Computing Differential Properties of 3-D Shapes from Stereoscopic Images without 3-D Models

Frédéric Devernay, Olivier Faugeras

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Computing Differential Properties of 3-D Shapes from Stereoscopic Images without 3-D Models

Frédéric Devernay   Olivier Faugeras

N° 2304

Juillet 1994

PROGRAMME 4

Robotique,

image

et vision

rapport de recherche

1994

ISSN 0249-6399

Computing Differential Properties of 3-D Shapes from Stereoscopic Images without 3-D Models

Frédéric Devernay Olivier Faugeras

Programme 4 — Robotique, image et vision
Projet Robotvis

Rapport de recherche n° 2304 — Juillet 1994 — 28 pages

Abstract:

We are considering the problem of recovering the three-dimensional geometry of a scene from binocular stereo disparity. Once a dense disparity map has been computed from a stereo pair of images, one often needs to calculate some local differential properties of the corresponding 3-D surface such as orientation or curvatures. The usual approach is to build a 3-D reconstruction of the surface(s) from which all shape properties will then be derived without ever going back to the original images. In this paper, we depart from this paradigm and propose to use the images directly to compute the shape properties. We thus propose a new method extending the classical correlation method to estimate accurately both the disparity and its derivatives directly from the image data. We then relate those derivatives to differential properties of the surface such as orientation and curvatures.

We present the results of the reconstruction and of the estimation of the surface orientation and curvatures on some stereo pairs of real images.

Key-words: Stereoscopy, Differential properties of surfaces, Disparity interpretation, Correlation, Three-dimensional vision.

(Résumé : tsvp)

Unité de recherche INRIA Sophia-Antipolis
2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex (France)
Téléphone : (33) 93 65 77 77 – Télécopie : (33) 93 65 77 65

Calcul de propriétés différentielles de surfaces 3-D à partir de paires stéréoscopiques sans modèle 3-D

Résumé :

Nous considérons le problème qui consiste à retrouver la géométrie tridimensionnelle d'une scène à partir de la disparité stéréoscopique binoculaire. Une fois qu'une carte dense de disparité a été calculée à partir d'une paire d'images stéréoscopique, on a souvent besoin de calculer des propriétés différentielles de la surface 3-D correspondante, comme l'orientation ou les courbures. L'approche habituelle consiste à construire d'abord une reconstruction 3-D de la surface, puis à calculer les propriétés de cette surface à partir de cette reconstruction, sans jamais revenir aux images originales. Dans cet article, nous proposons une toute autre méthode qui permet de calculer les propriétés des surfaces directement à partir des images. Cette méthode est une extension des méthodes classiques de stéréo par corrélation qui permet d'estimer à la fois la disparité et ses dérivées à partir des données d'intensité d'image. Nous établissons également le lien entre ces dérivées de la disparité et les propriétés différentielles de la surface comme l'orientation ou la courbure.

Nous présentons les résultats de la reconstruction 3-D et de l'estimation de l'orientation de la surface sur quelques paires d'images réelles.

Mots-clé : stéréoscopie, propriétés différentielles de surfaces, interprétation de la disparité, corrélation, vision tridimensionnelle.

Contents

1 Introduction

1.1 Motivation

Three-dimensional shape analysis in computer vision has often been considered as a two step process in which a) the structure of the scene is first recovered as a set of coordinates of 3-D points and b) models are fitted to this data in order to recover higher-order shape properties such as normals or curvatures which are first and second order differential properties of the shape surface. The original images are not used anymore whereas this is clearly where the information lies.

In some applications one may want to use even higher order properties such as affine or projective differential invariants [7], that would be especially useful in the situation of an uncalibrated stereo rig [6]. All these quantities can be expressed, using the perspective projection matrices of the two cameras (or the fundamental matrix in the case of an uncalibrated system), in terms of the derivatives of the disparity field. As a consequence, we are confronted to the task of estimating the spatial derivatives of the disparity map and we explore the possibility of estimating these derivatives directly from the images rather than applying to the disparity map the same paradigm as to the 3-D shape.

1.2 Related Work

We very briefly review the work that has been done in interpreting stereo disparity. Up till now the major part of the existing studies is on the interpolation or approximation of the possibly sparse disparity map by a surface. This was done using either minimization of spline functions [1, 15, 9], or interpolation by polynomial surface patches [10]. To calculate the orientation of the observed surface one approach was to simply differentiate numerically the point-by-point distance reconstruction [3]. Surface orientation can also be recovered using a set of differencing operators that are involved during the minimization process in the spline approach [15] or directly from the coefficients of the surface patches in the polynomial approach [10]. For both methods the presence of surface discontinuities can be detected and taken into account.

Theoretical results were also obtained for the calculation of surface orientation by studying the local projections of a surface and the displacement vector field generated by movement (i.e. the optical flow) [13, 11], or the disparity field in the case of stereoscopy [12]). The common point of these approaches is the application of tensor analysis to the classical field theory. A result that is closer to the approach presented here is the calculation of the three dimensional surface orientation from the disparity field and its first derivatives [16], but, as we show in this article, it can be done much more simply, thus allowing to reach the second order, i.e. curvatures.

1.3 Contributions

If we want to calculate the local differential properties of a 3-D surface, we can go at least two ways; first, we can reconstruct the scene points in three dimensions, fit some surfaces to the reconstructed points and compute the differential properties from the fitted surfaces. A second possibility is to avoid the explicit reconstruction and work directly from the disparity map. Because of the computational effort involved in the first approach and the extra noise added by the reconstruction process, we choose here the second approach.

We thus first have to compute the derivatives of the disparity. Since the precision of the dense disparity map calculated by a standard correlation technique is only about one pixel, we must either regularize the disparity map or compute its derivatives differently. The first solution implies the use of *local* regularization techniques, because we must keep in mind that the disparity map may contain holes due for example to object discontinuities or occluding contours. As a consequence we cannot use in a straightforward manner derivating filters that are nicely behaved, such as Gaussian derivative filters [4].

For these reasons we chose to explore a second solution and present a new method to compute precise values of the disparity and its derivatives, up to second derivatives, directly from the image intensity data, i.e. without using explicitly any derivation operator. For this we use the local deformations of the image of a surface from the left view to the right view, which give information on the derivatives of the disparity. We implemented an extension of a correlation technique based on this principle which gives good results, up to the second order derivatives of disparity. The counterpart is that it is much slower than calculating the derivatives from the disparity map itself, but the quality of the results is worth the effort.

We then present a method to compute the three-dimensional surface orientation from the first-order derivatives of the disparity. The analytic expressions are very simple when working in standard coordinates (i.e. when the images are rectified so that epipolar lines are horizontal). We also extend this to the computation of surface curvature from the second-order derivatives of the disparity, but the resulting expressions are less simple.

We tested our algorithms on real images successfully, and the results are presented at the end of this paper.

Our method can be easily extended to the case of an uncalibrated stereo rig [6], but in that case we will need to use either affine or projective invariants of the surface instead of Euclidean invariants. The notion of surface orientation is the same in an affine or projective world, but we will have to use the affine or projective normals and affine or projective curvatures instead of the Euclidean curvatures to characterize the higher-order differential properties of the surface. The use of weak calibration seems to be a very promising and still mostly unexplored field of research.

2 Computing derivatives of disparity

2.1 Difficulties of the computation

If we want to know the local surface orientation and curvatures from a stereoscopic pair of images, we have to calculate somehow the derivatives with respect to the image coordinates of the disparity map of a three-dimensionnal object. As always we encounter the most common problem in computer vision: the data contains too much noise to be simply derived because the precision of a disparity map obtained by a standard correlation technique is not much better than one pixel.

Another obstacle to the computation of the derivatives from the disparity map is that the map is *dense* but nevertheless contains holes in it due to occluded regions, discontinuities of the disparity or positions where no maximum was found for the correlation criterion. For these reasons we cannot use the nice deriving operators [4] because they have an infinite support, and if we want to regularize the disparity map we will have to use *local* operators like the one we present below.

Because the first order derivatives that we computed by regularization of the dense disparity map were somewhat unsatisfactory we thought of another method that would compute the derivatives directly from the image intensity data, without using any deriving operator. We developed an enhanced correlation method that takes into account the local deformations from one image to the other and is described later. Using this method gives more accurate results, but at a much higher computational cost.

2.2 Computing derivatives from the disparity map

As we pointed out, if we want to use the disparity map itself to calculate its derivatives we have to perform some local regularization that can handle holes in the disparity map. We chose to do a least square approximation of the data located in the neighborhood of a point where we want to calculate the derivatives by a model, and then deduce the derivatives from the equations of the fitted model.

Since we first want to calculate the first derivatives of the disparity from the disparity map, the model we use is simply a plane of equation:

$$z(x, y) = ax + by + c \tag{1}$$

and the neighborhood in which the approximation is made is an $h \times v$ window centered at the current point, so that the fitting algorithm is only a planar regression between a set of 3-D points (x_i, y_i, d_i) where d_i is the disparity at point (x_i, y_i). The solution of the regression is:

$$a = \frac{S_{xz}\, a_1 + S_{yz}\, a_2 + S_z\, a_3}{\Delta} \tag{2}$$

$$b = \frac{S_{xz}\, a_2 + S_{yz}\, b_2 + S_z\, b_3}{\Delta} \tag{3}$$

$$c = \frac{S_{xz}\, a_3 + S_{yz}\, b_3 + S_z\, c_3}{\Delta} \tag{4}$$

with:

$$
\begin{array}{ccc}
 & S = \sum_{i=1}^{N} \frac{1}{\sigma_i^2} & \\
S_x = \sum_{i=1}^{N} \frac{x_i}{\sigma_i^2} & S_y = \sum_{i=1}^{N} \frac{y_i}{\sigma_i^2} & S_z = \sum_{i=1}^{N} \frac{z_i}{\sigma_i^2} \\
S_{xx} = \sum_{i=1}^{N} \frac{x_i^2}{\sigma_i^2} & S_{xy} = \sum_{i=1}^{N} \frac{x_i y_i}{\sigma_i^2} & S_{xz} = \sum_{i=1}^{N} \frac{x_i z_i}{\sigma_i^2} \\
S_{yy} = \sum_{i=1}^{N} \frac{y_i^2}{\sigma_i^2} & S_{yz} = \sum_{i=1}^{N} \frac{y_i z_i}{\sigma_i^2} & S_{zz} = \sum_{i=1}^{N} \frac{z_i^2}{\sigma_i^2} \\
a_1 = S\,S_{yy} - S_y^2 & a_2 = S_x\,S_y - S\,S_{xy} & a_3 = S_y\,S_{xy} - S_x\,S_{yy} \\
b_2 = S\,S_{xx} - S_x^2 & b_3 = S_x\,S_{xy} - S_y\,S_{xx} & c_3 = S_{xx}\,S_{yy} - S_{xy}^2 \\
 & \Delta = S\,c_3 + S_y\,b_3 + S_x\,a_3 &
\end{array}
\tag{5}
$$

where $N = hv$ and σ_i is the uncertainty associated to the disparity value at pixel i. Once we have the equation of the fitted model, the coefficients a and b correspond to the first derivatives of the disparity over respectively x and y (the value of c was not used though it could be considered as a regularized value of the disparity). We can also calculate the standard deviations σ_a and σ_b on a and b:

$$\sigma_a^2 = \frac{S\,S_{yy} - S_y^2}{S\,S_{xx}\,S_{yy} - S\,S_{xy}^2 + 2\,S_{xy}\,S_x\,S_y - S_y^2\,S_{xx} - S_{yy}S_x^2} \tag{6}$$

$$\sigma_b^2 = \frac{S\,S_{xx} - S_x^2}{S\,S_{xx}\,S_{yy} - S\,S_{xy}^2 + 2\,S_{xy}\,S_x\,S_y - S_y^2\,S_{xx} - S_{yy}S_x^2} \tag{7}$$

and we decide to keep the values a and b as the derivatives of the disparity if there are enough data points, $\sigma_a < s$, $\sigma_b < s$, and $a > -1$, where the threshold s must be fixed (we chose $s = 0.05$ for our applications).

The last condition corresponds to the differentiation of the classical ordering constraint [8]. The ordering constraint says that if two points A_1 and B_1 lie close enough on the same epipolar line in a certain order (e.g. A_1 is closer to the epipole E_1 than B_1) in the first retina, then the images of the corresponding 3-D points in the second retina lie in the same order. For example in standard coordinates (i.e. the epipolar lines are horizontal) if $A_1 = (u_1, v_1)$ and $A_2 = (u_1 + d(u_1, v_1), v_1)$, let $B_1 = (u_1 + h, v_1)$, then a first order approximation of B_2 is:

$$B_2 = (u_1 + h + d(u_1, v_1) + h\frac{\partial d}{\partial u_1} + o(h), v_1)$$

so that the ordering constraint can be rewritten $\frac{\partial d}{\partial u_1} \geq -1$.

This last condition can also be replaced by the disparity gradient limit [14] which may be closer to the human visual system limit. Since the disparity gradient is related to local surface orientation, this is equivalent to setting a maximum angle between the local orientation of the surface and the optic ray.

A perhaps better method to compute the derivatives of disparity from the disparity map can be found in [10], but the quality of the results will still depend on the precision of the disparity map, which is the crucial problem.

2.3 An enhanced correlation method

We know that the disparity map we calculate by correlation has only a precision of about one pixel, so it cannot be derived to calculate local orientation and curvatures. We thought that instead of trying to look for the derivatives in the disparity map, where they may be definitively lost, why not look for these directly in the image intensity data? In this subsection we work in standard coordinates, i.e. the original images are rectified so that the epipolar lines are horizontal and consequently the disparity between the left and right images is only horizontal. The reference image used for the computation of disparity is the left image.

2.3.1 Principle

The idea comes from the following observation: a small surface element that is viewed as a square of pixels in the left image can be seen in the right image as a distorted square, and an approximation of this distorted square can be computed from the derivatives of the disparity. Let us first show it with a first order approximation.

The first order derivatives of disparity: Let us call $d(u,v)$ the disparity at point (u,v), that is to say that the point in the right image corresponding to (u,v) is $(u+d(u,v),v)$. Let $\alpha(u,v)$ and $\beta(u,v)$ be the derivatives of $d(u,v)$ with respect to u and v, respectively. Then the point corresponding to $(u+du,v+dv)$ is: