Expert Systems with Applications 41 (2014) 795–810

Contents lists available at ScienceDirect

Expert Systems with Applications

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

A novel evidence based model for detecting dangerous situations in level crossing environments

H. Salmane [a,*], Y. Ruichek [a], L. Khoudour [b]

[a] IRTES-SeT, University of Technology Belfort-Montbeliard, 13 rue Ernest-Thierry Mieg, 90010 Belfort cedex, France
[b] CETE SO, Center for Technical Studies of South West, 1 Avenue du Colonel Roche, 31400 Toulouse, France

A R T I C L E I N F O

Keywords:
Level crossing safety
Object detection and tracking
Energy vector
Kalman filter
Hidden Markov Model
Credibility model

A B S T R A C T

Considered as a weak point in road and railway infrastructure, level crossings (LC) improvement safety became an important field of academic research and took increasingly railways undertakings concerns. Improving safety of people and road-rail facilities is an essential key element to ensure a good operating of the road and railway transport. For this purpose, road and railway safety professionals from several countries have been focused on providing level crossings as safer as possible. Many actions are planned in order to exchange information and provide experiments for improving the management of level crossing safety and performance.

This paper aims to develop a video surveillance system to detect, recognize and evaluate potentially dangerous situations in level crossing environments. First, a set of moving objects are detected and separated using an automatic clustering process coupled to an energy vector comparison strategy. Then, a multi-object tracking algorithm, based on optical flow propagation and Kalman filtering correction with adaptive parameters, is implemented. The next step consists on using a Hidden Markov Model to predict trajectories of the detected objects. Finally, the trajectories are analysed with a particular credibility model to evaluate dangerous situations at level crossings. Real data sets are used to test the effectiveness and robustness of the method. This work is developed within the framework of PANsafer project, supported by the French work program ANR.

1. Introduction

1.1. Background

In compliance with the French call for tender ANR-VTT, PANsafer's main objective is to actively contribute on reducing level crossing accidents. To achieve this objective, we implement an intelligent video surveillance system that allows automatic recognition and evaluation of critical situations in level crossing environments. Before giving details about the video surveillance system, let us start by summarizing the PANsafer global context in which the whole technological strategy for reducing level crossing accidents is developed. PANsafer technological strategy can be summarized in three main points:

– One surveillance system dedicated to the detection of potentially dangerous situations thanks to video sensing and image processing (this is what we present in this paper) and one equipment of communication whose role is to send to the users approaching the LC the status of the LC. These two equipments are installed in the LC premises;
– The communication system will take the information on the dynamic status of the LC (status of the barriers, obstacle or not, level of dangerousness) to the road navigation terminals installed in the cars;
– A dynamic restitution of the status of the LC, with the mother tongue of the driver, is operated by the terminal of road navigation.

Since the surveillance system will mainly be used in a security context, the specific events occurring at the LC must be detected in real time and sent timely and safely to the users. So, for the surveillance system we try to find a trade-off between the accuracy of detection and a processing time not to prohibitive. We will see later in the paper what are the chosen options in terms of processing capabilities for the system and in conclusion one can find the performances of the system, so far, in terms of processing time.

The proposed video surveillance system is composed of five main software components. The first one aims to robustly detect and separate moving objects crossing the level crossing. The

* Corresponding author.
E-mail addresses: houssam.salmane@utbm.fr (H. Salmane), yassine.ruichek@utbm.fr (Y. Ruichek), louahdi.khoudour@developpement-durable.gouv.fr (L. Khoudour).

0957-4174/$ - see front matter
http://dx.doi.org/10.1016/j.eswa.2013.08.010

second one consists on performing object tracking using a Gaussian propagation optical flow based model. The objective of the third component is to segment moving objects into different regions, basing on the optical flow of the objects' pixels. In the fourth part of the system, for each segmented region, we use a real-time Hidden Markov Model to predict the ideal trajectory such as to avoid potentially dangerous situations. Finally, an analysis of all predicted trajectories in order to evaluate the degree of collision related to each moving object is carried out using evidence theory.

Methods based on background subtraction (Panahi, Sheikhi, & Gheissari, 2008) are often used to detect moving objects, assuming that the image is recorded by a static video sensor. For many years, this concept was used by many researches in order to perform background removal algorithms used in complex models. Thus, many approaches appeared in order to achieve the highest possible detection accuracy in real time, examples are the Independent Components Analysis (Fakhfakh et al., 2010), Histogram of Oriented Gradients (Bertozzi, Broggi, Rose, & Felisa, 2007), Wavelet (Toreyin, Cetin, Aksay, & Akhan, 2005), Sequential Kernel Density approximation (Han, Comaniciu, & Davis, 200), and Eigen backgrounds (Oliver, Rosario, & Pentland2000). However, to be efficient, these techniques require further development to distinguish between detected objects. For that reason, equivalence table based algorithms (Suzuki, Isao, & Noboru, 2003) is one of the solutions that can be used to separate the detected objects. These algorithms are used to label different connected clusters of pixels (Singh, Dunga, Shekhar, & Vohra, 2010). The problem of this technique is that it can only separate colored and textured objects. One can also use data association techniques, in which feature observations are associated and filtered to existing or new tracks (Kirubarajan & Bar-Shalom, 200). The performance of these methods depends on the distance between the objects and their size. Another important method from the literature consists on using sampling methods (Smith, 2004). In these methods, the multi-objects tracking model is resolved by a probabilistic exclusion principle (MacCormick & Blake, 1999). This approach is very accurate because it is mathematically rigorous, significantly reduces computation, and tracks robustly, but it requires a lot of complex sampling strategies (Deutscher, Blake, & Reid, 2000; Poon & Fleet, 2002; Sullivan, Blake, Isard, & MacCormick, 1989) to resolve the problem of multi object tracking.

Video tracking process can start when there are enough detected pixels belonging to moving objects. In the literature, there are two major kinds of methods to perform visual tracking. The first one is based on target representation and localization (Blob, kernel and contour tracking), and the second type of methods uses filtering (Kalman and particle filters) and data association. For example, Francois (2004) proposed a Blob tracking algorithm based on object segmentation in successive frames. This algorithm is accurate for tracking object in real-time but the major drawback is that this method uses the result of the background subtraction and blob detection modules without correction, therefore the errors generated at these stages are propagated to the tracking stage and affect the tracking performance. Li, Zhang, Shen, and Jiancheng (2010) use an adaptive Kalman filter combined with mean shift technique, and the center of each target candidate is tracked in a normalized color distribution. This method has a stable performance to track a moving object under certain real-world complex situations, such as occlusion, fast moving objects, and sudden changes in velocity of moving objects. However, the main problem of this kind of algorithms is that filtering precision maneuver will decrease when a predicted target of a moving object does not occur. Blake and Isard (1998) work is based on boundary objects tracks using snakes and temporal fusion by Kalman filter. Active contours have been widely used in tracking but this kind of methods is inherently slow and their accuracy depends on the convergence criteria used in the energy minimization technique. Yang, Duraiswami, and Davis (2005) tries to calculate the state-space distribution of the tracked object using a particle filter. This approach gives good result for occlusion problems, but a significant computation time can take place in this technique and it is hard to determine the optimal number of particles especially when the number of objects tracked increases. Each of these algorithms has advantages and drawbacks, but most of them cannot track correctly all the pixels from the given object.

Optical flow dataset obtained by tracking moving objects can be used to predict object trajectories in real time. Gaussian Mixture Model (GMM) (Bashir, Khokhar, & Schonfeld, 2006), Hidden Markov Model (HMM) (Bashir, Khokhar, & Schonfeld, 2007) and some of its extensions, such as hierarchical Hidden Markov Model (HHMM) (Nguyen, Phung, Venkatesh, & Bui, 2005) and couple Hidden Markov Model (CHMM) (Natarajan & Nevatia, 2007) are usually used for representing and recognizing objects' trajectories. However, these methods need a high number of statistical measures to be effective. Hence, it is difficult to apply these methods in real time. It was also important to implement a trajectory optimization process to predict trajectories that minimize the degree of danger related to the moving objects. The Dempster–Shafer theory (Dubois, 2006; Yager, Fedrizzif, & Kacprzyk, 1994) is an interesting concept for analyzing and evaluating danger situations. It allows to obtain a degree of risk that takes into account the evidence from different sources of danger.

1.2. Overview of the proposed model

This paper presents a new model to evaluate and recognize potential dangerous situations in a level crossing environment. Fig. 1 illustrates the synopsis of the proposed video surveillance system. The method starts by detecting and separating all moving objects that enter into a given surveillance zone. For that, an energy vector comparison strategy combined with a background subtraction technique is proposed. This method consists on clustering moving pixels by comparing a specific energy vector associated to each target and each moving pixel. Given a detected object, a new method that improves significantly the tracking performance of each pixel affected by motion within a detected object is proposed. This is achieved by a Harris points based optical flow propagation technique, followed by a Kalman filtering based correction. The whole tracking process (detection, optical flow, filtering) is discussed and evaluated in details by the authors in Salmane, Ruichek, and Khoudour (2011, 2012). In order to present the entire danger evaluation system, we give in this paper the summary of each step of the tracking process with different intermediate results. Once the tracking process is achieved, for each moving object, a real-time Hidden Markov Model is assessed to predict trajectories that minimize the degree of danger of the target in a level crossing environment. The predicted trajectories only depend on the optical flow of segmented regions in the current image. Indeed, each object may be characterized by different trajectories, where each trajectory is associated to a region extracted by segmenting optical flow of the moving pixels. The trajectory prediction procedure takes into account the infrastructure geometry of the level crossing environment. To estimate the degree of danger related to each object, each trajectory is analyzed considering different sources of danger (position, velocity, acceleration, distance). All the information obtained from the sources are fused using Dempster–Shafer theory. Once the analysis of the trajectories is completed, the degree of danger related to the object is simply the maximum of all the degrees of danger related to the trajectories corresponding to the different regions of the object.

The remaining of the paper is organized as follows. Section 2 is dealing with object detection and separation. In Section 3, the