orthogonalization method that we describe below computes the best rotation in closed form using unit quaternions. This orthogonalization method considerably increases the robustness of the method at the cost of very few extra computations. We characterize the best experimental setup that allows one to compute a precise pose even in the presence of camera calibration errors. Finally we provide a comparison between the results obtained with this method and the results obtained with a non linear minimization method [12], [13].

1.1 Background

The problem of object pose from 2-D to 3-D correspondences has received a lot of attention both in the photogrammetry and computer vision literature. Haralick et al. [8] cites a German dissertation from 1958 that surveys nearly 80 different solutions from the photogrammetry literature. Various approaches to the object pose (or external camera parameters) problem fall into 2 distinct categories: closed-form solutions and numerical solutions.

Closed-form solutions may be applied only to a limited number of correspondences [7], [9], [5]. Whenever the number of correspondences is larger than 4 then closed-form solutions are not efficient any more and iterative numerical solutions are necessary [15], [10], [8]. These approaches are, in general, very robust but they converge towards the correct solution on the premise that a good initial estimate of the true solution is provided. Phong et al. [13] describe a method that uses trust-region optimisation and that is less sensitive to initialisation than other minimization methods. However, the method of Phong et al. performs well for a relatively large number of correspondences. Whenever the number of correspondences is between 3 and 10, then the trust-region minimization method either requires a large number of iterations or doesn't converge towards the correct solution. From a practical point of view, it is important to have an object pose algorithm which doesn't necessarily require a large number of correspondences and which doesn't suffer from the limitations that are inherent to closed-form methods.

The method recently proposed by Dementhon & Davis [4] requires an arbitrary number of point correspondences. The 3-D points may be either in general position (i.e., non-coplanar) or coplanar and there should be at least

4 point correspondences. The method is fast and it is robust with respect to image measurements and to camera calibration errors. However, the object pose algorithm as suggested by Dementhon and Davis has convergence problems if the object is too faraway from the optical axis of the camera — a limitation that can be overcome as explained farther in this paper.

1.2 Paper organization

The remainder of this paper is organized as follows. In Section 2 we briefly recall the perspective camera model and two possible approximations of this model, weak and para perspective. Section 3 describes the iterative weak perspective algorithm as it has been proposed in [4]. Section 4 describes the new iterative para perspective algorithm that we propose as well as how to compute the object pose from a para perspective approximation. Section 5 provides details for solving the linear equations associated with pose computation in the case of both non coplanar and planar sets of object points. Section 6 analyses the convergence of both the weak and para perspective iterative algorithms and Section 7 shows how to improve the pose algorithms using a simple and straightforward formulation of the orthogonality constraint. Section 8 analyses the sensitivity of both algorithms to camera calibration and Section 9 provides an experimental comparison of the two methods described in this paper together with a comparison of these two methods with a non-linear minimization method. Finally Section 10 provides a discussion and gives directions for future work.

2 Camera models

Let us consider a simple setup, as depicted on Figure 1. We denote by P_i a 3-D point with coordinates X_i, Y_i, and Z_i in a frame that is attached to the object – the object frame. The origin of this frame is the object point P_0. An object point P_i projects onto the image in p_i with camera coordinates x_i and y_i and we have ($\mathbf{P}_i$ is the vector from point P_0 to point P_i):

$$x_i = \frac{\mathbf{i} \cdot \mathbf{P}_i + t_x}{\mathbf{k} \cdot \mathbf{P}_i + t_z} \tag{1}$$

$$y_i = \frac{\mathbf{j} \cdot \mathbf{P}_i + t_y}{\mathbf{k} \cdot \mathbf{P}_i + t_z} \tag{2}$$

These equations describe the classical perspective camera model where the rigid transformation from the object frame to the camera frame is:

$$T = \begin{pmatrix} \mathbf{i}^T & t_x \\ \mathbf{j}^T & t_y \\ \mathbf{k}^T & t_z \\ 0\,0\,0 & 1 \end{pmatrix} = \begin{pmatrix} R & \mathbf{t} \\ 0\,0\,0 & 1 \end{pmatrix}$$