Region tracking through image sequences

Bénédicte Bascle, Rachid Deriche

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Region tracking through image sequences

Benedicte Bascle, Rachid Deriche

N° 2439

Decembre 1994

PROGRAMME 4

rapport de recherche

ISSN 0249-6399

Region tracking through image sequences

Benedicte Bascle, Rachid Deriche

Programme 4 — Robotique, image et vision
Projet Robotvis

Rapport de recherche n° 2439 — Decembre 1994 — 21 pages

Abstract: This paper describes a new approach to the tracking of complex shapes through image sequences, that combines deformable region models and deformable contours. A new deformable region model is presented: its optimisation is based on texture correlation and is constrained by the use of a motion model, such as rigid, affine or homographic. The use of texture information (versus edge information) noticeably improves the tracking performances of deformable models in the presence of texture. Then the region contour is refined using an edge-based deformable model, in order to better deal with specularities and non planar objects. The method is illustrated and validated by experimental results on real images.

Key-words: Deformable models, region, contour, tracking

(Résumé : tsvp)

We would like to thank Gerard Medioni for his suggestions and advice during the progress of this work, and Andrew Blake and Andrew Zisserman for their help with the writing of this article. Email: bascle@sophia.inria.fr

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

Suivi de régions sur une séquence d'images

Résumé : Ce rapport décrit une nouvelle approche de suivi de formes complexes sur une séquence d'images, qui combine les modèles déformables de région et de contours. Un nouveau modèle déformable de région est présenté: son optimisation est basée sur la corrélation des niveaux de gris, et est contrainte par un modèle de mouvement, tel que rigide, affine ou homographique. L'utilisation de l'information donnée par les niveaux de gris (par rapport à l'information de contour) améliore notablement les performances de suivi des modèles déformables en présence de texture. Ensuite le contour de la région est raffiné par un contour déformable, de manière à corriger les erreurs dus aux spécularités et à la non-planarité des objets. La méthode est illustrée par des résultats expérimentaux sur des images réelles.

Mots-clé : Modèles déformables, région, contour, suivi

Summary

1. **What is the original contribution of this work?**

   This paper describes a new deformable model for region tracking, and presents a tracking approach that combines deformable regions and deformable contours. Previous deformable models found in the literature are either edge-finding curves or homogeneity-maximising regions. In contrast, this paper presents a deformable model for the tracking of regions based on texture correlation. Furthermore, other works that aim to use region information rely on a pre-computation of the region or contour optical flow for providing additional information about the region motion, and use it to guide the optimisation of the region outline by a deformable contour. But such methods optimise the region contour and, moreover, the computation of the optical flow introduces an external source of error in the tracking algorithm. The approach presented here optimises the region itself (and not the region contour), and is directly based on texture correlation, without resorting to external information such as optical flow. Furthermore, the combination of deformable regions and deformable contours allows us to track textured regions with a certain robustness to specularitites, occlusion and non-planarities.

2. **Why should this contribution be considered important?**

   This new texture-based deformable model of a region greatly improves the robustness of tracking when texture appears in the image, and provides a complementary approach to deformable contours which are more efficient for tracking purposes when texture is poor. Further, while this work clearly takes inspiration from correlation algorithms, it differs from these approaches in the following aspects. Firstly, given a window in the first image, correlation algorithms scan all possible corresponding windows in the second image and compute a correlation score for each of them. Then the window that has the best correlation score is said to correspond to the new location of the initial window. The method presented in this paper relies on the "active" behaviour of deformable models to find the new location of the region in the next image: the region is optimized iteratively by minimizing the correlation score. Thus an exhaustive search of the possible locations is avoided. It can also be noted that our method does not require the knowledge of the epipolar lines (i.e. prior calibration or auto-calibration of the system) to perform region tracking, though this knowledge can be used an additional constraint if it is available. Secondly, the shape of the correlation window is variable, depending on the previous shape of the region and the current estimation of the region motion (which is updated iteratively), whereas most correlation algorithms use fixed-size rectangular correlation windows.

3. **What is the most closely related work by others and how does this work differ?**

   Berger optimises the rigid 2D motion of a deformable contour so that it fits the optical flow measured along the contour by a separate algorithm. Then, in a second step, the deformable contour is optimised freely (without the rigid motion constraint) to better fit image edges. Bascle *et al* tracks a region with a deformable contour, using an affine motion model. Given the outlines of the tracked region, an affine approximation of the optical flow inside the region is computed, and used to make a prediction of the region position in the next image. This prediction is used to initialise the deformable contour in the most likely searching area. In contrast, this paper describes a deformable model of region, instead of contour. And the optimisation of this region model is based directly on the image texture, and not on an external optical flow measurement. Ivins and Porrill [IP94] describe a deformable region that expands from a seed as much as possible while remaining consistent with the statistics of the region measured at the beginning. These statistics are the mean and covariance associated with the texture potentials measured inside the region. The (eight) texture potentials are those defined by Laws [Law79] and result from combinations of smoothing, edge detection, and spot detection. This method is very efficient for the segmentation of textured regions, but it is oriented towards the detection of regions of homogeneous texture. It is also based on first derivative measurements inside the region. On the contrary, our deformable model is based on grey-level information, and can track regions of non-homogenous texture, being given a segmentation of the initial region in the first image. Szeliski [Sze94] recently used an intensity-matching technique similar to ours for image registration of planar patches. As in our case, the parameters of the homographic transformation between the images of a planar patch are optimized in order to minimize the discrepancy in intensities between the pair of images. However, our method differs in a number of important respects: first we track a region of interest delineated by a flexible contour in the image, while Szeliski optimizes the position of the whole image (i.e. a more global and less flexible rectangular window) with respect to a mosaic that positions the successive frames relative to each other. Second, our correlation criterion is normalized and the means of the intensities inside the regions are subtracted from the intensities to be compared between the two images: this is substantially more robust to lighting changes between the images; previous works on correlation have shown that normalization and use of the intensity mean greatly improves the performances of correlation algorithms. Third, our deformable region tracking is combined to deformable contour tracking, to better deal with specularities and non-planar regions.

4. **How can other researchers make use of the results of this work?**

   This region deformable model can be used as a complementary tracking tool to contour deformable models, being efficient for tracking of highly textured regions, whereas contour-based approaches are more efficient when texture is poor.

1 Introduction

Feature tracking in image sequences provides rich support to the analysis of time-varying environments. For instance, target tracking is used in surveillance tasks. Tracking algorithms are also useful to determine the 2D motion and deformations of an object: Herlin [HBG+94], Cohen [CAS92], Nastar [NA93] study the deformations of a cardiac ventricle in echographic images; Leymarie [Ley90] studies the deformation of biological cells in order to understand cell migration and defence mechanisms. Some collision avoidance and path planning strategies are also based on contour tracking: Cipolla [CB92] and Meyer [MB92] track contours in order to estimate time-to-collision; Curwen [CB92] track occluding contours in order to compute free space and to plan a possible path for a robot arm. Moreover feature tracking provides a good basis for the estimation of 3D motion and/or structure [Sha86] [BR86] [Har92] [DF90] [SJ87] [SA91] [FP93].

First investigations were concerned with the tracking of points [SJ87] and edge segments [CSD88] [DF90] [GCDVF93]. However, this data is sparse and sensitive to occlusion. Consequently, interest was driven to the tracking of more complex and global primitives, such as contours and regions. Among theses tracking methods - aimed at tracking global primitives, deformable models ("snakes") have had a considerable amount of success, due to their "active" behaviour, aimed at energy minimisation [Ter86] [KWT87]. Indeed this active behaviour enable tracking methods based on deformable models to perform simultaneous extraction and matching of the searched feature.

Different types of deformable model have been proposed in the literature for tracking purposes. A number of them are edge-based and seek high gradient areas of the image [KWT87] [YH92] [CB92] [CAS92] [CB92] [TS92] [NA93] [CCCD93] [KWM94] [NFSK94] [MSV94]. Some approaches uses an additional a-priori model of the contour displacement to constrain the optimisation of the deformable contour: Blake [BCZ93] optimises the deformable contour points (with Kalman-like dynamic equations) under an affine motion constraint whereas Bascle [BD93b, BD93a] actually optimises the (rigid, affine or homographic) motion parameters themselves instead of the contour points. Other approaches incorporate outside information given by optical flow measurements (made separately) to guide the optimisation of deformable contours. For instance, Berger [Ber94] computes the optical flow along the image contour and then constrains the deformable contour optimisation accordingly. Bascle *et al* [BBDM94] determine an affine approximation of the velocity field inside the region, and use this information to initialise the tracking of the region contour in the next image. A last class of approaches optimises the region contour, not by looking for edge areas in the image, but by maximising the grey-level homogeneity inside the region: Fua [FL89] optimises the region contour by maximising homogeneity inside the region edge, whereas Deriche and Blaszka [DB93] optimise a full model of an homogeneous region over an homogeneous background (including the region edge curve, the intensity values inside and outside the region, and the image smoothing parameter).

This paper describes a new approach for the tracking of moving regions, that combines region and contour tracking. First the whole region is tracked by a new deformable model of region. The optimisation of the deformable region is based on the minimisation of a correlation-like criterion, measuring the similarity between the grey levels measured inside