Improved building roof type classification using correlation-based feature selection and gain ratio algorithms

ABSTRACT

Of late, application of data mining for pattern recognition and feature classification is fast becoming an essential technique in remote sensing research. Accurate feature selection is a necessary step to improve the accuracy of classification. This process depends on the number of feature attributes available for interactive synthesis of common characteristics that discriminate different features. Geographic object-based image analysis (GEOBIA) has made it possible to derive varieties of object attribute for this purpose; however, the analysis is more computationally intensive. The aim of this study is to develop feature selection technique that will provide the most suitable attributes to identify different roofing materials and their conditions. First, the feature importance was evaluated using gain ratio algorithm, and the result was ranked, leading to selection of the optimal feature subset. Then, the quality of the selected features was assessed using correlation-based feature selection (CFS). The classification results using SVM classifier produced an overall accuracy of 83.16%. The study has shown that the ability to exploit rich image feature attribute through optimization process improves accurate extraction of roof material with greater reliability.

Keyword: Optimization building classification; Feature selection; Gain ratio; Correlation-based selection (CFS)