

Probabilistic classifiers using Nearest Neighbor balls

Bo Ranney & Jun Yu
Centre of Biostochastics
Swedish University of Agricultural Sciences
SE 901 83 Umeå
Sweden
Bo.ranneby@sekon.slu.se

Abstract

Accurate and quality assured detection of spatial and temporal variation in land use or landscape patterns is needed as direct or indirect input for the assessment of ecosystem health. For these purposes remote sensing methods have potential to be an efficient tool. Unfortunately, several of the existing remote sensing methods will not give satisfactory results. Although, it is possible to improve traditional remote sensing classification methods, for several applications it is necessary to introduce a new concept, pixel-wise probabilistic classifiers. Instead of classifying each pixel to a specific class, each pixel is given a probability distribution describing how likely the different classes are.

The probabilistic classifier is derived from a non-parametric classification method where multispectral and change detection techniques are combined. After the images have been denoised using the wavelet transform, the information value for each component in the feature vector is calculated and a suitable metric is defined. With this metric the distances to the different classes are measured for each pixel. In order to derive proper probability distributions the distances cannot be raised to any power but nearest neighbor balls have to be calculated.

The probabilistic classifier can be used in several different ways. The pixel-wise vectors of probabilities can be used to judge how reliable a traditional classification is and to derive measures of the uncertainty (entropy) for the individual pixels. Another application is for area estimation. As proper probability distributions have been derived the probabilistic classifier gives unbiased area estimates for arbitrary areas. Using sub sampling methods it is possible to derive variance estimates of the area estimates.

There are several areas of application. For example, the method can be used to derive cost-efficient monitoring systems and in source apportionment models.