Chapter 5
Adding Context Information to Video Analysis for Surveillance Applications

Solmaz Javanbakhti
Eindhoven University of Technology, Netherlands

Willem P. Sanberg
Eindhoven University of Technology, Netherlands

Xinfeng Bao
Eindhoven University of Technology, Netherlands

D.W.J.M. (Denis) van de Wouw
Eindhoven University of Technology, Netherlands

Ivo Creusen
Eindhoven University of Technology, Netherlands

Gijs Dubbelman
Eindhoven University of Technology, Netherlands

Lykele Hazelhoff
Eindhoven University of Technology, Netherlands

Svitlana Zinger
Eindhoven University of Technology, Netherlands

Peter H.N. de With
Eindhoven University of Technology, Netherlands

ABSTRACT

Smart surveillance systems become more meaningful if they both grow in reliability and robustness, while simultaneously offering a higher semantic level of understanding. To achieve a higher level of semantic scene understanding, the objects and their actions have to be interpreted in the given context, so that the extraction of contextual information is required. This chapter explores several techniques for extracting the contextual information such as spatial, motion, depth and co-occurrence, depending on applications. Afterwards, the chapter provides specific case studies to evaluate the usefulness of context information, based on: (1) region labeling of the surroundings of objects, (2) motion analysis of the water for moving ships, (3) traffic sign recognition for safety event evaluation and (4) the use of depth signals for obstacle detection. The chapter shows that the previous cases can be solved in an improved way with respect to robustness and semantic understanding. Case studies indicate up to 6.8% improvement of reliable correct object understanding and the novel possibility of labeling scene events as safe/unsafe depending on the object behavior and the detected surrounding context. In this chapter, it is shown that using contextual information improves automated video surveillance analysis, as it not only improves the reliability of moving object detection, but also enables scene understanding that is far beyond object understanding.

INTRODUCTION

Automatic surveillance video understanding is one of the ultimate application fields of computer vision research. The objective of visual surveillance systems is not only to use cameras instead of human eyes, but also to perform surveillance automatically using video analysis. One of the key objectives for automatic surveillance is to selectively guide the attention of human operators to potentially suspicious activities. The important arguments for doing so are twofold. First, automation reduces labor costs so less human operators have to observe many video feeds simultaneously, a job that is both error-prone and tedious. Second, with the huge amount of information contained in parallel viewing of video channels, the chance of missing an important event in one of the many surveillance videos is high. Smart surveillance systems should be able to at least detect and track moving objects, classify these objects and interpret their activities. A large number of surveillance systems have been proposed in recent years. These systems still need improvement in terms of reliability and robustness with respect to event interpretation and a real semantic understanding of scenes. These improvement points can be realized by adding additional information about those objects and/or scenes, so that a better classification and understanding is achieved. This extra information is typically the context of the behavior of objects or of the scene. This chapter aims at exploiting contextual information in two ways:

- To help to better interpret events based on object behavior with higher reliability and robustness.
- Obtaining a higher semantic level of scene understanding by adding contextual information about the scene itself.

Although it is present in scenes, the automated interpretation of the events and associated object detection in a monitored space is typically completely based on object detection and recognition, while the contextual information (e.g. about the surroundings of the objects) is overlooked. For example, a car detected on a parking place is a normal situation, whereas a car standing on tramway rails is a reason to raise an alarm. In this example, the rails are static object information from the surroundings acting as contextual information. In general, context can be applied at different levels: the involved features of an object usually at the pixel level, information about the object itself, and at the level of scene understanding (e.g. event classification and event detection). Challenges here include several aspects. First, what kinds of algorithms are needed for extracting the additional information from a surveillance video in order to contribute to the semantic meaning? Second, the ways for including the context information at the various levels is another important challenge. However, this chapter considers that object detection and eventually event detection should not happen in isolation, i.e., the process of recognizing one object in a scene can be influenced by the presence of additional information such as motion- or depth-based features, presence of other objects, as well as by the semantic context of the scene.

This chapter intends to show that using contextual information enables:

- A higher robustness of the object detection considered, and
- The automated analysis of complicated traffic surveillance scenarios that were previously not possible using conventional object classification techniques.

Although there is no standard view on how context information should be classified, it is evident that such information can contribute to various analysis levels for a surveillance system. This chapter explores the following aspects of contextual information:

Related Content

Exploring SMEs Adoption of Broadband in the Northwest of England

Boumediene Ramdaniand Peter Kawalek (2008). *Handbook of Research on Global Diffusion of Broadband Data Transmission (pp. 504-523).*

www.irma-international.org/chapter/exploring-smes-adoption-broadband-northwest/20459

Optimizing the Traffic of Voice Core Domain in UMTS Network through RNC Re-Homing

Ye Ouyangand Hosein Fallah (2012). *Research, Practice, and Educational Advancements in Telecommunications and Networking (pp. 227-240).*

www.irma-international.org/chapter/optimizing-traffic-voice-core-domain/62768

Tutorial for Space-Time ICI Parallel Cancellation Techniques for OFDM Systems

Hen-Geul Yeh (2023). *International Journal of Interdisciplinary Telecommunications and Networking (pp. 1-14).*

www.irma-international.org/article/tutorial-for-space-time-ici-parallel-cancellation-techniques-for-ofdm-systems/322788

Need of Algorithm Selection in Next Generation Optical Networks

P. Selvarajand V. Nagarajan (2018). *International Journal of Business Data Communications and Networking (pp. 81-92).*

www.irma-international.org/article/need-of-algorithm-selection-in-next-generation-optical-networks/193574

Secure Neighbor Discovery Protocol: Review and Recommendations

Junaid Latief Shah (2019). *International Journal of Business Data Communications and Networking (pp. 71-87).*

www.irma-international.org/article/secure-neighbor-discovery-protocol/216432