Chapter 1

Introduction

Robotics and automation have undergone an outstanding development in the manufacturing industry over the last decades owing to the increasing demand for higher levels of productivity and quality regarding numerous kinds of industrial activities. Many tedious and potentially harmful tasks, such as car welding and painting, have managed to be automated, shifting the human workforce towards more skillful occupations, such as assembly or quality control monitoring.

However, in spite of those significative advances, the automation levels accomplished in modern factories have not yet met the expectations that many foresaw for the beginning of the twenty-first century. On the one hand, the wide majority of repetitive tasks that have been able to be automated must be performed in highly controlled environments, where everything has been designed in order to simplify the process, and where even apparently insignificant failures may lead to unexpected and costly shutdowns of the whole production line. On the other hand, the continuous advances of robotics and artificial intelligence are still far away from allowing the automation of many tasks, which are still the stronghold of experienced human operators.

In order to change this scenario and move towards a truly automated industry, it is necessary to increase the flexibility and reliability of the whole production system. The solution to this challenge requires the implementation of a variety of technologies that, in many cases, are borrowed from disciplines that have nothing to do with real time automation, and that, therefore, must be properly adapted to fulfill new requirements, especially regarding time constraints.

Such an increase in flexibility and reliability is impossible to achieve without endowing the production systems with advanced sensory capabilities that allow them to detect unexpected events or to assess the quality of the products they are producing. Among the sensory systems, the visual perception systems provide the largest amount of information in the least time. Therefore, vision systems are the perfect candidates to acquire most of the necessary information about the working environment where the industrial activity takes place, as well as the status of the products being manufactured. The availability of such information is vital to accomplish the new challenges of robotics and automation.

At this point it seems apparent that both computer vision and image processing are two disciplines that will play an important role in the future development of robotics and automation. So far, their application to the automation industry has been rather incipient, mainly oriented towards controlling the quality of the manufactured products. However, the new problems require new solutions that will lead to the development of complex algorithms capable of reasoning in real time about the images being acquired by the perceptual system.

Computer vision and image processing are two complementary disciplines that are sometimes confused. Computer vision can be defined as a process of extraction of information about the physical world from image data by using a computer (González Jiménez, 1999), whereas image processing can be defined as a process that transforms a degraded image to another image of better quality in order to facilitate its posterior interpretation. It has traditionally been considered that images processed by computer vision applications are utilized by computers, while image processing applications generate images for human consumption (Umbaugh, 1998). In practice, image processing techniques can be applied in

order to facilitate the interpretation of images by further computer vision algorithms. Therefore, in this context, image processing can be considered to be a preprocessing stage for computer vision.

1.1 Problem Statement

The growing need for efficiently processing and analyzing the information contained in digital images is a continuous challenge in order to apply image processing and computer vision technologies to robotics and automation. Digital images are commonly processed in a brute force style, by analyzing all the pixels contained in the images, no matter how big and redundant they are. Since images may contain hundreds of thousands of pixels, any sequential processing upon them becomes easily tedious, even for the fastest processors, limiting thus the complexity of the visual tasks that can be performed with success.

In practice, hardware accelerated processors are frequently utilized in industry in order to perform simple image processing and computer vision algorithms in real time. However, these hardware solutions become quickly overcome as faster and faster off-the-shelf processors are developed. Thus, the high investments necessary for acquiring such specialized hardware are difficult to be justified in the end, especially when considering that maybe a year after the acquisition of the equipment, any personal computer can perform at similar speed or even faster than the dedicated hardware for a fraction of its cost. Hence, from the industry standpoint, it is important to develop fast image processing and computer vision techniques that manage to generate results efficiently by relying just on software algorithms and not on specialized hardware. This is one of the current challenges for the image processing and computer vision communities.

The efficiency problem related to the processing of digital images comes from the fact that digital images are large representations, typically containing hundreds of thousands of data that must be repetitively processed. Reducing the size of digital images is a problem that has been extensively tackled from the perspective of image storage and transmission. As a result, a wide variety of standard techniques have already been devised for compress-

ing digital images, achieving spectacular size reductions through popular algorithms, such as GIF or JPEG compression.

However, those techniques were not designed for applying image processing or computer vision operations upon the compact representations obtained as a result of the compression process. Thus, even though the images are kept in a compact form, they must be uncompressed prior to being able to process them.

Nonetheless, some researchers have managed to apply a limited number of basic operations (arithmetic operations, scaling, feature extraction, ...) upon such compressed representations (Smith, Rowe, 1993; Chang, 1995; Natarajan, Bhaskaran, 1995; Shen, Sethi, 1996). However, this technology is still far away from allowing the application of even the most common image processing and computer vision algorithms directly in the compressed domain.

An alternative way of representing images in a more compact way is through the utilization of geometric representations, such as adaptive triangular meshes (García, Sappa, Basañez, 1997; Wilson, Hancock, 1997; Terzopoulos, Vasilescu, 1991; DeFloriani, 1989; Soucy, Croteau, 1992). The advantage of using those representations is that they allow the modeling of large areas of pixels with basic geometric primitives. For example, a large white region can be represented by a few triangles instead of by hundreds of pixels. Additionally, these representations allow the application of further processing operations directly in the geometric domain. Thus, by taking advantage of the benefits of those representations, some processing operations, such as segmentation (Gevers, Smeulders, 1997; García, Basañez, 1996) and integration (Soucy, Laurendeau, 1995b; Sappa, 1999; Turk, Levoy, 1994; Pito, 1996a), have already been proposed to be efficiently performed in the geometric space instead of in the image space.

However, the suitability of these alternative representations for the application of other basic computer vision and image processing operations is still an open issue. This has been the basic motivation for the work done in this dissertation.

1.2 Research Focus

This dissertation has focused on the development and evaluation of a set of techniques to perform typical image processing operations upon adaptive triangular meshes representing digital images. Three types of digital images have been considered: gray-level images, range images and digital elevation models.

G*ray level images* are two-dimensional arrays in which every element is a pixel that represents a certain gray level (light intensity) between black and white. *Range images* are two-dimensional arrays of pixels in which each element does not store a level of light intensity, but the distance from a point on the surface of a 3D object to a virtual plane referred to the range sensor utilized to acquire the image (Sappa, 1999). Finally, *digital elevation models* (*DEMs*) are two-dimensional arrays that keep terrain elevations at regularly spaced horizontal intervals. DEMs are utilized to model terrain surfaces.

The processing operations developed in this work belong to two broad categories within image processing: *image analysis* and *image enhancement*, according to the classification utilized by (Jain, 1989) and (Umbaugh, 1998). The image analysis operations that have been developed include: geometric transformations, thresholding and quantization, algebraic operations, selection of regions of interest, generation of synthetic images, edge detection, region segmentation and labeling, feature extraction and histogram generation. On the other hand, the developed image enhancement operations include gray-scale and histogram modification algorithms, and image filtering techniques. One of the aims of this work has been the identification of those functions that can perform faster in the geometric domain than in the image domain. This may occur due to the fact that an adaptive triangular mesh may be able to approximate the contents of a digital image with fewer data.

Besides the acceleration of some image processing operations, an important spin-off of this research has been the possibility of applying typical operations utilized in image processing to $2^1/_2$D adaptive triangular meshes in general. These meshes can represent any kind of surface, such as terrain. In this way, it is possible to process terrain surfaces as if they were digital images, utilizing the know-how available to image processing and com-

puter vision in order to reason about terrain shapes. This has, for instance, direct application to path planning and obstacle detection and recognition in robotics, as well as potential applications in other areas related to the environmental sciences.

Besides the development of image processing operations applied to triangular meshes, this work has also tackled the problem of generating adaptive triangular meshes from digital images and vice-versa. All in all, this dissertation provides a set of tools to map digital images to triangular meshes, to process triangular meshes in the geometric domain and to map the obtained triangular meshes back to image space.

1.3   Organization of the Dissertation

The next chapters of this dissertation are organized as follows. Chapter 2 describes previous related work related to the topics dealt with in this dissertation. Chapter 3 describes two techniques to approximate digital images with adaptive triangular meshes without applying optimization, as well as efficient techniques to generate digital images from adaptive triangular meshes (García, Vintimilla, Sappa, 1999a; García, Vintimilla, Sappa, 1999b; García, Vintimilla, Sappa, 2000).

Chapter 4 describes a set of techniques to perform analysis operations upon adaptive triangular meshes. The chapter includes operations to perform: geometric transformations (García, Vintimilla, Sappa, 2000), thresholding (García, Vintimilla, 2000a) and quantization, algebraic operations, selection of regions of interest, generation of synthetic images, edge detection, region segmentation and labeling (García, Vintimilla, 2000a), feature extraction and histogram generation.

Chapter 5 describes the implementation of image enhancement techniques from adaptive triangular meshes (García, Vintimilla, 2000b). The chapter includes gray-scale and histogram modification algorithms, as well as image filtering techniques.

This dissertation concludes with Chapter 6, which summarizes the whole work, highlighting its principal contributions and proposing future lines of research.