International Journal of Remote Sensing & Geoscience (IJRSG)
www.ijrsg.com

USE OF CI TEST IN CHANGE DETECTION ANALYSIS

Subhanil Guha, Department of Geography, Dinabandhu Andrews College, Kolkata, India

Abstract

Bayesian network approach is a recent and sophisticated method used in the processing of satellite data. In this paper, this new approach based on Conditional independence (CI) test algorithm has been applied significantly to detect land use changes in two multi date satellite image for a part of Dehradun area. The experimental classification results with an overall accuracy of 90% establish that Bayesian network is a very effective approach for change detection analysis.

Keywords: Bayesian network; Conditional independence test; Change detection

Introduction

Remote sensing technique is frequently using to detect changes in land use and land cover in any region. In the last 40 years, a large number of change detection techniques have been introduced. The remote sensing change detection method usually can be divided into change mask development (CMD) and categorical change extraction (CCE) (Dai and Khorram, 1999). Most change detection methods act as CMD method, e.g. Image Rationing, Image Differencing. These methods usually not able to find the differences in the types of land use changes. The CCE method can provide the change category by comparing two different time imageries. Post-Classification Comparison can get the changes through comparison of the classification results map of the multi temporal data, depends on the accuracy of the individual classification. At present the neural network change detection (Chen et al., 2004) belongs to the Direct Multi-Date Classification method. It does not have data distribution premise (Hasi et al., 2003). In this paper, a comparatively new direct multi-date classification method based on Bayesian network has been applied. It is based on conditional independence (CI) test and try to avoid the Post classification error and the directed acyclic graph (DAG) of Bayesian network can well express the mutual information of multi-temporal data. The basis of this network (Pearl, 1988) is probability theory which is a powerful reasoning tool under conditions of uncertainty. In recent years, the approach has been frequently applied in medicine or artificial intelligence (Heckerman et al., 1995; Pearl, 1988; Friedman and Elidan, 2001). However, the technique has seldom been applied to process remote sensing image data. Bayesian network classifier has been constructed by identifying the CI relationships among the nodes. With the help of the mutual information test we can find the CI relationships among the attributes. Part of Dehradun has been taken as a study area due to its recent rapid changing nature in land use type.

Bayesian network method

Bayesian network method has recently been established as an authentic as well as successful technique for representing knowledge and inference under uncertain conditions. It represents the inter-relationships among the dataset attributes (Cheng and Greiner, 2001) which indicates its superiority over the other probabilistic models.

Bayesian network theory

The Bayesian network theory entirely follows the probability theory and graph theory. Here the nodes are considered as random variables and edges are considered as conditional dependencies in a directed acyclic graph (DAG). A DAG is a directed graph without having any directed cycles. The links or arcs indicate the interactions among the variables. The links are not capable to form a closed loop, that keeps to the DAG. A node *A* with no entering arcs called a source node and a table having unconditional probabilities *p(A)* will be assigned to it. Again, when a variable *A* receives edges from other variables *B1;...; Bn*, then *A* will be known as child variable and the variables *B1;...; Bn* will be considered as parental variables. The conditional node probabilities will be contained in the node probability table of variable *A*. The edges are calculated using a set of conditional probability tables. The tables express the probability for each specific variable in a particular state, given the states of its parents. The structure of network represents joint probability which is given by Eq. (1), where pa_i stands for the parental nodes of node x_i:

$$P(x) = \overset{n}{\underset{i=1}{P}} p(x_i/pa_i) \quad (1)$$

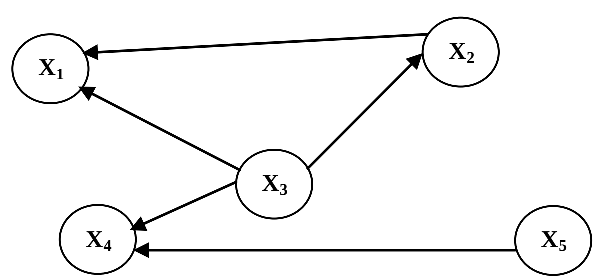


Figure 1. Example of a Bayesian network graph

An illustration of a Bayesian network has been given in Fig. 1.

$$p(x_1, x_2, x_3, x_4, x_5) = p(x_1/x_2, x_3)p(x_2/x_3)p(x_3)p(x_4/x_3, x_5)p(x_5) \quad (2)$$

Conditional Independence test (CI test)

The dependency relationships of Bayesian network construction are measured by using CI test. The CI test is based on the examination that two nodes x_i and x_j are conditionally independent given a set of nodes C or not. Two nodes x_i and x_j are considered as conditionally independent given a set of nodes C only when the conditional mutual information of the nodes is smaller than a threshold values ε. The mutual information of two nodes x_i and x_j may be defined as-

$$I(x_i, x_j) = \underset{x_i, x_j}{S} p(x_i, x_j) \log p(x_i, x_j) / p(x_i) p(x_j) \quad (3)$$

The conditional mutual information can be computed by the following equation-

$$I(x_i, x_j | C) = \underset{x_i, x_j, c}{S} p(x_i, x_j, C) \log p(x_i, x_j | C) / p(x_i | C) p(x_j | C) \quad (4)$$

Bayesian network algorithm is derived from CI relationships among nodes and CI test method (Cheng et al., 2002). It includes three phases i.e., drafting, thickening and thinning:

- Computing numerical statistics of each pair of nodes and hence create a draft;
- Adding the edges in the network using CI test when the two nodes of the edge are not conditionally independent;
- Removing the edges in the network using CI test if the two nodes of the edge are conditionally independent.

Change detection

The remote sensing data change detection procedure includes data preprocessing and change detection process. Bayesian network change detection method has been applied in the present study (Fig. 2).

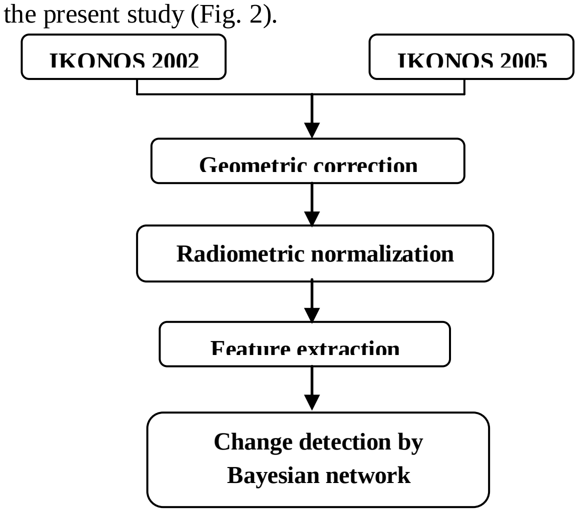
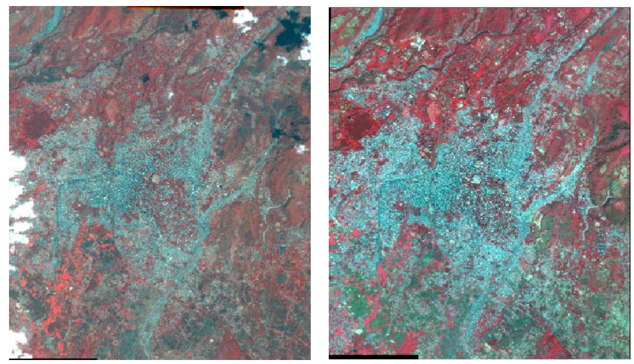


Figure 2. Flow chart of methodology

Study area and Satellite data

The study area is located within the Dehradun city, capital of the Uttarakhand state, India. Latitudinal and longitudinal extent of the study area is from 30°16'N to 30°22' N and from 78° 00' E to 78°06' E respectively. The two IKONOS images were acquired on November 4, 2002 and on November 7, 2005 (Fig.3). For the purpose of this research it was required that multi-temporal data should be of near anniversary dates so that the seasonal variation in atmospheric condition is kept at minimum level. The satellite data of first week of November has been considered suitable for this study.



IKONOS-2002 IKONOS-2005

Figure 3. Study area in IKONOS satellite image

Data preprocessing procedures

In the study data preprocessing procedures include geometric correction, radiometric normalization and feature extraction.

Geometric registration

Firstly, the IKONOS image 2002 has been geometrically registered with 1:50,000 topographical map on Universal Transverse Mercator projection. The reference points are evenly distributed in the study area, and the root mean square error of the registration process is 0.4638, and then IKONOS image 2005 has also been geometrically registered using image-matching method referring to the image of 2002.

Radiometric normalization

Here, pseudo invariant feature method has been used for radiometric normalization of 2002 MSS image with respect to 2005 MSS image (Schott, Salvaggio and Volchok, 1988). This method involves analysis of in-scene man induced elements (road, agricultural land, urban land etc.). Total 64 training sites satisfying the criteria that the sites should have approximately the same elevation, contain only minimal amounts of vegetation and must not change with time, and have a wide range of DN values (Fig.4).

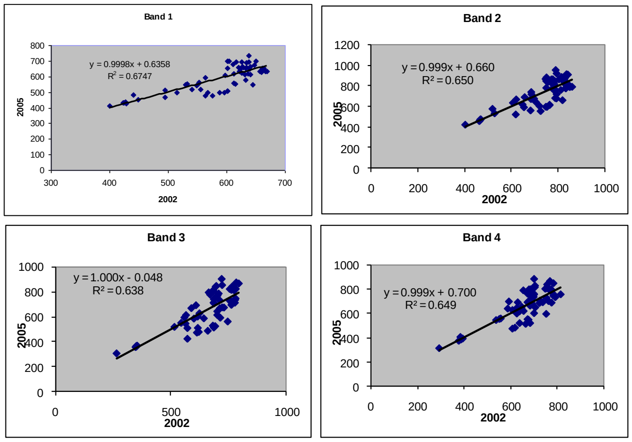


Figure 4. Band wise regression analysis of pseudo invariant features of image 2002 and 2005

Feature extraction

Method of Principal component analysis has been applied for extracting the suitable features so that redundancy of data is minimized.

Bayesian network change detection procedures and results

After completion of the preprocessing stage, the training data has been selected in the combination image of two dates and the standard false color composite image.

Training data

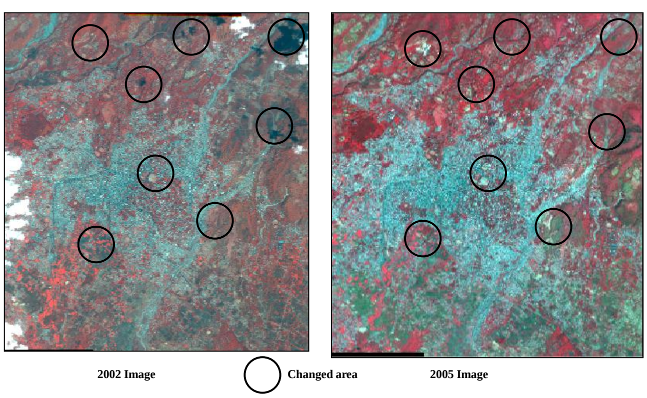


Figure 5. Visual interpretation of change classes (2002-2005)

To detect the changes in land use type brightness values of both the images have been compared. Water body, natural vegetation, agricultural land, fallow land and built up area have been used as different types of land use. Training data has also been selected from the given land use types and these are based on the DN values. The final part of the analytical study is accuracy assessment. Here, the changes in the land use are clearly visible from the two multi date images (2002 & 2005) through visual interpretation. Some of these changes have been marked by circles (Fig. 5). The change classes of land use training data are represented in Table 1.

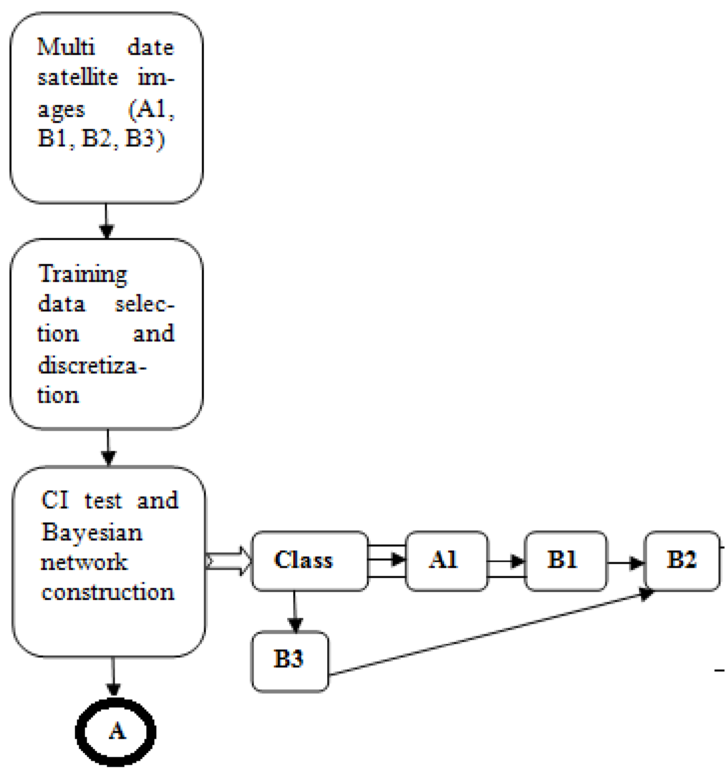
Table 1. Classes of land use change training pixels from 2002 to 2005

Change classes	Number of training pixels
Natural vegetation to Agricultural land	720
Natural vegetation to Fallow land	390
Natural vegetation to Built up area	1173
Water body to Agricultural land	108
Water body to Fallow land	200
Water body to Built up area	823
Agricultural land to Fallow land	114
Agricultural land to Built up area	759
Fallow land to Agricultural land	170
Fallow land to Built up area	844
Built up area to Built up area	822
Total Change	6123

Change detection procedures

The whole process of Bayesian network method can be divided into the following five steps given in Fig. 6.

(i) Combining the multi date data sets of four bands A1, B1, B2 and B3 into one data set as input data for change detection analysis.



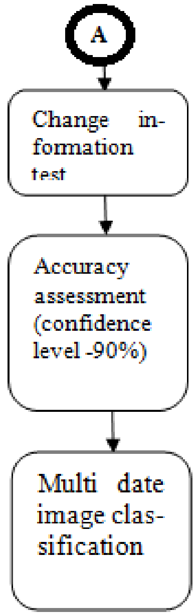


Figure 6. Flow chart of Bayesian network

(ii) The training data for change classes has five kinds of attributes, four attributes describe the brightness values of two images and one attribute denotes change class label. Then all the input attributes of training data have been discretized using entropy minimization discretization method (Fayyad and Irani, 1993).

(iii) Training the Bayesian network structure using CI test algorithm. This training result of Bayesian network structure is shown in Fig. 5. The graph is comprises with 5 Directed Acylic Arcs. These are - class to A1, class to B1, class to B2, class to B3, B3 to B2.

(iv) The accuracy assessment for change detection regarding to confidence level as 90%.

(v) Multi date satellite data land use classification using Bayesian network method to detect the changes.

Accuracy assessment

An overall accuracy of 90% has been achieved using this classification technique. Most of the land use changes have been noticed into the recent growing urban area (Table 2).

Table 2. Land use changes transformation matrix from 2002 to 2005

2002/2005	Natural vegetation	Water Body	Agricultural land	Fallow land	Built up area	Total
Natural vegetation	0.00	0.00	0.12	0.07	0.19	0.38
Water Body	0.00	0.00	0.02	0.03	0.13	0.18
Agricultural land	0.00	0.00	0.00	0.02	0.12	0.14
Fallow land	0.00	0.00	0.03	0.00	0.14	0.17
Built up area	0.00	0.00	0.00	0.00	0.13	0.13
Total	0.00	0.00	0.17	0.11	0.72	1.00

Discussion and concluding remarks

The present study has developed and implemented the methodologies and algorithms of constructing Bayesian network mainly based on CI test. On the basis of the experiments, the Bayesian network model for change detection shows a great potential as a new effective method. The merits of this method are: (1) It can examine the multi-temporal images because CI test algorithm can represent the information dependence between two images of different dates; (2) It can generate accurate land use change classes; (3) It is less erroneous; (4) It merges prior knowledge with data information. Hence, it can be applied as a authentic tool in land use change detection analysis for multi-temporal satellite imageries.

Acknowledgements

The author is thankful to IJRSG Journal for the support to develop this document. The author is also grateful to his guide Prof. S.C. Mukhopadhyay, UGC Emeritus Fellow and Former Head of the Department of Geography, Calcutta University for his guidance, mental support and valuable suggestions. The author is also thankfully acknowledged to Indian Institute of Remote Sensing and Survey of India, Dehradun for providing the satellite images and topographical maps.

References

[1] Chen, X., Dai, Q., Ma, J., 2004. Research on direct classification change detection for temporal and multi-spectral remote sensing data using neural networks. Computer Engineering and Applications 49 (28), 12–15 (in Chinese).

[2] Cheng, J., Greiner, R., 2001. Learning Bayesian belief network classifiers: algorithms and system. In: Proceedings of the 14th Canadian Conference on Artificial Intelligence. Ottawa, Canada, pp. 141–151.

[3] Cheng, J., Greiner, R., Kelly, J., Bell, D.A., Liu, W., 2002. Learning Bayesian networks from data: an information theory based approach. The Artificial Intelligence Journal 137 (1–2), 43–90.

[4] Dai, X.L., Khorram, S., 1999. Remotely sensed change detection based on artificial neural networks. Photogrammetric Engineering and Remote Sensing 65 (10), 1187–1194.

[5] Fayyad, U.M., Irani, K.B., 1993. Multi-interval discretization of continuous-valued attributes for classification learning. In: Proceedings of the 13th International Joint Conference on Artificial Intelligence. Chamberry, France, pp. 1022–1027.

[6] Friedman, N., Elidan, G., 2001. Learning the dimensionality of hidden variables. In: Proceedings of Seventeenth Conference on Uncertainty in Artificial Intelligence (UAI), Seattle, Washington, USA, pp. 144–151.

[7] Hasi, B., Ma, J., Li, Q., Han, X., Liu, Z., 2003. The self organizing feature map neural networks classification of the Aster data based on wavelet fusion. China Science (D) 33 (9), 896–902.

[8] Heckerman, D., Mamdani, A., Wellman, M.P., 1995. Real world applications of Bayesian networks. Communications of the ACM 38 (3), 24–26.

[9] Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers, San Francisco, CA.

Biography

SUBHANIL GUHA received the B.Sc. degree in Geography from the University of Calcutta, Kolkata, West Bengal, the M.Sc. degree in Geography from the University of Calcutta, Kolkata, West Bengal, and the P.G. Diploma in Geoinformatics from the Indian Institute of Remote Sensing, Dehradun, Uttarakhand and International Institute for Geoinformation Science and Earth Observation, Enschede, Amsterdam, The Netherlands, respectively. Currently, He is in the Department of Geography in Dinabandhu Andrews College, Kolkata, India. His teaching and research areas include Geography and Geoinformatics. He has authored/co-authored a textbook chapter: Change in Mangrove Forest Cover and Deltaic Islands in Sundarban Areas of West Bengal: A temporal analyses using NCI technique on LANDSAT TM5 data, In "Sundaresan, J. ; Sreekesh, S. ; Ramanathan, A.L. ; Sonnenschen, Lenand ; Boojh, Ram. , Climate Change and Environment", (Chapter 12). pp. 113-127. (ISBN: 9788172338336), 2013; an international paper: Vegetation dynamic of Kolkata and its surrounding area**,** Environment and Natural Resources Research, Canadian Center of Science and Education, 2(4), pp. 18-30. (ISSN: 1927-0488 (Print), ISSN: 1927-0496 (Online)), 2012; an international paper: Landuse/ Land Cover Change in Hugli Estuary using FUZZY C-mean Algorithm, International Journal of Geomatics and Geosciences, 2(2), pp. 613-626. (ISSN: 0972-4380), 2011. Author may be reached at subhanilguha@gmail.com