Part 2: Change detection for SAR Imagery

(based on Chapter 18 of the Book. Change-detection methods for location of mines in SAR imagery, by Dr. Nasser Nasrabadi, et al.)

* Let **x** and **y** be vectors from reference and test images respectively

- Distance (Euclidean, Mahalanobis) based change detection.

$$\mathbf{e}_{MD}(\mathbf{y}) = (\mathbf{x} - \mathbf{y})^T \mathbf{C}_X^{-1}(\mathbf{x} - \mathbf{y}).$$

- Image ratioing

$$\mathbf{e}_R = \frac{y}{x}$$

- Subspace projection-based change detection

$$e_{SP}(\mathbf{y}) = \mathbf{y}^T \mathbf{C}_X^{-1}\mathbf{y} - \mathbf{y}\mathbf{C}_X^{-1}\mathbf{P}_X\mathbf{y}^T$$

here P$_\mathbf{x}$ is a subspace projection operator.
- Wiener prediction-based change detection.

$$e_{W'}(\mathbf{y}) = (\mathbf{y} - \hat{\mathbf{y}}_W)^T \mathbf{C}_X^{-1}(\mathbf{y} - \hat{\mathbf{y}}_W)$$

here y$_w$ is the Wiener filtered image vector.

The following 15 slides are kindly provided by Dr. Nasser Nasrabadi.

Additional remarks

- Results indicate that in general the more complex algorithms exhibited superior performance as compared to the simple methods; the exception seems to be the ratio method.

- The ratio method, although a simple implementation, has performance which is competitive with the more complex change detection methods implemented.

- Results indicate that taking into account the local spatial and statistical properties of the image improves the change detection performance.

- The ROC curves also show that addition of the $\mathbf{C}_X^{-1}$ term may serve to mitigate false alarms that arise from speckle.

- As expected the simple difference method exhibits the worst performance, having a high occurrence of false alarms.

- The authors also examined Kernel versions of selected change detection methods.

- The results are based on very limited data.

Advanced approaches to change detection in SAR images

(based on chapter 17 of the Book, Unsupervised change detection in muti-temporal SAR images, by Dr. Bruzzone, et al.)

- Nice literature survey

- Deals with images of the same scenes at two different time instants.

- The proposed approach includes:

  Multi-resolution decomposition of log-ratio image.

  Adaptive scale identification

  Scale-driven fusion.

Part 3a: The Classification Problems

Based on Chapters 2,3,16,17,22,25,,26,and 27 of the Book

- Pixel classification and segmentation into regions are typical image classification problems.
- Event classification from signals is another popular classification problem.

* Statistical pattern classification is still very important. The use of ensemble classifiers is popular.

* Neural networks and support vector machines emerge as more practical approaches to the classification problems.

Statistical Classifiers

Advantages of statistical classifiers:

* Theory and techniques are now well developed.

*   The parametric classification rules include the (optimum) Bayes, maximum likelihood decision rules and discriminant analysis. The nonparametric (or distribution free) method of classification includes nearest neighbor decision rule and Parzen's density estimate.

*   Several hundred papers on NNDR have been published. One variation of NNDR is called the voting k near centroid neighborhood rule (k-NCN). An improvement of 1 to 1.5% over k-NNDR can be achieved. Still the original k-NNDR performs well with proper reference data sets.

   (S. Grqabowski, et al., "Nearest neighbor decision rule for pixel classification in   remote sensing, in Frontiers of Remote Sensing Information Processing.

• Feature extraction and selection algorithms have been well studied.

• Contextual information can be incorporated in the decision rule.

Remark: To seek for improved classifier is always a challenge.

Restrictions of Statistical Classifiers

- Accuracy of the statistical information is always of concern especially when the number of training samples

  is small. One solution to "enhance" the statistics is to use unsupervised samples. (work by Prof. Landgrebe).

- The "Hughes phenomenon": For a fixed training sample size, the classification performance first improves with

  the increase in number of features until it reaches a peak

  beyond which further increase in feature number will degrade the performance.

- Difficulty of using higher than second order statistics.

  Most algorithms assume underlying Gaussian statistics.

- Spatial dependence information cannot be fully captured by statistics. Syntactic, semantic, and structural pattern recognition approaches have their useful role though statistical classifiers are

  most popular.

Random Forest (RF) Classification of Remote Sensing Data (based on Chapter 15 of the Book)

- A random forest (RF) classifier is a classifier comprising of a collection of classification trees, each trained on a randomly chosen subset of the input data where final classification is based on a majority vote by the trees in the forest.

- Each tree is trained using a random subset of the training samples.

- During the growing process of a tree the best split on each node in the tree is found by searching through m randomly selected features. For a data set with M features, m is selected by the user and kept much smaller then M.

- Every tree is grown to its fullest to diversify the trees so there is no pruning.

- RF was investigated for hyperspectral and muti-source remote sensing data.

RF classifiers continued-1

- Random forests have been shown to be comparable to boosting in terms of accuracies, but are computationally much less intensive than boosting. Boosting is an ensemble classification method that train several classifiers and combine their results through a voting process.

- Two random forest (RF) approaches are explored; the RF-CART (Classification and Regression Tree) and the RF-BHC (Binary Hierarchical Classifier).

- The RF-CART approach is based on CART-like trees where trees are grown to minimize an impurity measure.

- The RF-BHC approach depends on class separability measures and the Fisher projection, which maximizes the Fisher discriminant where each tree is a class hierarchy, and the number of leaves is the same as the number of classes.
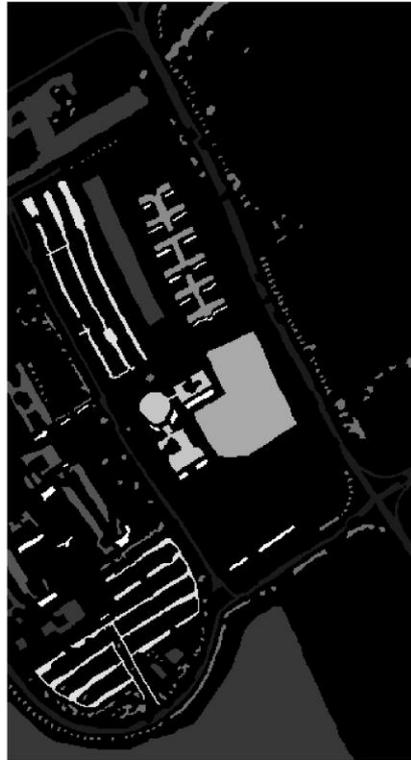
RF classifiers continued-2

- The different RF approaches are compared in experiments in classification of a multisource remote sensing and geographic data from the Anderson River in Canada and urban area from Pavia, Italy using

  hyperspectral ROSIS data.

- RF is attractive for both types of data.

- Both approaches perform well. For ROSIS data,RF-BHC performs better than BHC and RF-CART in test data

- RF approach is fast in training and classification and is distribution free.

* The problem of curse of dimensionality is addressed by the selection of low m.

ROSIS data: reference image (left), gray scale image (right)

Class color   a) University ground truth   b) University pseudo color(gray) image

The BHC tree used for classification of ROSIS data.