論文内容の要旨

LOCAL FEATURE BASED RECOGNITION AND CLASSIFICATION OF VEHICLES FOR TRAFFIC SURVEILLANCE

（局所特徴を使用した交通監視のための車両認識とクラス分類）

モホッタラ　ヤサンティ　シャミラ　ウィクラマーチ

With the rapid increase of motorization and population density, traffic problems such as congestion, road crashes, and environmental pollution have become serious and inevitable in and around urban areas.

Intelligent Transport Systems (ITS), which applies advanced computing, information and communication technologies to surface transport systems, is widely viewed as a major solution to many of these problems. ITS offers improved safety, efficiency in transport, and a positive environmental effect by reducing vehicle emissions and energy consumption.

In many ITS applications, sensing is a fundamental task involving different kinds of sensors that collect various traffic parameters. Point-oriented sensors, such as ultrasonic sensors and magnetic loop detectors, have been in practical use for decades to count the vehicles passing through. Compared to loop detectors, vision-based sensors provide more flexible and detailed traffic parameters such as vehicle speed, location, vehicle class, lane changes, and even descriptions of road conditions. Moreover, the ease of installation and maintenance is another attractive advantage of vision-based sensors. However, vision-based vehicle sensing is very challenging, because the appearance of a vehicle may vary greatly due to complex outdoor environmental conditions, cluttered backgrounds, shadows, and occlusion. This dissertation focuses on vision-based vehicle recognition, and proposes two systems, one that uses a stationary camera and one that uses an on-board camera, to overcome these issues.

In the first study, we present a vehicle detection system to detect vehicles from the images acquired through a single stationary camera mounted over the road. The recognition algorithm is a modification of the Eigen window method, which is based on local features. To minimize the effects from outdoor illumination changes, the images are converted to binary edge images. The system extracts local binary features from training images and compresses them to code features using the vector quantization method. The detection of vehicles is done by matching all the features from the input image with these code features, followed by a voting process.

The system gave promising results in outdoor experiments, showing an accuracy of over 98%. Not only the existence of vehicles, but also their location and vehicle area were recognized accurately. Moreover, the system worked robustly in the presence of shadows, outdoor illumination changes, and partial occlusion, proving the ability to be fixed in many ITS applications for practical use.

In the next step, we enhanced the system to recognize the classes of vehicles. Vehicle classification is an important requirement for traffic surveillance in many governmental and public systems, and also in investigation and traffic census to obtain information on the vehicle types that use a particular street. In the current situation, human operators manually count the vehicles, giving poor results in quality and cost. As a result, the road conditions are often described by outdated data. An automated vehicle classification system can offer many benefits for these applications, expanding their possibilities.

We modified the recognition algorithm with three improvements, combining them in a probabilistic approach, to fit vehicle classification better. The three improvements are:

1. Modified feature extraction: to extract more robust and reliable features.
2. Background verification: to eliminate misidentifications of a part of a large vehicle as a small vehicle.
3. Foreground verification: to eliminate misidentifications of small vehicles as a part of a large vehicle.

The classification is divided into five vehicle classes, namely, sedan, wagon, mini-van, hatchback, and others. Compared to a wide range of vehicle classification research studies that divide vehicles into classes such as buses, cars, and motorbikes, our target is hard, because our classes are not distinguishable by size and their interclass differences are very small.

Model-based classification requires good training images of all targeted vehicle classes. Collecting these images is generally a hard and time-consuming task. To overcome this issue, we propose using training images rendered through three-dimensional computer graphics (CG). Today's CG technologies make it possible to change the appearance, the illumination conditions, and the viewpoint freely, ensuring a high degree of reality at the same time. Moreover, small differences in appearance will not influence our classification, because the system only uses the binary-edge images obtained from these CG images.

Outdoor experiments were done to evaluate the classification algorithm as well as the effectiveness of CG images as training data. The results showed an accuracy of over 89% in classifying vehicles into our five vehicle classes. More specifically, hatchback was classified 100% accurately, while sedan and wagon followed with a high accuracy of over 90%. The experiments also proved that using CG training images is successful and causes no significant effect on accuracy, while dispensing much of the trouble of collecting real vehicle images for each class.

Vision sensors are good at wide and coarse detection, but they still hold difficulties in performing under outdoor environmental conditions. On the other hand, recent laser range sensors have become more simple, compact, and reliable, attracting more attention in many applications. But they are incapable of retrieving texture on objects. Thus, a fusion of a vision sensor and a laser range sensor will make it possible to obtain 3D models with textures, leading to many robust and reliable applications for practical use.

In the third study, we propose such a system that fuses laser range data and image data from a probe vehicle for robust recognition, and evaluates it in recognition of on-street parked vehicles. The two sensors are mounted on a probe vehicle, and they scan the environment while the probe vehicle runs in the lane next to parked vehicles. The laser range data is processed first to obtain a depth map, and to segment the vehicles from the background. Segmentation can be simply done based on depth, but for more reliable segmentation, we apply a method called the height curve based method, that calculates a histogram of scanned points to detect the vehicle body surface.

Next we calibrate the two sensors to obtain the projection matrix from the laser coordinate system to the image coordinate system. Once the laser data for vehicles and the calibration matrix are ready, we fuse the laser range data and image data. Assuming the probe vehicle progresses straight in a constant speed parallel to on-street parked vehicles, the points in the laser scanned lines are projected onto the corresponding images and lined up.

We obtained very good results in calibration, projecting the laser points onto the vehicle area quite correctly. But the system was not powerful enough to segment vehicles from images precisely. The main reason for that was the poor reflectance of laser from some points, especially from black vehicles. For more robust segmentation, we propose employing a graph cut method fused with laser data to initialize the background/foreground area. The proposed method enables segmenting the vehicles with high accuracy. This system can be applied to scan the 3D geometry of objects and retrieve the texture at the same time. Hence, the possibilities of this system can be expanded to applications such as 3D modeling of towns, 3D navigation systems, and many more applications.

Next we apply these segmented vehicle images to classification. We found a large number of unnecessary edges on vehicle bodies, because the vehicle body reflects the surroundings. We modify the feature extraction and voting process to avoid the misclassification that can occur due to these excessive edges. The modified system could classify our five vehicle classes with an accuracy of over 85%, and this stage is still to be improved. More efficient results can be expected by fusing with laser data in the classification stage as well.

We conclude that the contributions of this dissertation are as follows:

1. We proposed a vehicle detection system with high accuracy that is robust to outdoor environmental conditions.
2. Next we enhanced the system for vehicle classification. The system performed nearly 90% accurately in recognizing five vehicle classes.
3. In this system, we proposed and proved the effectiveness of using CG images as training data.
4. Then we proposed a sensor fused system for robust sensing. The calibration of laser range sensor data and camera sensor data gave very good results. Graph cut-based segmentation fused with laser range data proved capable of extracting the texture of 3D models accurately. The system can be expanded for practical use in many other applications such as 3D modeling of towns.
5. The effectiveness of our algorithm to recognize on-street parked vehicles is proved through experiments.