AUTOMATED LAND COVER CHANGE DETECTION: THE QUEST FOR MEANINGFUL HIGH TEMPORAL TIME SERIES EXTRACTION

*†‡B.P. Salmon, †*J.C. Olivier, †‡W. Kleynhans, ‡K.J. Wessels and ‡F. van den Bergh*

†Department of Electrical, Electronic and Computer Engineering, University of Pretoria, South Africa

‡Remote Sensing Research Unit, Meraka Institute, CSIR, Pretoria, South Africa
bsalmon@csir.co.za

*Defense, Peace, Safety and Security Unit, CSIR, Pretoria, South Africa

1. INTRODUCTION

The transformation of natural vegetation by practices such as deforestation, agricultural expansion and urbanization, has significant impacts on hydrology, ecosystems and climate [1]. Coarse resolution satellite data provide the only regional, spatial, long-term and high temporal measurements for monitoring the earth's surface. Automated land cover change detection at regional or global scales, using hyper-temporal, coarse resolution satellite data has been a highly desired but elusive goal of environmental remote sensing [2, 3].

A time series is a sequence of data points measured at successive time intervals. Time series analysis comprises methods that attempt to understand the underlying force structuring the data, identifying patterns, detecting changes and clustering. Subsequence clustering is performed on streaming time series that are extracted with a sliding window from an individual time series [4]. A subsequence $x_p(t)$ for a given time series $x(t)$ of length N, is given as

$$x_p(t) = \left[\, x(t_p)\; x(t_{p+1})\; \ldots\; x(t_{p+Q}) \,\right], \tag{1}$$

for $1 \leq p \leq N$-Q+1, where Q is the length of the subsequence. The sequential extraction of subsequences in (1) is achieved by using a temporal sliding window that has a length of Q and position p that is incremented with a natural number $\mathbb{N}$ to extract sequential subsequences $x_p(t)$ from $x(t)$. The signal processing and data mining communities have made wide use of the clustering of subsequence time series, $x_p(t)$, that were extracted using a temporal sliding window. However, it has found very limited applications on satellite time series data.

Recently the data mining community's attention was brought to a fundamental limitation of clustering subsequences of a time series that were extracted with a sliding window [4]. The sliding window causes clustering algorithms to form sine wave cluster centers regardless of the data set, and clearly makes it impossible to distinguish one dataset's clusters from another. This is due to the fact that each data point within the sliding window contributes to the overall shape of the cluster center as the window moves through the time series [4]. This limitation was illustrated by using data sets from various fields, i.e. stockmarket and a random walk data sets. Keogh and Lin [4] demonstrated a tentative solution that would not suffer from the aforementioned limitation when the procedure was applied to a periodic data set and the sliding window position p was incremented

This research was funded by the CSIR Strategic Research Panel.

by the exact length of the periodic cycle. Since remote sensing time series data has a very strong periodic component due to seasonal vegetation dynamics, the extracted sequential time series could potentially be processed to yield usable features. These features could enable effective subsequence clustering and potentially be used for change detection.

Land cover change in context is defined here as the assignment of subsequences that are extracted from a time series that transition from one cluster to a different cluster and remains there for the rest of the time series.

The objective of this paper is to introduce the concept of unsupervised land cover change detection algorithm that operates on a temporal sliding window of MODIS time series data that uses a feature extraction method that does not suffer from the limitation shown by Keogh and Lin [4]. Two well-known unsupervised clustering techniques were used within a land cover change detection algorithm and were evaluated specifically on new settlement development, both real and simulated land cover change, using the 8-day composite MODIS land surface reflectance data product.

2. METHODOLOGY

2.1. Study Areas

The area of interest was the Limpopo province which is situated in the northern part of South Africa. The province is still largely covered by natural vegetation used as grazing for cattle and wildlife. The development of settlements is one of the most pervasive forms of land cover change in South Africa. The area within the province was selected where settlements and natural vegetation occur in close proximity to ensure that the rainfall, soil type and local climate were similar over both land cover types. The selected areas of interest in the study area is composed of 433.75km^2 of natural vegetation and 374.25km^2 of human settlements.

2.2. Feature extraction - Subsequence Time Series

In this section a method is shown that will create usable features from time series $x_p(t)$ extracted from MODIS data. The fixed acquisition rate of the MODIS product and the seasonality of the vegetation in the study area makes for an annual periodic signal $x(t)$ that has a phase offset that is correlated with rainfall seasonality and vegetation phenology. The Fast Fourier Transform (FFT) of $x_p(t)$ was computed, which decomposes the time sequence's values into components of different frequencies with phase offsets. Because the time series $x_p(t)$ is annually periodic, this would translate into frequency components in the frequency spectrum that have fixed positions. This can be viewed as a fixed location for each of the features for the clustering algorithm in the feature space regardless of the sliding window position in time, which overcomes the main disadvantage to a sliding window [4]. Because of the seasonal attribute typically associated with MODIS time series and the slow temporal variation relative to the acquisition interval, the first few FFT components dominate the frequency spectrum.

Keogh and Lin [4] suggested that the sliding window position p should be shifted by a complete periodic cycle [4], but by computing the magnitude of all the FFT components removes the phase offset, which makes it possible to compensate for both the restrictive position p of the sliding window and the rainfall seasonality. The features $X_p(f)$ for the clustering method were extracted from the sliding window $x_p(t)$ by the methodology discussed above as

$$X_p(f) = |\,\mathcal{F}(\,x_p(t)\,)\,|, \tag{2}$$

where $\mathcal{F}(\cdot)$ is the FFT function. The mean and annual FFT components from (2) were considered as it was shown in [5] that considerable class separation can be achieved from these components.

2.3. Unsupervised change detection

The clustering method was required to process subsequences of time series data and detect land cover change as a function of time. Land cover change is declared when consecutive subsequences that are extracted from one MODIS time series, transitions from one cluster to another cluster and remains in the new assigned cluster for the rest of the time series. The temporal sliding window was designed to operate on a subsequence of the time series to extract information from two spectral bands from the MODIS product. These features were analyzed with two different clustering techniques: Ward and K-means.

3. CONCLUSIONS

In this paper, a method for unsupervised land cover change detection incorporating a temporal sliding window, operating on MODIS time series data was demonstrated. The unsupervised approaches reported *true positive* measurements of higher than 70.5% on all simulated land cover change using cross validation. The results for the detection of simulated land cover change was compared to real mapped settlement development and a *true positive* accuracy of 76.12% was achieved. The difference in change detection accuracy between the real and simulated land cover change were still acceptably close in these experiments, even though only a limited number of real land coversion examples were available.

Since the MODIS time series has a very strong periodic component due to seasonal vegetation growth, it provides the remote sensing community with a special type of data which, if processed correctly, is immune to the limitation pointed out by Keogh and Lin [4]. This is mainly due to the extraction process which produced a short-term FFT that fixed the features' positions, which allows the features to be analyzed and permits the temporal sliding window to be moved at any time increment. This should rekindle the remote sensing community's quest for automated change detection using time series as it allows them to use many different types of algorithms and methodologies on sequential time series extracted from satellite data.

4. REFERENCES

[1] R. S. DeFries, L. Bounoua, and G. J. Collatz, "Human modification of the landscape and surface climate in the next fifty years," *Global Change Biology*, vol. 8, no. 5, pp. 438–458, May 2002.

[2] J. R. G. Townshend and C. O. Justice, "Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations," *International Journal of Remote Sensing*, vol. 9, no. 2, pp. 187–236, February 1988.

[3] T.R. Loveland et al., "A strategy for estimating the rates of recent united states land-cover changes," *Photogrammetric Engineering and Remote Sensing*, vol. 68, no. 10, pp. 1091–1099, October 2002.

[4] E. Keogh and J. Lin, "Clustering of time-series subsequences is meaningless: implications for previous and future research," *Knowledge and Information systems*, vol. 8, no. 2, pp. 154–177, August 2005.

[5] S. Lhermitte et al., "Hierarchical image segmentation based on similarity of NDVI time series," *Remote Sensing of Environment*, vol. 112, no. 2, pp. 506–521, February 2008.