Image Region Classification based on Multiclass SVM

- *Final Report*

Lei Liu

Goal

The goal of image region classification is to simply the representation of an image into something that is easier to analyze. Detection of semantically meaningful regions in digital images could find its use in classification and retrieval systems for large image databases, image processing algorithms, etc.

The goal of the project is to build an image region classification tool that is capable of classifying each pixel in an image into five different categories using Multi-class SVM:

Sky: sky including clouds

Ground: road, sand, …

Water: sea, river, lake, …

Building: houses, man-made structures, ...

Vegetation: grass, tree

Approach

1. Data Collection

MSRC-V2 Image Database: publicly available "semantic segmentation" image databases which contain 591 photos manually segmented into 23 object classes, including the five classes we aim to classify.

Steps:

- Get a set of tiles (we choose a tile to be a square subdivision of the image with length of 15 pixel) for each class from the image database: For each image in the database, subdivide it into tiles and store tiles that are entirely inside one of the five target regions into corresponding groups.

- For each class, randomly select 3900 tiles as training set and another 3900 tiles as testing set.

2. Feature Extraction

We need to compute a low-level description of tiles before submitting them to the classifier. Joint histogram is used which combines color distribution with gradient statistics. Color distribution is represented by the HSV color space and the vertical and horizontal gradients are computed by applying Sobel's filter to the luminance image. The partition of gradient values is determined by random selection of two million pixels from the training set.

H_{min}	H_{max}	S_{min}	S_{max}	V_{min}	V_{max}
0	360	0	15	0	31
0	360	0	15	32	69
0	360	0	15	70	100
-18	18	16	100	32	100
19	40	16	100	32	100
41	62	16	100	32	100

63	158	16	100	32	100
159	208	16	100	32	100
209	288	16	100	32	100
289	330	16	100	32	100
331	341	16	100	32	100

Partition of HSV color space

Gradient	0.02	0.06	0.17

Partition of vertical and horizontal gradients

Since HSV color space is partitioned into 11 bins and both vertical and horizontal gradients are partitioned into 4 bins, the total number of bins for the joint histogram is 11*4*4 = 176.

3. Multi-class SVM algorithm

The soft margin Gaussian-kernel SVM algorithm was originally implemented using the built-in Matlab quadprog function. However, since quadprog uses primal form and convergence is extremely slow, I decide to use libsvm instead which is a third-party software package for support vector classification. It is much faster than the original implementation due to the fact that it tries to optimize the dual form. Multi-class SVM algorithm is constructed according to One-versus-the-rest strategy. Thus in total we train five SVM classifiers and each was trained to discriminate between one class and the others.

Each classifier defines a discrimination function $f(x) = (w*x+b)/||w||$ (w and b are the normal and bias of the hyperplane) which assumes positive value when the tile belongs to that class and negative otherwise. The output of the combined classifier is the label for the class who has the largest discrimination function value.

4. Pixel-wise classification

The tiles are sampled at one fifth of their size in both x and y dimensions, thus each pixel can be found in 25 tiles except for the pixels on the border of the image.

Each tile is independently classified and the pixel’s final label is determined by majority vote.

It is often the case that some regions of the image can not be labeled as one of the five classes and different labels may be assigned to overlapping tiles. To achieve a more reliable classification, we introduce a rejection rate: if the fraction of concordant votes of overlapping tiles lies below 0.55, then we label the pixels inside these tiles as unknown.

Results

1. Performance of individual SVM classifiers

To train the individual SVM classifiers, we use the 2500 tiles of the related class and a random selection of 2500 tiles form the other classes. To test the SVM classifier, we use all the 5*2500 = 12500 tiles from the test set to evaluate its performance.

We choose Guassian Kernel parameter Sigma = 0.1 and C = 50 through cross-comparing different combinations

The correct rates are as follows:

Building: 83.8%

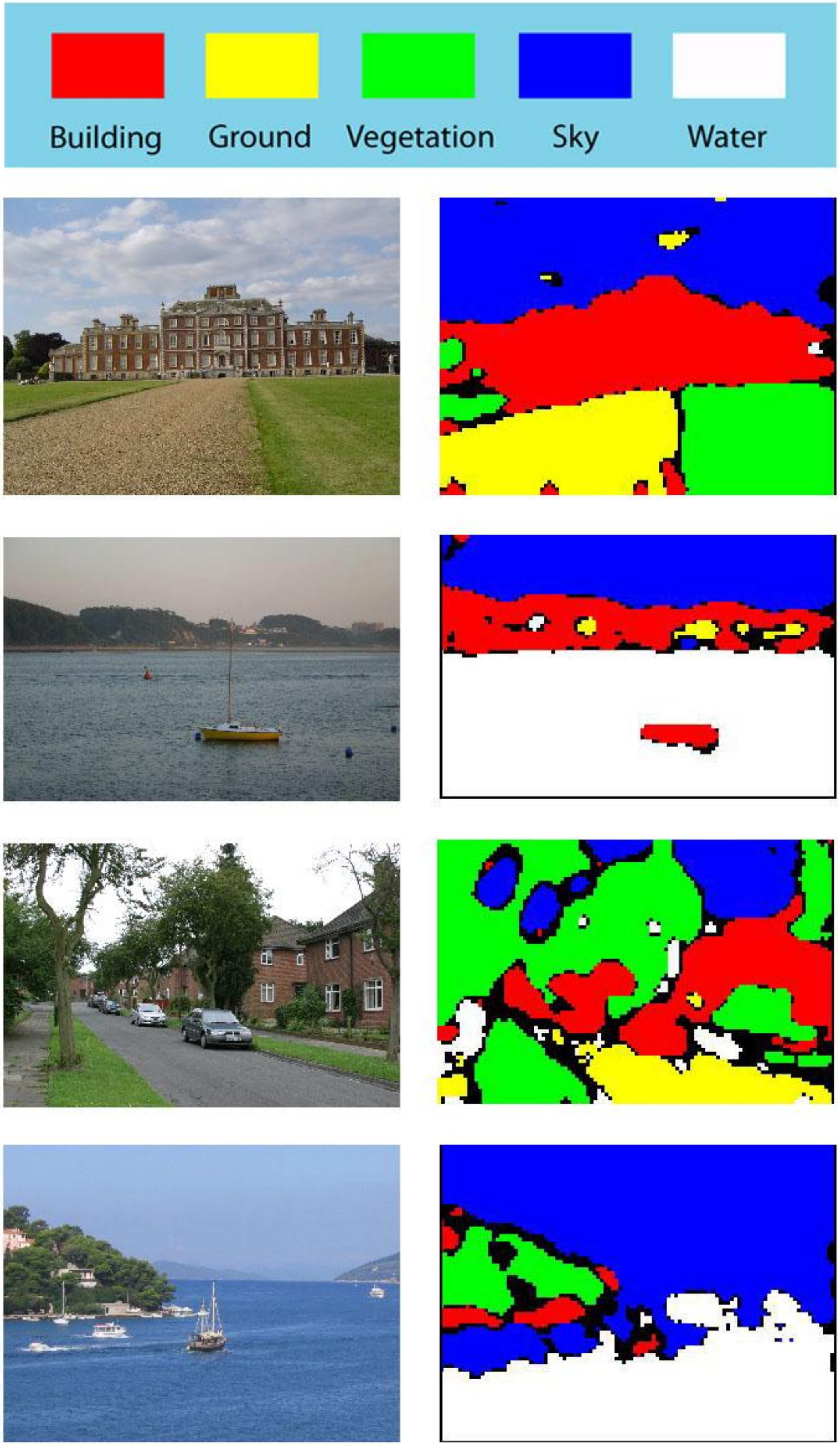
Ground: 84.7%

Sky: 92.0%

Water: 85.1%

Vegetation: 95.6%

2. Image classification results


Building
Ground
Vegetation
Sky
Water

References:

[1] Claudio Cusano, Gianluigi Ciocca, Raimondo Schettini, etc. "Image Annotation using SVM". Proceedings of the SPIE, Volume 5304, pp. 330-338, 2003.
[2] C.Faloutsos, R.Barber, M.Flickner, etc. "Efficient and effective querying by image content". Journal of Intelligent Information Systems, 3, pp. 231-262, 1994.
[3] MSRC-v2 image database, http://research.microsoft.com/en-us/projects/objectclassrecognition/ .
[4] M. Stricker, and M. Swain, "The Capacity of Color Histogram Indexing", Computer Vision and Pattern Recognition, pp. 704-708, 1994.