Increasing multi-trackers robustness with a segmentation algorithm

MARTA MARRÓN, MIGUEL ÁNGEL SOTELO, JUAN CARLOS GARCÍA
Electronics Department
University of Alcala
Campus Universitario. 28871, Alcalá de Henares. Madrid.
SPAIN
marta@depeca.uah.es   http://www.depeca.uah.es

Abstract: - Different solutions have been proposed for multiple objects tracking based on probabilistic algorithms. In this paper, the authors propose the use of an only particle filter to track a variable number of objects. The robustness and adaptability of the estimator are increased by the use of a clustering algorithm. The measurements used in the tracking process are extracted from a stereovision system, and thus, the 3D position of the tracked objects is obtained at each time step. Tracking results are presented at the end of the paper, showing the reliability of the proposals.

Key-Words: - Probabilistic, Multi-tracking, Robustness, Clustering, Particle Filters.

1   Introduction

Probabilistic algorithms in their different implementations (Multi-Hypothesis Techniques - MHT- [1], Particle Filters -PF- [2] and their diversifications [3][4]) have fully shown their reliability in estimation tasks. Nowadays these methods are widely applied to solve positioning problems in robot autonomous navigation [5].

The idea of tracking multiple objects appeared with the first autonomous navigation robot, to overcome the obstacle avoidance problem and soon probabilistic algorithms, such as PF ([6], [7]) and KF ([8]), were applied to achieve this aim. The objective is, in any case, to calculate the posteriori probability ($p(\vec{x}_t \mid \vec{y}_{1:t})$) of the state vector in the recursive two steps standard estimation process (see (1)), in which, at least, some of the involved variables are stochastic, and by means of the Bayes rule:

$$p(\vec{x}_t \mid \vec{y}_{1:t}) = \eta \cdot p(\vec{y}_t \mid \vec{x}_t) \cdot \int p(\vec{x}_t \mid \vec{x}_{t-1}) \cdot p(\vec{x}_{t-1} \mid \vec{y}_{1:t-1}) \cdot \partial \vec{x} \quad (1)$$

To solve the multiplicity problem, an expansion of the state vector to include that of all the elements to track (see (2)) was the first solution proposed in [9].

$$\vec{\chi}_t = \left\{ \vec{x}_t^1, \vec{x}_t^2, ..., \vec{x}_t^k \right\} = \bigcup_{i=1}^{k} \vec{x}_t^i \qquad (2)$$

The computational load of the resultant estimator does not allow achieving a real time execution of the algorithm for more that 4 or 5 objects [10].

Another solution for the multiple objects tracker is to use a standard estimator to track each object but, apart of the inefficiency of the final algorithm, it cannot deal easily with a dynamic number of objects.

On the other hand, in order to work with independent estimators it is necessary to include an association algorithm to assign each measurement to one of the trackers. Most of the association solutions are based on the Probabilistic Data Association (PDA) theory [11], such as the Joint Probabilistic Particle Filter (JPDAF) like in [12]. Again, the problem related to these techniques is the execution time.

In this context the authors propose in [13] another solution to the multi-tracking problem based on a single PF whose multimodality is exploited to perform the multi-tracking task. The algorithm is called Extended Particle Filter (XPF). This solution has been tested in complex indoor environments with sonar [13] and stereo-vision data [14] with good results. A clustering algorithm is used in this point to organize the measurements and the information involved in the estimation process, and to increase the robustness and reliability of the final estimator.

In this paper, a general revision of the tracking system developed, based on stereovision measurements, is presented, a slight description of the XPF developed in [13] is then included; the functionality of the two proposed clustering methods is presented; and finally some results and conclusions obtained from the segmentation processes applied to the probabilistic estimator are exposed.

2   General description

Most of the tracking systems developed in the last years for autonomous navigation and surveillance applications are based on visual information.

Measurements obtained from a vision system are generated both by the environment and the objects to track. To use this information in the tracking task a process to distinguish the data set that comes from the

objects to track is needed. This pre-processing algorithm is especially important to achieve a real time execution, due to the big amount on measurements included in the vision image.

Fig. 1 shows the functional block diagram of the global tracking system in which a classification pre-process has been included to organize the measurements that come from the vision system in two groups:

- those that inform about the position of the objects in the image.
- those that are generated by the elements in the environmental structure (such as floor and walls).

With the mentioned functionality, the classification pre-process behaves as an obstacle detection module.

As it has been exposed in Fig. 1, measurements classified as environmental structure ones can be used to implement a partial reconstruction of the environment in which the tracked objects and the robot itself are moving. This process, that is independent of the taking application, is still under development by the authors.

3 The classification algorithm

Developing obstacle avoidance tasks requires 3D information about the objects position in the environment. To obtain these data, a stereo-vision system has been used as shown in Fig. 1.

Fig. 2 shows the process used to detect and extract the array of measurements from the stereo-vision system, which is explained in the following text.

The stereo-vision system proposed is formed by two black and white digital cameras whose acquisition is synchronized. The 3D points are obtained from each pair of frames, by means of an epipolar technique.

As the amount of information in each image is too big to be computed it in real time, matching pixels for the epipolar equation are taken among the edge image obtained applying a canny filter to the input frame.
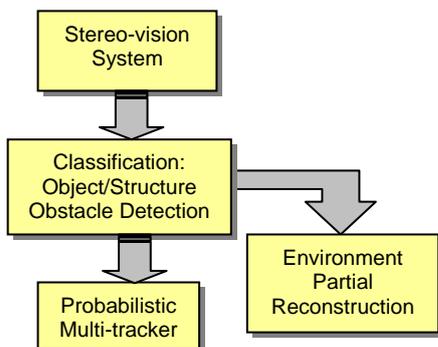

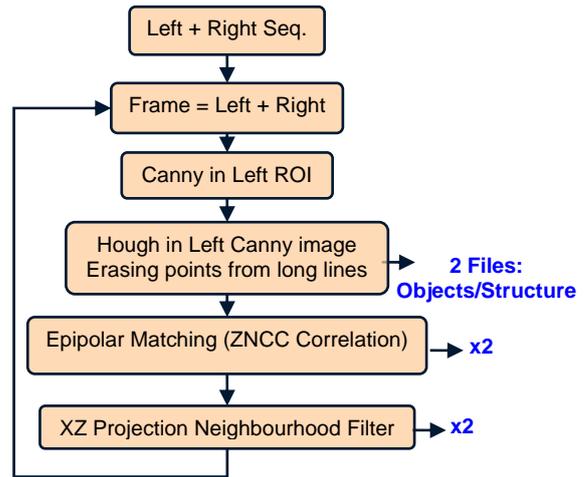Figure 1. Block diagram of the global tracking system

Figure 2. Block diagram of the stereo-vision classifier and object detector

With the pixels in the edge image, the classification process is performed. Different strategies have been tested to achieve the best results classifying the measurement set, both in execution time and in reliability.

Edges corresponding with environmental structures have the common characteristics of forming long lines, and thus, the Hough transform has been selected as the best procedure to erase the pixels from the edge image that should not be part of the object measurement array.

The execution time of the developed classifier is around 60ms, which is an acceptable solution to implement a real time acquisition image process of 15fps to 33fps.

Different solutions are still under development to achieve a more robust classification process.

4 The multi-tracking algorithm

As mentioned in the introduction, a PF is used as a multimodal tracker to estimate the position and speed of the objects in the environment, from the measurement array obtained in the classification process.

PF is a particularization of the Bayesian estimator in which the densities related to the posterior estimation (also called belief) is discretized. A detailed description of the PF functionality and mathematical base can be found in [2] and [4].

As the state vector is not discretized, like it is in most of the rest Bayes filter implementations, the PF is more exact in its estimation than the KF or estimators based on a grid (MonteCarlo estimators). Moreover, due to the same reason, the computational load of this Bayes filter form is smaller than in other implementations, and thus, more adequate to implement real time estimation.

Finally, PFs include an interesting characteristic for multi-tracking applications: the ability of representing multiple estimation hypotheses with an only algorithm, through the multimodality of the belief. This facility is not available in the optimal implementation of the Bayes estimator, the KF. For all these reasons, the PF has been thought as the most appropriated algorithm to develop a multi-tracking system.

4.1 The XPF

Most of the solutions based on a PF presented in the related scientific literature do not use the multimodal character of PF to implement the multi-tracker. These alternative solutions, mentioned in the introduction of this paper, are however less efficient in execution time.

In order to adapt the standard PF for its use tracking a variable number of elements, some modifications must be included in the basic algorithm. In [13] and [14], the authors propose what is called the extended particle filter (XPF), an adaptation of the standard PF for tracking a variable number of elements.

The XPF first approach appears in [15]. The algorithm proposed there was nevertheless not used in real tracking tasks because it is not robust enough. The XPF proposed by the authors in [14] includes a clustering algorithm to improve the behaviour of the first extended PF, giving as a result the XPF-CP process, whose functionality is shown in Fig. 3.

The clustering algorithm whose functionality is presented next in this document, organize the data in clusters that represent all objects in the scene.

These clusters are then used in the tracking algorithm. With the clustered data two important innovations are included in the standard PF in order to facilitate the multi-tracking process:

a) *Re-initialization step:* M from the N total number of particles that conform the belief distribution in the PF are directly included in this step previous to the prediction step. With this step, measurements related to newly appearing objects in the scene have a representation in the estimator belief.

To improve the robustness of the estimator, the inserted data are not selected randomly from the array but from its segmentation. Choosing measurements from every cluster ensures a probable representation of all objects in the scene, and therefore the robustness of the multi-tracker is increased.

Thanks to this re-initialization step the posterior distribution dynamically adapts itself to the different objects present in the scene.

b) *Resampling step:* This step is also modified from the one in the standard PF. On one hand, only (N-M) samples of the belief have to be extracted in this step, as the M resting particles would be inserted in the re-initialization step.

On the other hand, the clustering process is also used in this step, because the importance sampling function used to weight the particles is obtained from the similarity between each one of them and the centroid of the clustered input data set.

Using the cluster centroids to weight the particles related to the newly appeared objects is increased, improving the robustness of the final estimator.

Without the clustering process, the solution proposed in [15] rejects the hypothesis related to new objects, and thus, the multimodality of the PF cannot be exploited.

The robustness problem has been the main reason in the researching community for not using the PF multimodal capability. With the algorithm proposed in this paper, a new solution for multiple tracking using an only estimator is designed.
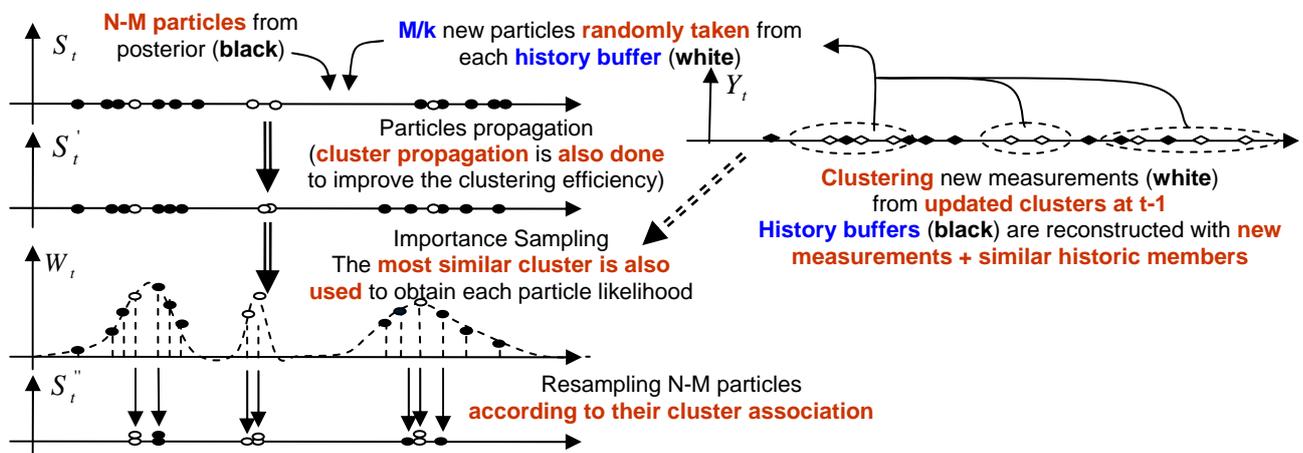

Figure 3. Description of the XPFCP functionality