Chapter 12
Object Analysis with Visual Sensors and RFID

Gour C. Karmakar
Monash University, Australia

Laurence S. Dooley
The Open University, UK

Nemai C. Karmakar
Monash University, Australia

Joarder Kamruzzaman
Monash University, Australia

ABSTRACT

Object analysis using visual sensors is one of the most important and challenging issues in computer vision research due principally to difficulties in object representation, segmentation, and recognition within a general framework. This has motivated researchers to investigate exploiting the potential identification capability of RFID (radio frequency identification) technology for object analysis. RFID however, has a number of fundamental limitations including a short sensing range, missing tag detection, not working for all objects, and some items being just too small to be tagged. This has meant applying RFID alone has not been entirely effective in computer vision applications. To address these restrictions, object analysis approaches based on a combination of visual sensors and RFID have recently been successfully introduced. This chapter presents a contemporary review on these object analysis techniques for localisation, tracking, and object and activity recognition, together with some future research directions in this burgeoning field.

INTRODUCTION

The innate desire for humans to have their computers and machines do exactly whatever they are able to do has, to date remained largely unfilled. While humans can effortlessly analyse and recognise an object and its activity or event, computer-based automatic systems still remain stubbornly in an embryonic state. This has lead to computer vision being one the most challenging research topic areas over the past few decades. The application of computer vision is now however, rapidly in-

DOI: 10.4018/978-1-4666-1616-5.ch012

creasing due to the advancement of software and hardware technologies, including RFID (*radio frequency identification*), digital image analysis, artificial intelligent and robotic vision systems.

Computer vision has been extensively applied in a wide range of diverse domains. These include, but are not limited to: robotic vision (Hsiao et al., 2008), autonomous car and mobile robot (Milanés et al., 2010), event detection, surveillance and tracking (Yilmaz et al., 2006; Lee et al., 2007), airport identification from aerial photographs, object-based image identification and retrieval (Veltkamp & Tanase, 2000), object and activity recognition (Smith et al., 2005; Wu et al., 2007; Bashir et al. 2007), behaviour prediction (Skinner, 1953; Pentland, 1999; Chen et al., 2004), object-based image and video coding (ISO/IEC 14496-2, 1999), criminal investigation using video footage analysis, computer graphic, and medical diagnoses (cancerous cell detection, segmentation of brain images, skin treatment, and intrathoracic airway trees) (Paul & Paul, 1993; Pham & Prince, 1999; Liu et al., 1997). An essential preprocessing step in all the aforementioned real-world examples is object detection and separation (segmentation), which both involve the core element of image analysis to retrieve and understand an event's information in real-time (Computer vision, 2011). Image analysis, typically includes image segmentation, feature extraction, and object classification (Baxes, 1994).

Image segmentation is the process of separating mutually exclusive homogeneous regions of interest (i.e., having similar pixel intensities) from other regions in an image. In reality, most natural objects are not homogeneous and comprise numerous objects in a hierarchical order. Furthermore, there are generally a huge number of objects and a myriad of variations amongst them. This makes object segmentation in an image an interesting if intractable task as it somewhat contradicts the above definition of object-based image segmentation. This is because there is no universally accepted standard definition of image segmentation. The properties of an object to be segmented depend on its applications and human perception, so segmenting an object is in essence an ad hoc process, which depends on the emphasis given to the set of desired properties and a trade-off among them (Paul & Paul, 1993; and Karmakar et al., 2001). To reduce the complexity of detecting and separating an object and capturing its dynamic movement pattern (i.e., motion), computer vision researchers are progressively exploiting temporal information from visual sensors such as digital video cameras (Song & Fan, 2006; Ahmed et al., 2007).

Even if an object has been successfully separated, there still remains the question of *how can the object be automatically identified or recognized*? Automatic object identification is regarded as the most difficult problem in computer vision because of being unable to define a unified feature set to represent generic natural objects. Any identification system requires an accurate representation and supervised machine learning of an object, its interaction with other objects and a physical context. Many techniques (Lou et al., 2002; Yu et al. 2002) have evolved though their performance is limited. In addition, the development of a suitable training set requires considerable manual intervention and so is time consuming (Shirasaka et al. 2006). This has prompted the adoption of RFID technology in object analysis, which has been already proved as an effective approach in application domains like object tracking in automated assembling lines (Wang, 2007), workflow optimization (Faschinger et al., 2007) and inventory control (Goodrum et al., 2006; Ko, et al., 2007).

The main drawbacks of RFID are very short sensing range, missed detection, sensitivity to environmental noise and the inherent fact that the technology does not work for certain materials like metals and food items, while some items are simply too small to have a tag attached. For these reasons, RFID alone is often ineffectual, which motivated the investigation of alternative approaches based on combining RFID and com-

Related Content