Chapter 4
From Object Recognition to Object Localization

Rigas Kouskouridas
Democritus University of Thrace, Greece

Antonios Gasteratos
Democritus University of Thrace, Greece

ABSTRACT

Recognizing objects in a scene is a fundamental task in image understanding. The recent advances in robotics and related technologies have placed more challenges and stricter requirements to this issue. In such applications, robots must be equipped with a sense of location and direction with a view to the efficient accomplishment of navigation or demanding pick and place tasks. In addition, spatial information is required in surveillance processes where recognized targets are located in the working space of the robot. Furthermore, accurate perception of depth is mandatory in driver assistance applications. This chapter presents several recently proposed methods capable of first recognizing objects and then providing their spatial information in cluttered environments.

INTRODUCTION

Computer vision generally interferes with recognizing patterns and targets. A wealth of research is devoted to the building of algorithms capable of either detecting simple blob-like structures or recognizing complicated patterns. Generally, the efficiency of a pattern recognition technique depends on its ability to decode, with as much accuracy as possible, vital visual information contained in the natural environment. During the past few years, remarkable efforts were made to build new algorithms for robust object recognition in difficult environments. To this end, researchers emphasized in developing recognition paradigms based on appearance features with local potency (Nister, D., & Stewenius H. 2006, Sivic, J., & Zisserman, A. 2003). Algorithms of this field extract features with local extent that are invariant to possible illumination, viewpoint, rotation and scale changes.

Another aspect that has received much attention in the literature is to exploit the data derived during recognition with a view to provide objects' spatial information. Apart from its identity, several other object-related characteristics, such as its distance to the camera or its pose (orientation relative to the camera's plane), could be obtained (Thomas, A., et al. 2009, Sandhu, R., et al. 2009, Ekvall, S., 2005). As a result, assigning spatial attributes to recognized objects provides solutions to numerous technical problems. In robotics applications robots must be equipped with a sense of location and direction with a view to the efficient accomplishment of navigation or demanding pick and place tasks (Kragic, D., et al. 2005, Wong, B., & Spetsakis, M. 2000). In addition, spatial information is required in surveillance processes, where recognized targets are located in the working space of the robot. Furthermore, accurate perception of depth is mandatory in driver assistance applications (Borges, A. P. et al. 2009). Quality control procedures of industrial production frameworks demand accurate acquisition of enhanced spatial information in order to reject faulty prototypes. To sum up, the ultimate challenge for computer vision society members is the building up of advanced vision systems capable of both recognizing objects and providing their spatial information in cluttered environments.

This chapter is mainly devoted to two major and heavily investigated aspects. Initially, the current trend in recognition algorithms suitable for spatial information retrieval is presented. Several recently proposed detectors and descriptors are analytically presented along with their merits and disadvantages. Their main building blocks are examined and their performance against possible image alterations is discussed. Furthermore, the most known techniques emphasizing in the estimation of pose and location of recognized targets are presented.

LOCAL APPEARANCE-BASED OBJECT RECOGNITION AND MULTI-CAMERA SYSTEMS

State-of-the-art object recognition frameworks rely on local appearance-based features extracted and organized by detectors and descriptors respectively. Generally, the main idea behind interest location detectors is the pursuit of points or regions containing exceptional information about an object or a scene. These spots or areas hold data that distinguish them from others in their local neighborhood. Thus, regions in a scene that enjoy solitary quality and quantity of information can be easily detected. It is apparent that, detector's efficiency is directly related to its ability to locate as many distinguishable areas as possible in an iterative process. Harris and Stephens (Harris, C., & Stephens M. 1988) were the first to implement an interest point detector, known as Harris Corner detector. Due to the fact that it provides significant repeatability, many recent proposed studies (Rothganger, F. et al. 2006, Schmid, C., & Mohr, R. 2006) have adopted it in order to perform demanding object recognition tasks. Furthermore, several variations of Harris Corner detector, such as Harris-Laplace (Mikolajczyk, K., & Schmid, C. 2004) and Harris-affine (Mikolajczyk, K., et al. 2005), were presented due to the fact that they provide significant efficiency. In turn, another profitable detector, the Maximal Stable Extremal Regions Detector (MSER), was recently proposed in (Matas, J. et al. 2004). In short, regions darker or brighter than their surroundings are detected, while the efficiency of the algorithm relies on the trade-off between pixels' intensity value on the center of the mask and those on the local neighborhood. In turn, the "Features from Accelerated Segment Test (FAST)" feature detector constitutes the most recently interest point extractor and it was proposed in (Rosten, E., & Drummond, T. 2006).

Related Content

Modelling and Simulation Platform for Chemical Plume Tracking and Source Localization
Tien-Fu Luand Mohamed Awadalla (2020). *Robotic Systems: Concepts, Methodologies, Tools, and Applications (pp. 1489-1518).*
www.irma-international.org/chapter/modelling-and-simulation-platform-for-chemical-plume-tracking-and-source-localization/244069

Stereo Vision Depth Estimation Methods for Robotic Applications
Lazaros Nalpantidisand Antonios Gasteratos (2014). *Robotics: Concepts, Methodologies, Tools, and Applications (pp. 1461-1481).*
www.irma-international.org/chapter/stereo-vision-depth-estimation-methods-for-robotic-applications/84960

Could Emotions Be Modelled through Information Processing?
Ebrahim Oshni Alvandi (2015). *Handbook of Research on Synthesizing Human Emotion in Intelligent Systems and Robotics (pp. 63-82).*
www.irma-international.org/chapter/could-emotions-be-modelled-through-information-processing/127555

Which is Better?: A Natural or an Artificial Surefooted Gait for Hexapods
Kazi Mostafa, Innchyn Herand Jonathan M. Her (2011). *International Journal of Intelligent Mechatronics and Robotics (pp. 63-72).*
www.irma-international.org/article/better-natural-artificial-surefooted-gait/58323

A Logical Model for Narcissistic Personality Disorder
Mohamad Ab Salehand Ali Awada (2016). *International Journal of Synthetic Emotions (pp. 69-87).*
www.irma-international.org/article/a-logical-model-for-narcissistic-personality-disorder/172104