Expert Systems with Applications 41 (2014) 6646–6661

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Review

# Context aided pedestrian detection for danger estimation based on laser scanner and computer vision

Fernando García [a,*], Jesús García [b], Aurelio Ponz [a], Arturo de la Escalera [a], José María Armingol [a]

[a] Intelligent Systems Lab, University Carlos III de Madrid, Leganés, Spain
[b] GIAA Research Group, University Carlos III de Madrid, Colmenarejo, Spain

ARTICLE INFO

Article history:
Available online 14 May 2014

Keywords:
Data fusion
Context
Laser scanner
Computer vision
Pedestrian detection

ABSTRACT

Road safety applications demand the most reliable sensor systems. In recent years, the advances in information technologies have led to more complex road safety applications able to cope with a high variety of situations. These applications have strong sensing requirements that a single sensor, with the available technology, cannot attain. Recent researches in Intelligent Transport Systems (ITS) try to overcome the limitations of the sensors by combining them. But not only sensor information is crucial to give a good and robust representation of the road environment; context information has a key role for reliable safety applications to provide reliable detection and complete situation assessment. This paper presents a novel approach for pedestrian detection using sensor fusion of laser scanner and computer vision. The application also takes advantage of context information, providing danger estimation for the pedestrians detected. Closing the loop, the danger estimation is later used, together with context information, as feedback to enhance the pedestrian detection process.



## 1. Introduction

Most of the accidents in roads are connected with human errors. Wrong decision making or driver inattentions are the two main errors that cause traffic accidents. These kind of errors are related to human nature and cannot be eliminated, although efforts can be made to decrease them. Recent researches in Intelligent Vehicles have focused on using the advances in information technologies to prevent these errors. One example of these kind of applications are the Advanced Driver Assistance Systems (ADAS), which try to warn driver in case of hazardous situations.

Sensor trustability is one of the main issues when dealing with road safety applications. The ADAS systems need the most reliable set of sensors to fulfill the requirements of these demanding applications. Thus, to accomplish such a difficult task, it is mandatory to combine different information sources so we can overcome the limitations of each independent sensor. Here is where context information can be helpful for both increasing the accuracy of each sensor independently and providing a new information source to improve the performance of the Fusion process.

In this paper, a novel approach, based on data fusion for pedestrian detection is presented. It makes use of state of the art pedestrian detection algorithms for both laser scanner and computer vision and Joint Probabilistic Data Association (JPDA) for data association, which was specially adapted to be used in a real time automotive environment. The application also takes advantage of some available contextual information, (including static knowledge as well as some online information). Thus, by combining a strong association technique and context information, classical ADAS detection is enhanced. Three main sensors were used for this application:

**Laser scanner**. Recent researches have focused on the use of this well-known sensor in automotive applications. Its robustness and reliability has been proved in different test and contests (e.g. DARPA Grand and Urban Challenge) (Defense & Agency, 2006; Iagnemma & Buehler, 2006; Buehler et al., 2005; Iagnemma and Buehler, 2007.
**Computer vision**. It is a common topic in Intelligent Vehicles research, and nowadays it can be found in commercial systems.
**Inertial sensor**. It is an improved GPS with inertial correction that allows accurate estimation, not only of position and velocity, but also of Euler angles, acceleration, etc. This information about vehicle state is added as dynamic contextual knowledge to enhance the fusion process.

* Corresponding author. Tel.: +34 6248325; fax: +34 916249430.
E-mail addresses: fegarcia@ing.uc3m.es (F. García), jgherrer@inf.uc3m.es (J. García), apv@ing.uc3m.es (A. Ponz), escalera@ing.uc3m.es (A. de la Escalera), armingol@ing.uc3m.es (J.M. Armingol).

http://dx.doi.org/10.1016/j.eswa.2014.04.034
0957-4174/

## 2. Sensor fusion in vehicles. State of the art

Fusion approaches in vehicle safety can be divided in relation to the processing architecture employed. In Intelligent Vehicle applications, this is usually the way of dividing them. Most of these approaches take advantage of complementary properties of the available sensors mainly at fusion level 1 (Object Assessment), focusing on the detection and tracking of the different actors involved in road environments:

In **feature vector fusion** approaches, some preprocessing is performed for each sensor to create a set of features for each one. These individual sets are combined to create a compound set that is used to perform the obstacle detection and classification. In Premebida, Ludwig, Silva, and Nunes (2010) and Premebida, Ludwig, and Nunes (2009)features are extracted for each sensor independently and a new data set is created; authors present different approaches whether combining or not the different features of the different sensors and comparing results. The final classification after fusion is compared with alternative methods such as Naïve Bayes, Gaussian Mixture Models, Neural Networks. In Zhao, Chen, Zhuang, and Xu (2014) multiple feature fusion is performed, based on different feature extraction approach for vision approaches i.e. Support Vector Machine (SVM), Naïve Bayesian and Minimum Distance Classifier.

**Decentralized fusion architecture** based approaches perform detections and classifications for each sensor independently and a final stage combines the detections according to the certainty of the detections and sensors trustability. Spinello and Siegwart (2008) uses Adaboost vision based pedestrian detection and Gaussian Mixture Model classifier (GMM) for laser scanner based pedestrian detection, finally a Bayesian decisor is used to combine detections of both subsystems. In Premebida et al. (2009) pedestrians are detected by a laser scanner using multidimensional features that describe the geometrical properties of the detections, and features of Histograms of Oriented Gradients (HOG) with SVM classification for computer vision based pedestrian detection; Bayesian modeling approach is used for the final fusion. Premebida and Nunes (2013) takes advantage of lidar ROI detection, computer vision and contextual information from digital map to provide high level fusion based on a Bayesian approach.

Generally, data fusion approaches among Intelligent Vehicles researches use data from laser scanner to detect regions of interest (ROI), and computer vision to classify among different obstacles that can be found. Such as Ludwig, Premebida, Nunes, and Ara (2011) where HOG features combined with SVM approach provide pedestrian detection and Pérez Grassi, Frolov, and Puente León (2010) based on Invariant Features, again with SVM, to perform the vision based pedestrian detection. These approaches take advantage of the trustability of the laser scanner for obstacle detection but fusion is limited to speed up the process by detecting robust ROIs. Thus, the information added by the fusion process is limited and it could barely be considered real data fusion.

Some other fusion approaches in the scope of the Intelligent Transport Systems researching field take advantage of different sensors' properties in systematic approaches, although no explicit data fusion processes or algorithms are included in the works: Broggi, Cerri, Ghidoni, Grisleri, and Jung (2008) uses information from laser scanner to search those zones of the environment where pedestrians could be located and visibility is reduced (e.g. space between two vehicles) and performs detections using a vision approach. In Bohmlander, Doric, Appel, and Brandmeier (2013) mono camera and a capacitive sensor are used for pedestrian detection, and in Garcia, Cerri, Broggi, de la Escalera, and Armingol (2012) radar and computer vision by means of featured based optical flow, are used for vehicle overtaking detection.

Fusion is also widely used for vehicle positioning, based on the fusion of different positioning and tracking techniques, such as GPS, inertial measurements and odometry. By fusing GPS signal with inertial measurements, problems regarding to GPS signal loss can be overcome, improving the precision of the positioning systems, as presented on Bhatt, Aggarwal, Devabhaktuni, and Bhattacharya (2014) and Martí et al. (2012).

Within the scope of data fusion, target tracking is one of the main aspects, several works have been presented that enhance the detections by the use of advance tracking procedures in Intelligent Transport Systems (ITS) and expert systems researching fields. By combining strong classification algorithms and trustable tracking procedures, reliable pedestrian detection can be performed: In Li, Xu, Goodman, Xu, and Wu (2013) background-foreground identifications enhances the detection, the later combination with Camshift tracker and a Kalman Filter (KF) allows trustable pedestrian detection and tracking. In Fan, Mittal, Prasad, Saurabh, and Shin (2013) deformable part models and KF are used for visual based pedestrian detection and tracking with JPDA association technique. In Schneider and Gavrila (2013) comparative between Extended KF and Interacting Multiple Models (IMM) tracking methods is provided for stereovision based pedestrian detection. Works presented on Sánchez, Patricio, García, and Molina (2009) and Gómez-Romero, Patricio, García, and Molina (2011)provide tracking procedures for surveillance applications, taking advantage of the context information in complex scenarios.

Context information can aid safety applications by both adding inference development (i.e. checking the consistence of the detections with the previously defined model) and adding explanatory aspects when the inference is consistent with the context. Some works commented before already take advantage of this useful information for pedestrian detection based on digital maps (Premebida & Nunes, 2013), video surveillance applications (Sánchez et al., 2009; Gómez-Romero et al., 2011) or for vehicle positioning (Martí et al., 2012). This work uses contextual information to complete the available sensor data, representing accurately the current situation of the vehicle and the objects surrounding it, affecting to the predictable behavior of the driver according to the safety regulations.

Finally it is important to remark that traffic and road safety applications are common topic within expert systems, the availability of modern IT technologies allows the development of modern and advance algorithms that make use of the these advances to prevent road accidents or mitigate its consequences: In Castro, Delgado, Medina, and Ruiz-Lozano (2011) an expert system based on fuzzy logic is presented, designed specifically to avoid pedestrian accidents. Guo, Ge, Zhang, Li, and Zhao (2012) presents an Adaboost and SVM based system for pedestrian detection. In Conesa, Cavas-Martínez, and Fernández-Pacheco (2013) vehicles driving in opposite direction are detected by means of agent based architecture. In Jo, Lee, Park, Kim, and Kim (2014) driver drowsiness is analyzed by means of the used of driver specific biological measurements and computer vision based algorithms for eye state and blinking detection. Finally authors in Abellán, López, and De OñA (2013) provide an algorithm to analyze and identify the severity of the accidents using decision trees.

In this work, a fusion-based expert system for pedestrian detection and danger avoidance is presented. Previous works, commented above, are particular solutions developed for specific sensor inputs not taking advantage of all the contextual knowledge. Furthermore, present work tries to focus on a system-level approach, making use of information and processes performed at different levels, and taking into account the final application of the fusion process. All these aspects are not common among ITS researches. Finally, the model makes use of a JPDA approach for multiple sensors, able to overcome difficult situations in the data