Presentation Summary

Mean Shift: A Robust Approach Toward Feature Space Analysis

Robert F.K. Martin
27 April 2009

Given an arbitrary dataset, it is often useful to be able to find the most common values. In a vision task for a grayscale image this would correspond to finding the peaks of the histogram. For color space and other high-dimensional feature spaces, the modes are found in the most densely pack hyper-clusters.

The mean shift method, first proposed in 1975, is an iterative way to find those modes. Suppose we have a histogram from some unknown data distribution and we make an initial guess. The mean shift will give us a vector that points toward the nearest mode. With out initial guess and this vector, we make another guess. And so on and so forth until we converge within some tolerance to a mode.

In order to find all the modes with in a dataset, we run mean shift several times from different starting points. From this we keep a list of found modes. With use these found modes as starting points with some added random perturbation. If we converge to the same mode, we keep it on the list. If not, we prune. Then, since gradient descent methods are not always exact, we group nearby modes into one. The final list should be all the maxima of some feature dataset.

Datasets are often modeled with kernel density estimators. Kernel density estimators converge to the true underlying probability density function given a proper kernel and a sufficient sample set of data. Since KDE is simply a summation, its derivative can be found quite easily. From this derivative we can formulate a solution to a gradient descent problem

Mean shift can be applied to several different problems. A very simple application is anisotropic diffusion. By adding in spatial information to the color information, it is possible to group similar colors in similar regions of space. Another simple application is segmentation. By applying anisotropic diffusion to an image, we can gather clusters of pixels that belong to the same region.

A more complex application is object tracking. Given a region in an image, mean shift can be used to find a similar region in space in a successive frame. However, this application tends to be more difficult and not only relies on a good parameterization of the feature space, but also a video sequence with few occlusions or scale changes.