EFFICIENT INCORPORATION OF MARKOV RANDOM FIELDS IN CHANGE DETECTION

H. Aanæs[1], *A. A. Nielsen*[2], *J. M. Carstensen*[1], *R. Larsen*[1] *and B. Ersbøll*[1]

DTU Informatics[1] & DTU Space[2],
Technical University of Denmark,
DK-2800 Kgs. Lyngby, Denmark
e-mail: `haa@imm.dtu.dk`

Remotely sensed Earth observation data are typically collected as reflection or emission of electromagnetic radiation from the surface of the Earth. The data are collected globally and often in digital image format and they cover the same geographical area after some period of time. For many Earth observation platforms data have been available for decades. Therefore one of the most important applications of this type of data is the study of spatio-temporal dynamics of the surface of the Earth related to areas as agriculture, forestry, environment, oceanography, etc. Such studies include change detection over time in multi- and hyperspectral digital image data.

A large class of change detection methods works by calculating a statistical probability of change for each pixel, as seen in Figure 1. This has the disadvantage that in many applications one is looking for regions of change, and such information is, by definition, not used in pixel-wise classification. To elaborate, an underlying problem is that the probability distributions describing change and no-change have considerable 'overlap', and thus there is a lot of noise in the pixel-wise change detection result, as seen in Figure 1 bottom center. In particular clearly defined regions of change are hard to identify. Thus there is a need to incorporate a notion of local homogeneity in the modelling of the problem, i.e. the solution to the change detection should be made under consideration to the fact that regions are sought, and thus a singe pixel of change in a large region of no-change is very unlikely. This is the issue which we address here.

A good way of incorporating such homogeneity constraints into the modelling of the problem is via Markov Random Fields (MRF) cf. e.g. [1, 2], which have been used successfully to solve similar problems before, and which have a very well developed and understood underlying theory. This is what we propose doing here. So instead of just classifying based on the probability measure, $P_{image}(x_i)$, which is solely dependent on the individual pixels, x_i, we propose classifying based on the following measure:

$$P(x) = \frac{1}{Z} \exp\left(\sum_i \left\{ \log\left(P_{image}(x_i)\right) - \beta \sum_{j \in \mathcal{N}(x_i)} n_{ij} \right\} \right) \quad , \tag{1}$$

where Z is a normalization constant and n_{ij} is a function that is one if pixel x_i and x_j are assigned the same class[1] and zero otherwise. The neighborhood, $\mathcal{N}(x_i)$, is here chosen to be a standard 4-neighborhood. This is the so-called Ising model [2], which was first used for describing magnetism in iron. The β parameter determines the degree of regularization, and is in our experiments set to 1. The result is depicted in Figure 1 bottom right, and it is seen that the result has clearly defined regions, which still correspond well to the probability measure in Figure 1 bottom left.

The use of such homogeneity constraints in change detection is ongoing work, and we aim at comparing with many other probability functions, and extending to more than two classes (i.e. change vs. no-change). We, however, feel that the results presented here are encouraging. The pixel-wise probability measure, $P_{image}(x_i)$, used here is from the so-called IR-MAD change detector [3, 4] which in turn is based on canonical correlation analysis, [5]. IR-MAD stands for iteratively re-weighted multivariate alteration detection.

Previously, a problem with using Markov Random Fields for such tasks has been optimization, or numerically finding an optimal solution. Typically simulated annealing or other stochastic optimization techniques were used. These methods were slow and seldom gave the optimal result c.f. e.g. [6]. It was however, relatively recently discovered that graph based techniques could be used to solve these problems, yielding highly efficient algorithms achieving guaranteed optimal results in the two class case, cf. e.g. [7]. We naturally employ these methods in this work. Doing so ensures that the results achieved are a result of one's modelling of the problem, and are not influenced by stochastic artifacts due to the optimization method, and enables more rigorous experimentation with regularization parameters — here β — e.g. in a regularization path framework [8, 9].

We thus propose using the novel combination of the IR-MAD change detector with Markov random fields, and do it in a efficient way which is guaranteed to achieve the optimum of the objective function. The results are promising. Other change detection schemes based on MRF are [10, 11], which use automatic threshold selection and the expectation-maximization (EM) algorithm, [12], for unsupervised estimation of the statistics of the changed and unchanged pixels in the simple difference image.

1. REFERENCES

[1] G. Winkler, *Image Analysis, Random Fields and Markov Chain Monte Carlo Methods: A Mathematical Introduction (Applications of Mathematics)*, Springer Verlag, 2003.

[1]Here there are two classes; change and no-change.

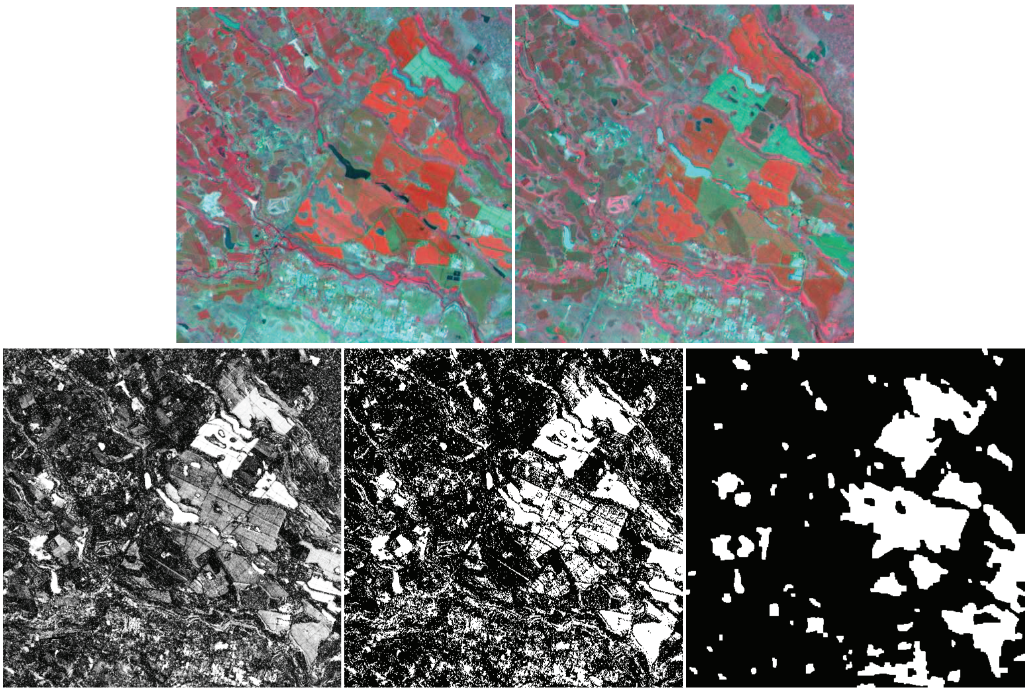

Fig. 1. Illustration of our proposed method. The data used in the example are 512 by 512 20 m pixels SPOT HRV multispectral data acquired on 5 Feb 1987 and 12 Feb 1989. The geographical region covered is an agricultural area with large scale pineapple (center and right), small scale coffee (left) and the town of Thika (bottom) which lies to the immediate north of Nairobi, Kenya. (here shown with the photo-infrared band as red, red as green, and green as blue) **Top:** The input images of Thika taken in 1987 (left) and compared to an image of the same region in 1989 (right). **Bottom left:** A probability measure of change for each pixel, i.e. $P_{image}(x_i)$, white indicating a high probability of change. **Bottom center:** Change detection result, by finding pixels in b) with a higher probability for change then no-change. Black indicates change, i.e. $P_{image}(x_i) > 0.5$. **Bottom right:** The proposed method where the probability in b), i.e. $P_{image}(x_i)$, is combined with a homogeneity term.

[2] S.Z. Li, *Markov Random Field Modeling in Image Analysis*, Springer Verlag, 2nd edition, 2007.

[3] A.A. Nielsen, "The regularized iteratively reweighted mad method for change detection in multi- and hyperspectral data," *IEEE Transactions on Image Processing*, vol. 16, no. 2, pp. 463–478, 2007.

[4] M.J. Canty and A.A. Nielsen, "Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted mad transformation," *Remote Sensing of Environment*, vol. 112, no. 3, pp. 1025–1036, 2008.

[5] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3/4, pp. 321–377, 1936.

[6] D. M. Greig, B. T. Porteous, and A. H. Seheult, "Exact maximum a posteriori estimation for binary images," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 51, no. 2, pp. 271–279, 1989.

[7] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.

[8] K. Sjöstrand, M.S. Hansen, H.B. Larsson, and R. Larsen, "A path algorithm for the support vector domain description and its application to medical imaging," *Medical Image Analysis*, vol. 11, no. 5, pp. 417–428, 2007.

[9] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net.," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301, 2005.

[10] L. Bruzzone and D.F. Prieto, "Automatic analysis of the difference image for unsupervised change detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 3, pp. 1171–1182, 2000.

[11] L. Bruzzone and D.F. Prieto, "An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images," *IEEE Transactions on Image Processing*, vol. 11, no. 4, pp. 452–466, 2002.

[12] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the em algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.