Knowledge-Based Systems 54 (2013) 255–268

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Hierarchical fuzzy logic based approach for object tracking

Nuno Vieira Lopes [a,b,*], Pedro Couto [b], Aranzazu Jurio [c], Pedro Melo-Pinto [b]

[a] School of Technology and Management, Polytechnic Institute of Leiria, Morro do Lena, Alto do Vieiro, Apartado 4163, 2411-901 Leiria, Portugal
[b] CITAB – Centre for the Research and Technology of Agro-Environmental and Biological Sciences, University of Trás-os-Montes e Alto Douro, Quinta de Prados, 5000-911 Vila Real, Portugal
[c] Department of Automatic and Computation, Public University of Navarra, Campus Arrosadía, 31006 Pamplona, Spain

ARTICLE INFO

Article history:
Received 6 April 2013
Received in revised form 9 September 2013
Accepted 12 September 2013
Available online 20 September 2013

Keywords:
Fuzzy logic
Dynamic fuzzy sets
Inference engine
Hierarchical
Multiple object tracking

ABSTRACT

In this paper a novel tracking approach based on fuzzy concepts is introduced. A methodology for both single and multiple object tracking is presented. The aim of this methodology is to use these concepts as a tool to, while maintaining the needed accuracy, reduce the complexity usually involved in object tracking problems. Several dynamic fuzzy sets are constructed according to both kinematic and non-kinematic properties that distinguish the object to be tracked. Meanwhile kinematic related fuzzy sets model the object's motion pattern, the non-kinematic fuzzy sets model the object's appearance. The tracking task is performed through the fusion of these fuzzy models by means of an inference engine. This way, object detection and matching steps are performed exclusively using inference rules on fuzzy sets. In the multiple object methodology, each object is associated with a confidence degree and a hierarchical implementation is performed based on that confidence degree.

1. Introduction

Object tracking plays an important role in computer vision. During the last years, extensive research has been conducted in this field and many types and applications of object tracking systems have been proposed in the literature such as automated surveillance, vehicle navigation, human computer interaction, and traffic analysis [1–7].

Tracking is essential to many applications and robust tracking algorithms are still a huge challenge. Difficulties can arise due to noise presence in images, quick changes in lighting conditions, abrupt or complex object motion, changing appearance patterns of the object and the scene, non-rigid object structures, object-to-object and object-to-scene occlusions, camera motion and real time processing requirements. Typically, assumptions are made to constrain the tracking problem in the context of particular applications. For instance, almost all tracking algorithms assume that the object motion is smooth or impose constrains on the object motion to be constant in velocity or acceleration. Multiple view image tracking or prior knowledge about objects, such as size, number or shape, can also be used to simplify the process. In this work, the word "object" refers to the template image pattern being tracked (e.g. person's hair, briefcase, etc.).

Normally, tracking is seen as a main task involving several subtasks such as image segmentation for object detection, object matching and object position estimation. A myriad of algorithms has been developed to implement this subtasks but each one have their strengths and weaknesses and, over the last years, extensive research has been made in this field to find optimal tracking systems for specific applications. Many approaches of tracking techniques have been proposed in the literature, however, they are not completely accurate for all kind of scenarios and just provide good results when a certain number of assumptions are verified. Moreover, tracking methodologies that are not designed for particular applications, where specific and well established assumptions or constrains can easily be imposed, tend to be very complex. These reasons are the motivation to study and implement new tracking approaches where the introduction of soft computing techniques, such as fuzzy logic, is intended for:

- Reducing the tracking task complexity by endowing the methodology with the capability of incorporating reasoning in the same sense that human reasoning simplifies real tracking problems (e.g. most tracking problems are not complex for humans, they are indeed trivial in most situations).
- Endowing the methodology with the needed scalability in order to cope with the specific needs of different tracking problems by easily adding, changing or adapting the used fuzzy sets while maintaining its general framework.

* Corresponding author at: School of Technology and Management, Polytechnic Institute of Leiria, Morro do Lena, Alto do Vieiro, Apartado 4163, 2411-901 Leiria, Portugal. Tel.: +351 244820300.

E-mail addresses: nuno.lopes@ipleiria.pt (N.V. Lopes), pcouto@utad.pt (P. Couto), aranzazu.jurio@unavarra.es (A. Jurio), pmelo@utad.pt (P. Melo-Pinto).

0950-7051/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.knosys.2013.09.014

The presented methodology intents to be an ease and feasible general framework for object tracking that can easily be adapted to specific applications or problems.

The remainder of this paper is organized as follows. In Section 2 the definition and a review of object tracking is presented. A general description of fuzzy set theory is presented in Section 3. The proposed approach is presented in Section 4. A possible implementation of the proposed approach is given at Section 5. Section 6 shows the experimental results to illustrate the effectiveness of the proposed approach and a comparative study with well known tracking approaches is performed. Finally Section 7 presents the final conclusions and future directions.

2. Object tracking

Object tracking can be described as the problem of estimating the trajectory of an object as it moves around a scene. Although this general concept is almost consensual, the specific definition of tracking can change in the literature. Nevertheless, tracking systems must address two basic processes: figure-ground segmentation and temporal correspondences [8]. Figure-ground segmentation is the process of extracting the objects of interest from the video frame. Segmentation methods are applied as the first step in many tracking systems and therefore they are a crucial task. Object detection can be based on motion [9,10], appearance [11–13], etc. Temporal correspondence concerns to the association of the detected objects in the current frame with those in the previous frames defining temporal trajectories [14,15].

In [16], tracking is described as a motion problem and a matching problem. In this work, the motion problem is related with the prediction of the object location in the next frame. The second step is similar to the explained above. However, [1,17] present a wider description of tracking with three steps: detection of interesting objects, tracking such objects and analysis of object tracks to recognize their behavior. In [18] this behavior analysis is seen as a further interpretation of tracking results.

The selection of the most suitable feature to track is a critical role in tracking systems. The uniqueness of such feature provides an easy way to distinguish the object in the scene along time. Properties as intensity, color, gradient, texture or motion are commonly used to perform object tracking.

According to its properties, object tracking could be categorized in three groups: point, kernel and motion based approaches.

2.1. Point based tracking

Point based tracking approaches are suitable for tracking objects that occupy small regions in an image or they can be represented by several distinctive points. These points must be representative of the object and invariant to changes in illumination, object orientation and camera viewpoint. Points denoting significant gradient in intensity are preferred and commonly used by different detectors such as Harris [19], KLT [20] and SIFT [21]. To deal with the point correspondence problem between frames, deterministic constraints such as proximity, maximum velocity and small velocity change could be used. An alternative is to use statistical methods such as Kalman or particle filters. KLT and SIFT approaches provide internal methodologies to address the correspondence problem. Scale-invariant feature transform (or SIFT) is a well-known algorithm for object recognition and tracking. Interesting points are extracted from the object to provide a set of descriptors. These descriptors must be detected on the new image even among clutter, partial occlusion and uniform object scaling and rotation. In order to reduce computational time consumption, a research region in the next frame is defined according the last known location or based in a motion model of the object. However this method would typically not work with deformable or articulated objects since the relative positions between the descriptors differ from the original representation. To overcome this limitation an update scheme could be used and the object descriptors are recomputed after a predefined elapsed time.

2.2. Kernel based tracking

In this approach it is required a template or an appearance model of the object. Template tracking consists of searching in the current image for a region similar to the object template. The position and, consequently, the object matching between two consecutive frames is achieved by computing a similarity measure such as the cross-correlation. The cross-correlation concept is presented in [13]. Instead of templates, other object representations can be used for matching, for instance, color, color statistics, texture or histogram based information. The mean shift tracking algorithm is an efficient approach to tracking objects whose appearance can be described using histograms [22]. This iterative method maximizes the appearance similarity by comparing the histograms of the object and the region around the predicted object location. The Bhattacharya and Kullback–Leibler distances are commonly employed to measure the similarity between the template and the current target region. It fails in the case of occlusions and quick appearance changes.

2.3. Motion based tracking

This group of approaches perform tracking based on displacement or optical flow of image pixels. The optical flow of a pixel is a motion vector represented by the translation between a pixel in one frame and its corresponding pixel in the following frame. This computation has been proved to be difficult to achieve due to issues such as the brightness constancy assumption and the aperture problem. The classic tracking algorithm Kanade–Lucas–Tomasi (KLT) was firstly proposed by Lucas and Kanade in 1981, being perfected by Tomasi and Kanade in 1991 and explained in detail by Shi and Tomasi in 1994 [20]. The method proposed by Lucas and Kanade computes the optical flow for each pixel of an image, while the method proposed by Tomasi and Kanade known as KLT, extracts optimal points in the image and then computes the optical flow on the subsequent images to only this subset of points. The KLT is a complete method that provides a solution for two problems in computer vision: the problem of optimal selection of suitable points in an image and the problem of determining the correspondence between points in consecutive frames. It has little tolerance in image brightness variation and difficulty in detecting rapid object movements. Tracking moving objects can also be achieved by constructing a reference representation of the environment called background model and then finding deviations between this model and each incoming frame. A significant change between the background model and an image region denotes a moving object. This process is referred as background subtraction and represents a popular method especially under those situations with a relatively static background. An alternative approach to detect changes and, consequently the movement, between two consecutive intensity image frames $I(x,y,t)$ and $I(x,y,t-1)$ taken at times t and $t-1$, respectively, is to perform a pixelwise difference operation. Frame differencing is very adaptive to dynamic environments, but generally does a poor job of extracting all the relevant pixels, i.e., there may be holes left inside slowly moving objects.

Since the arise of fuzzy logic theory, it has been successfully applied in a large range of areas such as process control systems, automotive navigation systems, information retrieval systems and image processing. As presented beforehand, a tracking system