Chapter 10
A 3D Vision–Based Solution for Product Picking In Industrial Applications

Mirko Sgarbi
Scuola Superiore Sant’Anna, Italy

Valentina Colla
Scuola Superiore Sant’Anna, Italy

Gianluca Bioli
Scuola Superiore Sant’Anna, Italy

ABSTRACT
Computer vision is nowadays a key factor in many manufacturing processes. Among all possible applications like quality control, assembly verification and component tracking, the robot guidance for pick and place operations can assume an important role in increasing the automation level of production lines. While 3D vision systems are now emerging as valid solutions in bin-picking applications, where objects are randomly placed inside a box, 2D vision systems are widely and successfully adopted when objects are placed on a conveyor belt and the robot manipulator can grasp the object by exploiting only the 2D information. On the other hand, there are many real-world applications where the 3rd dimension is required by the picking system. For example, the objects can differ in their height or they can be manually placed in front of the camera without any constraint on the distance between the object and the camera itself. Although a 3D vision system could represent a possible solution, 3D systems are more complex, more expensive and less compact than 2D vision systems. This chapter describes a monocular system useful for picking applications. It can estimate the 3D position of a single marker attached to the target object assuming that the orientation of the object is approximately known.

INTRODUCTION
Vision systems are widely used in industrial applications for different tasks (Malamas, Petrakis, Zervakis, Petit, & Legat, 2003; Chin & Harlow, 1982; Nelson, Papanikolopoulos, & Khosla, 1996). For example, visual inspection systems allow 24-hour inline quality control of the parts during manufacturing processes. If some parts are out of tolerance
or defective, they can be discarded at a very early stage or reallocated. Vision systems are used to verify the correctness of assembly operations by verifying the position and the orientation of each component. Code readers allow identification and full tracking of each product. Vision systems are used to create fully automated solutions for handling dangerous or toxic materials.

One of the most interesting fields where vision systems are successfully applied is the control of robots for pick and place operations. Thanks to increasing computational power and after years of extensive research, vision systems are now emerging and proving their validity in very complex applications like grasping parts that are randomly placed inside a bin. In this case, a 3D vision system has to recognize and localize the target object among many other ones, under any perspective transformation and even if it is partially visible. Typical 3D vision systems use either a single camera combined with a laser projector (Alshawish & Allen, 1995) or two cameras for stereo vision (Brown, Burschka, & Hager, 2003). On the other hand, less complex one-camera 2D vision systems (Cheng & Denman, 2005) are widely used in more structured pick and place applications, where the target objects lie, for instance, on a conveyor belt. These types of systems have reached a high level of reliability. However, most of them require the images to be taken with the target object located at a predefined distance from the camera. Perspective deformation errors of the image are compensated through the particular shape of the end-effectors. Actually the end-effectors are designed to allow a firm grasp despite of an imprecise localisation (obviously within reasonable limits).

Unfortunately, there are many real-world applications that cannot be represented by the above-described situation. The objects can differ in height or be manually placed in front of the camera without any constraints on their distance from the camera itself. Some production lines can adopt a mix of manual and automatic operations on the parts conveyed from one working unit to another. For instance, automatic operations could be directly performed by a robot. It would be convenient to exploit the flexibility of the robot system in order to pick the object from a line, machine it and to place it back on the line. Unfortunately, picking objects on a conveyor that is used both by human operators and robot manipulators can be a very difficult task. For example, parts can be randomly placed over a conveyor belt or hung up on a conveyor chain without precise restrictions on orientation and position.

The detection of the picking point without a well-structured conveying system could, in practice, prevent efficient use of the automatic station, as human intervention is always required in order to pick the objects, place them on the machine and perform the reverse operation. An alternative approach consists in changing the entire production line by adding custom fixtures and specific constraints. This solution could result in high costs and poor flexibility. The optimal solution would be to adopt automatic units capable of substituting human intervention in pick and place operations. This goal can be achieved through the adoption of a 3D vision system to estimate the position and orientation of the part to be picked. As standard 3D systems are complex, expensive and cumbersome, other solutions have also been explored (Oh & Lee, 2007).

This chapter offers a solution that requires some minor constraints on the orientation and position of the objects that are placed on the conveyor by a human operator, i.e. the objects must be placed on the conveyor belt with approximately the same position and orientation and with a predefined side facing the camera. The proposed system estimates the 3D position of the object by using only one camera and only one marker on the target object. The object orientation has to remain almost the same for each picking operation. System accuracy must be sufficient for a correct grasp, but an exact object or marker location is not required.

The chapter is organised as follows: the following section provides the background on vision...