Object Detection by Combing Local Features' Motion,

Appearance, and Location Information in a Robust Voting System

by

Zhipeng Wang

A Doctor Thesis

Submitted to

the Graduate School of the University of Tokyo

on June 14, 2013

in Partial Fulfillment of the Requirements

for the Degree of Doctor of Information Science and Technology

in Computer Science

Thesis Supervisor: Katsushi Ikeuchi

Professor of Computer Science

ABSTRACT

Object detection is a fundamental perceptual skill in human, and plays an important role in machine vision area. Effective object detection methods can help with video surveillance, driving assistance, etc. Researchers improve performance of detection methods mainly by proposing better representative models, better classifiers, or more efficient methods for solution space exploring.

In this work, the performance of detection methods is improved from a different aspect. The work explores the information which are not made fully use of by previous methods. And the efforts are two-fold: 1) utilizing of motion information by combining it with appearance information, and 2) utilizing of mutual information encoded among the image features of the same object. Three detection methods are proposed accordingly.

The first detection method is developed for real-time applications. In a hierarchical way, this method makes time-consuming steps in its pipeline deal with fewer instances, and combines motion information with appearance information very efficiently. This method gives promising detection results in real time, and gives 100% detection rate and 0% false alarm rate in one of the experiments.

Since the performance of the first method in complex scene is not promising, the second method is proposed. This method extends the Implicit Shape Model to incorporate motion information, and outperforms the state-of-the-art method on two datasets. This method also performs well in distinguishing near objects and similar different-class objects.

To improve the efficiency of the second method, the third method is proposed, which is Pyramid Match Score for detection. The method does pyramid matching during training and detection for efficiency, and makes full use of the visual and spatial information encoded among the image features of the same object, which improves detection performance.

2   PMS(Pyramid Match

Score)   PMS   2   100