Chapter 2

FEATURE SYNTHESIS FOR OBJECT DETECTION

2.1 Introduction

Designing automatic object detection and recognition systems is one of the important research areas in computer vision and pattern recognition [7], [35]. The major task of object detection is to locate and extract regions of an image that may contain potential objects so that the other parts of the image can be ignored. It is an intermediate step to object recognition. The regions extracted during detection are called regions-of-interest (ROIs). ROI extraction is very important in object recognition, since the size of an image is usually large, leading to the heavy computational burden of processing the whole image. By extracting ROIs, the recognition system can focus on the extracted regions that may contain potential objects and this can be very helpful in improving the recognition rate. Also by extracting ROIs, the computational cost of object recognition is greatly reduced, thus improving the recognition speed. This advantage is particularly important for real-time applications, where the recognition accuracy and speed are of prime importance.

However, the quality of object detection is dependent on the type and quality of features extracted from an image. There are many features that can be extracted. The question is what are the appropriate features or how to synthesize features, particularly useful for detection, from the primitive features extracted from images. The answer to these questions is largely

dependent on the intuitive instinct, knowledge, previous experience and even the bias of algorithm designers and experts in object recognition.

In this chapter, we use genetic programming (GP) to synthesize composite features which are the output of composite operators, to perform object detection. A composite operator consists of primitive operators and it can be viewed as a way of combining primitive operations on images. The basic approach is to apply a composite operator on the original image or primitive feature images generated from the original one; then the output image of the composite operator, called composite feature image, is segmented to obtain a binary image or mask; finally, the binary mask is used to extract the region containing the object from the original image. The individuals in our GP based learning are composite operators represented by binary trees whose internal nodes represent the pre-specified primitive operators and the leaf nodes represent the original image or the primitive feature images. The primitive feature images are pre-defined, and they are not the output of the pre-specified primitive operators.

This chapter is organized as follows: chapter 2.2 provides motivation, related research and contribution of this chapter; chapter 2.3 provides the details of genetic programming for feature synthesis; chapter 2.4 presents experimental results using synthetic aperture radar (SAR), infrared (IR) and color images. Various comparisons are given in this section to demonstrate the effectiveness of the approach, including examples of two-class and multi-class imagery; finally, chapter 2.5 provides the conclusions of this chapter.

2.2 Motivation and Related Research

2.2.1 Motivation

In most imaging applications, human experts design an approach to detect potential objects in images. The approach can often be divided into some primitive operations on the original image or a set of related feature images obtained from the original one. It is the expert who, relying on his/her experience, figures out a smart way to combine these primitive operations to achieve good detection results. The task of synthesizing a good approach is