A Component-Based System for Car Detection

Kevin Thompson and Bernd Heisele, with T. Poggio

The Problem: The goal of this project is to detect passenger cars based on the position and confidence of several different component classifiers. We also hope to develop a component selection process which can be generalized enough to apply this technique to other domains.

Motivation: Heisele [1, 2] was able to classify faces very accurately using a component-based approach and an SVM classifier. We want to extend that technique first to recognize cars, and then other objects. Cars were selected because intuitively, they would seem to have strong characteristic components. Recent work in Agarwal & Roth [9] suggests that a component-based approach is suitable for cars.

Previous Work: In Papageorgiou and Poggio [5], cars were detected using overcomplete Haar wavelet features and a support vector machine classifier. Perona et al. [6] applied a framework for unsupervised feature extraction, based on the use of an “interest operator” from [8] to select potentially useful feature regions from high-pass filtered images. After correlating and clustering potential features in the positive data set, features of sufficient cluster size were validated in terms of how well they predict the object within a statistical model. Schneiderman and Kanade [7] used features that were probability distributions on one or more levels of wavelet coefficients. Histograms of these single- or multi-level wavelet features were used in a statistical model to detect cars. In addition, separate detectors were trained to recognize eight different car orientations, so that by combining their results the system was tolerant to a large variance in orientation.

The idea of using complex, spatially localized component inputs to classify objects is evident to some degree in Perona [6], in the form of the image regions found by the interest operator. Mohan et. al. [3, 4] describes a pedestrian detection system which uses SVM classifiers on Haar wavelet data to detect head, arm, and leg components within specific regions of an image. These classifier outputs are then given to a final SVM, which classifies the pattern as a pedestrian or not. Heisele et al. [1, 2] applies a similar technique to face detection, using histogram-equalized grayscale components. This problem proved particularly suitable to component-based detection, due to the availability of 3D morphable head models for generating synthetic data. Two dimensional images generated from this data have known correspondence points, based on their projection from the 3D model, which can be exploited to automatically extract features around these reference points.

Most recently, Agarwal & Roth [9] have attempted automated component extraction and component-based recognition on cars. They used the same interest operator as Perona [6] for selecting fixed-size components. For classification, they generate very large, very sparse binary feature vectors which contain information on these components’ multiplicity and their spatial relationships. They use the SNoW learning system to classify these sparse feature vectors.

Approach: The goal of this project is to test a component-based SVM approach, similar to Mohan [3, 4] and Heisele [1, 2], on the domain of car detection. This technique will involve several steps, including component selection, component classifier training and testing, feature selection for the final SVM, and optimization. Some of the potential benefits of this system include high recognition performance, based on the good results of those two previous component-based SVM systems, and a degree of rotation invariance (as described in Heisele [2]), thanks to minimal changes in specific components under object rotation.

For component selection, we have tried manually selecting a small number of example components and then using correlation within windowed constraints to build component training sets. These training sets are hand-refined based on their errors and re-correlated. We will also be testing automated extraction techniques, using an interest operator to pick out starting points for component building and “growing” the components as in Heisele [1, 2].

The architecture of the completed system will be similar to those described in the Mohan and Heisele papers. First, components will be searched for within restricted search windows on a particular image. Second, the outputs of these classifiers, along with positional data about where they maximize within the search windows, will be passed to a final SVM classifier.
Once the component classifiers have been constructed, the final steps in building a whole-car classifier will be to select the best features from the available component classifiers, and to train the output SVM. We select classifiers based on which ones have the best classifiers (i.e., the classifiers that can best distinguish component from non-component).

The output SVM will be trained on the output levels of the selected component classifiers, as well as on their position data. The component classifier output levels are fairly straightforward to get, but there are a few options for how to represent the position data. We will be testing deviation from mean location within the search windows (used by Heisele), inter-component distances (used by Perona), and normalized inter-component distances. The results in [9] suggest that the angle between components may also be important.

**Difficulty:** The system must have as general a component selection process as possible, while still achieving high classification accuracy.

**Impact:** We hope to show that SVM component-based classification, which has already been successful at classifying faces, can be applied successfully to other domains. A general technique for component selection will allow this technique to be applied to a broad variety of problems.

**Future Work:** Over the course of this research, we will be trying different techniques at component selection and different metrics for component classification. Afterwards, we will see how well our component selection and classification technique works on other types of objects.


**References:**


