Abstract

Digital image classification of remote sensing (RS) data is viewed as the process of categorising all the pixels in an image automatically into a finite number of land use/land cover (LU/LC) classes of interest to extract useful thematic information. Since the traditional hard classification techniques are parametric in nature and expect the data to follow a Gaussian distribution, they have been found to be performing poorly on high resolution satellite images of urban/semi-urban environments where LU/LC classes exhibit extensive overlapping in spectral space. Another major drawback of such classifiers lies in the difficulty of integrating ancillary data like digital elevation model, slope, texture, surface temperature, contextual information, etc. into spectral bands since ancillary information gives rise to a non-Gaussian distribution of the resultant data. Consequently, the parametric classifiers fail to exploit the information available through today’s advanced sensor systems and various bands of ancillary data to the highest extent. Hence, generating a satisfactory classified image from the higher spectral, spatial and temporal resolution, and high-dimensional (bands) data is one of the present-day challenges in RS data analysis. In this direction, the current research work is aimed at developing a better classification strategy by integrating advanced image processing techniques and classification algorithms to improve classification accuracy of high spatial resolution RS data over semi-urban LU/LC features.

The research work has been carried out in three stages on a panchromatic sharpened IRS P-6 LISS-IV (2.5m) imagery of the campus of the National Institute of Technology Karnataka, Surathkal, India. The first stage performs the visual and quantitative comparison between the three conventional image fusion techniques in order to select a suitable fused image for the intended study. The reliability of the metrics used to measure the quality of the images is studied in respect of classification accuracy. The second stage, in order to overcome the limitations experienced in the parametric classification, concentrates on introducing a non-parametric classifier based on the J4.8 decision tree classification (DTC) algorithm -an improved version of Quinlan’s C4.5 algorithm. The classifier’s behaviour in respect of its prediction accuracy, 10-fold cross validation accuracy, overall classification accuracy (OCA), tree complexity, etc. has been studied at various sizes of training data and on two class hierarchy levels by integrating band transformation and band ratioing techniques. In the third stage, to bring out a greater separability between the spectrally overlapping classes, a texture based image
classification approach has been introduced and a methodology is developed to determine the optimal window size, interpixel distance and the best combinations of texture bands in multi-spectral data. The five texture measures, viz. entropy, angular second moment, contrast, mean and homogeneity derived from the grey-level co-occurrence matrix (GLCM) are investigated in the study. The work further focuses on studying the border-effect and effectiveness of employing texture measures at class hierarchy level I and II. Besides, an evaluation of the most commonly employed statistical interclass separability measures - the transformed divergence (TD) and Jeffries-Matusita (JM) distances as feature selection criteria in determining the most effective texture bands in classification has also been carried out.

The major observations and contributions of the thesis are: Among the three conventional image fusion techniques, the Brovey transform method is found to be the best in improving classification accuracy. The statistical measures, which are adopted for the evaluation of image fusion techniques - except TD, do not exhibit concurrency with OCA. The decision tree classifier shows a very promising performance on high resolution data at class-hierarchy levels I and II in comparison with the maximum likelihood classifier. Addition of ancillary image bands and band ratioing significantly reduces the tree complexity in DTC without compromising on accuracy. Combining texture measures with spectral features at optimal window size and interpixel distance has proven to be significantly effective in classification of spectrally overlapping LU/LC features - especially at higher class hierarchy level. The border-effect is observed on the image boundary at higher window sizes but unnoticeable at interclass boundaries within the image. The increase in data dimensionality due to addition of texture bands worsens the performance of the statistical classifier; on the contrary, the non-parametric classifier is consistent in its performance. The statistical feature selection criteria do not exhibit a reliable performance in determining the best performing texture features and optimal feature combinations since texture violates the normal distribution of data in multi-spectral bands.

In summary, this research work has successfully developed a framework to effectively integrate and extract information from non-spectral data to improve classification accuracy of LU/LC features in high spatial resolution RS data. Attempt has also been made to provide answers through empirical studies to some of the dubious issues and contradictory findings in RS image classification with regard to image evaluation metrics, statistical feature selection criteria and border-effect.