CS771A Project: Vehicle Classification

Group Number: 28

Objective

- Detect and classify objects in the traffic surveillance video into pedestrians and vehicles
- Further classify the vehicles into:
  - 2 wheelers
  - Rickshaw
  - Autorickshaw
  - Car

Roadmap

- Feature extraction and Classification
- Foreground-Background Separation
- Classify the detected objects

Feature Extraction and Classification

- Extracted features like HOG(Histogram of Oriented Gradient) and SIFT(Scale-Invariant Feature Transform) for training
- Experimented with the following classifiers:
  - LinearSVM
  - Random Forest
  - Adaboost

Feature Extraction and Classification

HOG

A feature descriptor used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of image.It is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy.

Our experimentation

Tweaked the following parameters:

- Number of pixels in a cell
- Number of cells in a block

HOG Feature 5 Fold Cross Validated Results

Cells per Block: (3, 3) Pixels per Cell: (8, 8)

Classifier	Accuracy	Precision	Recall
Linear SVM	0.852951	0.829102	0.816334
Random Forest n_estimator = 100	0.815646	0.862853	0.862853
Random Forest n_estimator = 200	0.823252	0.866324	0.866324
Random Forest n_estimator = 250	0.821079	0.860848	0.860848
Adaboost	0.674393	0.689259	0.689259

Cells per Block: (2, 2) Pixels per Cell: (8, 8)

Classifier	Accuracy	Precision	Recall
Linear SVM	0.761412	0.684711	0.593173
Random Forest n_estimator = 200	0.805340	0.761680	0.739426

HOG Feature 5 Fold Cross Validated Results

Cells per Block: (1, 1) Pixels per Cell: (8, 8)

Classifier	Accuracy	Precision	Recall
Linear SVM	0.770887	0.694461	0.689056

Cells per Block: (3, 3) Pixels per Cell: (6, 6)

Classifier	Accuracy	Precision	Recall
Linear SVM	0.808354	0.757420	0.722123
Random Forest n_estimator = 200	0.742032	0.675545	0.557587

We selected cells per block= 3,3 and pixels per cell = 8, 8 and Linear SVM as our final classifier.

Feature Extraction and Classification

SIFT

SIFT extracts high contrast regions which are a good feature representation of the objects. It can robustly identify objects even among clutter and under partial occlusion, because the SIFT feature descriptor is invariant to uniform scaling, orientation, and partially invariant to affine distortion and illumination changes.

Feature Transformation

Mapped the feature vectors of objects to clusters using K-Means Clustering

Our experimentation

Tweaked the following parameters:

- Number of clusters(100 to 1000)

Feature Extraction and Classification

5 Fold Cross Validation Results with SIFT features

No. of clusters	Classifier			
	LinearSVC	Random Forest n_estimator = 100	Random Forest n_estimator = 200	AdaBoost Classifier n_estimator = 100
100	0.625053	0.635853	0.645788	0.573218
200	0.619438	0.655723	0.655723	0.554643
300	0.598272	0.647084	0.651835	0.556371
400	0.621598	0.661771	0.662203	0.557235
500	0.609071	0.656587	0.665226	0.525701
600	0.597840	0.661771	0.659611	0.546004
700	0.623758	0.667818	0.669114	0.531317
800	0.628509	0.664794	0.669114	0.536501
900	0.637149	0.666090	0.679049	0.543844
1000	0.611231	0.665226	0.672138	0.537796

Foreground-Background Separation

Sliding Window

- Detection and classification of objects simultaneously
- Slide a window of fixed size across the frame
- Label each window using the classifier(including the None type)

  Our experimentation

  - Pyramid of a frame in order to accommodate objects of different size
  - Non-Maximum-Suppression to remove redundant labeled windows
  - Hard-negative mining to improve our prediction

Foreground-Background Separation

Blob detection

- Threshold(current frame - background frame) = objects

  Our experimentation

  - GaussianBlur to remove noise from the frame
  - BackgroundSubtractorMOG2 from OpenCV to detect objects
  - Morphology Transformations to separate adjacent objects
  - Threshold on the size of blobs to filter noise