IMPROVING FEATURES USED FOR LAND COVER CHANGE DETECTION BY REDUCING THE UNCERTAINTY IN THE FEATURE EXTRACTION METHOD

*‡B.P. Salmon, *†W. Kleynhans, ‡J.C. Olivier and *†C.P. Schwegmann*

‡School of Engineering and ICT, University of Tasmania, Australia
†Remote Sensing Research Unit, Meraka Institute, CSIR, South Africa
*Department of Electrical, Electronic and Computer Engineering, University of Pretoria, South Africa
bpsalmon@utas.edu.au

ABSTRACT

The well-being of the environment is one of the major factors that contributes to sustainability. Sustainable human settlements require local governance to plan, implement, develop, and manage human settlements expansions. This is important as the number anthropogenic activities is directly correlated to the increase in human population within a geographical region. Regional mapping of land cover conversion of natural vegetation to new human settlements is essential. In this paper we explore the effect which the length of a temporal sliding window has on the success of detecting land cover change. It is shown using a short Fourier transform as a feature extraction method provides meaningful robust input to a machine learning method. In theory, the performance is increased by improving the estimates on the features by increasing the length of the sliding window. Experiments were conducted in the Limpopo province of South Africa and were found that increasing the length of the sliding window beyond 12 months yield minor improves due to other seasonal and external factors.

***Index Terms*—** Change detection, Fourier transform, Satellite, time series

1. INTRODUCTION

Land cover classification and change detection is the spectral and spatial analysis of either bi-temporal (2) or multi-temporal (3 or more) images [1]. Many different approaches attempt to extract meaningful features and classifying/detect change land cover using a machine learnong method. This result in many different applications being processed fast and efficiently, leaving a fraction of processes still reliant on intense human-operator-dependent mapping.

In our study area the mapping of informal human settlement usually requires human interpretation as these settlements vary quite significant in imagery. The physical area of an informal settlement in our study area only spans a few contiguous pixels but does cause major disturbance in surrounding vegetation. This allows the layout of the settlement to be easily camouflage with the distributed vegetation. An robust change detection method can reduce the amount of work on human interpreters by detecting these new settlements reliably.

Most land cover change detection methods can be broadly categorized into change index and post-classification change detection methods. The change index method derives a change index which declares change if a predefined threshold is exceeded. Post-classification change detection methods derives the change in land cover class between images. Change index methods usually have a higher change detection accuracy than post-classification change detection methods as it is only computes if a change has occur, not the type of transition.

Bi-temporal change detection methods are very accurate as long as seasonal compensation is made to the image, e.g. images are taken in the same season or using training data to adjust for it. It also unfortunately can not assist with ensuring that persistent land cover change has occurs. Multi-temporal change detection methods can be more complex but have the ability to derive the normal seasonal variations of the land cover classes to reliably detect change. Medium resolution remote sensing data is used to ensure that the temporal revisit time is high enough to decipher the seasonal variation from change events.

In this paper we asked the question what is the fundamental lower bound which a change detection method can achieve and how can we lower that bound? Before we answer, we would first mention that certain research questions can not be answered using satellite imagery which will not be covered here. Second we are restricting the investigation to our imagery and ground truth maps with the adequate support of mathematics for our statements. Based on these assumptions, we will derive the Cramer-Rao Lower bound for our features and show in experimental results how these bounds can be used to improve change detection accuracy [2].

The disadvantage is some method's true validity is lost in highly convoluted algorithms which only report minor if even any significant improvements. We start our analysis by investigating performance of a method and explain mathematical the limitations of this method and fundamental challenges faced by the remote sensing community.

(a) 23 April 2001 - beginning of study period

(b) 17 June 2009 - after study period

(c) 31 March 2016 - most recent image

Fig. 1: High resolution image taken near Daring ($23°49'21.3''$S, $29°21'30.24''$E) in the beginning of the study period (top), shortly after the end of the study period (middle) and the most recent image available (bottom) respectively (courtesy of Google Earth).

The paper is organized as follows. Section 2 discusses the data set. In section 3 we present and discuss the lower bound on our feature extraction method for detecting land cover change and how we can improve it. In section 4 we present our results to support our findings. Section 5 presents our discussion and conclusions.

2. STUDY AREA AND SATELLITE DATA

Our study area is based in the Limpopo province which is in north South Africa. The area is mainly covered by natural vegetation and subjected to massive expansions of new informal settlements. The expansion of informal human settlements is often unplanned and result in a disturbance to the local vegetation.

Both natural vegetation and informal human settlement areas undergo seasonal variations in their recorded satellite reflectance values. Multi-temporal images enable the extraction of an time series for each pixel to detect transition between land cover classes. Time series in our experiments were extracted from the MCD43A4 Bidirectional reflectance distribution function corrected MODIS product.

The product combines both recording from the Terra and Aqua satellites to provide 8-day composite 500 meter spatial resolution acquisitions in seven spectral bands. Missing data were filled by interpolating through temporal neighbors using a cubic spline algorithm. The study period for our experiments was February 2000 to December 2008 with all time series inspected visually using high resolution SPOT2 imagery in 2000-2001 and SPOT5 imagery 2006-2008 (figure 1). All time series mapped as change had at least 70% of their area altered. Due to the lack of available high resolution imagery the exact date of new human settlements change could not determined.

3. PROPOSED METHOD

A time series of N-samples extracted from the MODIS product for a pixel is

$$\vec{x}_b = \{x_{n,b}\}_{n=1}^{n=N} = \{x_{1,b}\, x_{2,b} \ldots x_{N,b}\}, \tag{1}$$

where the time index is denoted by $n, \in [1, N]$ and the spectral band by $b \in [1, \ldots, 7]$. Time series $\vec{x}_b$ is wide-sense stationary if the mean μ_b and variance σ_b^2 is independent of time. If a time series is not undergoing any land cover change (or any severe environmental changes such as drought) it can shown the time series approximates a wide-sense stationary process as it exhibits a symmetrical response in its auto-correlation function.

A temporal sliding window is used to extract sub-sequences from the time series to process with a machine learning method to detect land cover change. A sub-sequence $\vec{x}_b(t)$ for a given time series $\vec{x}_b$ is given as

$$\vec{x}_b(t) = \{x_{n,b}\}_{n=t}^{n=t+Q}, \tag{2}$$

where t denotes the time index within the time series and Q the length of the temporal sliding window. In [3], the meaningful extraction of time series was investigated and it was determined that a proper feature extraction method is needed to ensure proper analysis of a temporal sliding window. These limitations were overcome by computing the magnitude of all the Fourier transform components which removes all the phase offsets. This made it possible to compensate for both the restrictive position of the temporal sliding window and the rainfall seasonality [3]. The features presented to the machine learning method can be expressed as

$$\vec{X}_b(f) = \left|\mathcal{F}\left(\vec{x}_b(t)\right)\right|, \tag{3}$$

where $\mathcal{F}(\cdot)$ denotes the Fourier transform function. Machine learning methods attempt to compute the maximum likelihood class estimations based on the features presented. The feature extraction method estimates a set of parameters from an initial data set intended to be informative and non-redundant. It does not create additional information but more intuitive for methods to learn differences in land cover classes. An assumption of the Fourier analysis is that the frequencies within the temporal sliding window are multiples of the fundamental frequency $1/Q$ and that only white Gaussian noise is present. A windowing technique such as Hanning, Hamming or Blackman window could be used to minimize the spectral smearing. If this assumption holds for orthogonality between frequencies then feature estimates could be accurately approximated using

$$\hat{f}_k = \frac{2}{Q}\sum_{q=t}^{t+Q-1} x_{n,b}(t)\cos\left(\frac{2\pi kq}{Q}\right), \tag{4}$$

$$\hat{g}_k = \frac{2}{Q}\sum_{q=t}^{t+Q-1} x_{n,b}(t)\sin\left(\frac{2\pi kq}{Q}\right). \tag{5}$$

These estimates $\hat{f}_k$ and $\hat{g}_k$ are offset due to the present of white Gaussian noise, which given the linear model the estimates would approximate

$$f_k = \mathbb{E}\left[\hat{f}_k\right] \tag{6}$$

$$g_k = \mathbb{E}\left[\hat{g}_k\right] \tag{7}$$

with covariance matrix $\mathbf{C}_b^*$ of

$$\mathbf{C}_b^* = \frac{2\sigma_b^2}{Q}\mathbf{I}, \tag{8}$$

where $\mathbf{I}$ denotes the identity matrix. This is Cramer-Rao lower bound to estimating Fourier components [2]. Note to lower the variance on each component either the signal to noise ratio (SNR) needs to be increased or more samples need to be included in the time series. The signal to noise ratio is very important as improvement can only be verified if two competing methods are compared on imagery with similar SNR. The second statement is also intuitive as a more accurate estimates of the features can be derived if there are more samples available. The covariance matrix $\mathbf{C}_b^*$ is only valid under ideal conditions, as with most orthogonality and fundamental harmonics conditions are seldom satisfied. A more accurate covariance matrix for an application could be expressed as

$$\mathbf{C}_b = \frac{2\sigma_b^2}{Q}\mathbf{I} + \epsilon \geq \mathbf{C}_b^*, \tag{9}$$

where ϵ denotes an increase in estimation uncertainty. The variable ϵ unfortunately is not constant for all time indices. There are many reasons for this to list: (1) different start of rainfall season causing significant increases in variations of time series representing natural vegetation, (2) seasonal drought, (3) land cover change, etc. The limitation is as we increase the length Q of the temporal sliding window we reduce the first term, but we might also include non-stationary events into the window which in effect increase ϵ.

This was important in our study area as certain parts of the province received irregular volume of annual rainfall which causes damage to the natural vegetation. This significantly varies annual spectral reflectance values which could negatively impact feature estimation if the window increases to longer than one annual seasonal cycle. There is also the problem that the distribution of features (given no noise) might be overlapping between different classes, which means you can never achieve perfect separation of land cover classes.

4. EXPERIMENTAL RESULTS

The data set consist of $29.5\,\text{km}^2$ changed and $808\,\text{km}^2$ no changed time series (1497 vegetation time series and 1735 settlement time series). Sequential sub-sequences were extracted from each time series to classify. Land cover change was declared when a natural vegetated time series permanently transition to the settlement class. A Multi-Layer Perceptron (MLP) neural network was used to classify the sub-sequences in a post-classification approach.

In the feature extraction step the magnitude of the fast Fourier transform on the sub-sequences was computed. Selecting only a few Fourier components for analysis is standard [4, 5], and for our experiments we only considered the mean and annual Fourier components. The MLP comprises an input layer, one hidden layer and an output layer. All hidden and output layers used a tangent sigmoid activation function in each node. The weights in the training phase of the MLP were determined using a steepest descent gradient optimization method, with gradients estimated using backpropagation. The results for all our experiments are shown in table 1.

We investigate three different spectral band combinations in our experiments [6]: (1) NDVI, (2) first two spectral bands and (3) all seven spectral bands. There was a general improvement in classification and change detection accuracy (increase true positive rate and decreased false positive rate) when using more combinations of spectral bands. This was because

Table 1: Classification and change detection accuracy using an MLP on the extracted sub-sequences. Classification entries presents the average accuracy in percentage along with the corresponding standard deviation. The change detection entries give the average true positive rate along with the false positive rate in parenthesis.

Spectral band	**Class**	**Sliding window length Q**		
		6 months	12 months	18 months
NDVI	Vegetation	69.7 ± 7.8	72.8 ± 5.3	73.9 ± 4.8
	Settlement	81.5 ± 5.0	83.2 ± 3.7	84.8 ± 3.1
	Change detection	$69.2(30.0)$	$70.2(29.5)$	$71.9(29.2)$
Bands 1–2	Vegetation	81.4 ± 4.3	83.1 ± 4.1	85.2 ± 3.7
	Settlement	86.3 ± 3.4	86.8 ± 2.7	88.1 ± 2.2
	Change detection	$77.6(22.4)$	$78.2(21.3)$	$78.7(20.7)$
Bands 1–7	Vegetation	93.1 ± 2.1	94.4 ± 1.6	94.5 ± 1.5
	Settlement	93.8 ± 1.6	95.2 ± 1.1	96.3 ± 1.0
	Change detection	$90.5(9.6)$	$90.8(9.4)$	$91.0(8.9)$

the MLP could derive a more complex hyper-plane to separate the land cover classes. We also observe an significant improvement when increasing the length of the temporal sliding window Q in all experiments. This supports our hypothesis that a longer sequence of measurements reduces the error variance in estimating the features which results in more accurate computation of their respective class label. We do not have the ability to reduce the SNR in our images, but we can easily decrease the SNR to verify the second part of the bound. By adding white Gaussian noise to our time series it was found that all experimental results deteriorated. Increasing the length of the temporal sliding window also increases the variable ϵ, which provides an upper bound to the classification/change detection accuracy.

5. CONCLUSION AND DISCUSSION

In this paper we show the covariance matrix derived from estimating features used to classify land cover classes. The level of uncertainty has a direct effect on the performance of the change detection capabilities when using a machine learning method. This creates a lower bound on the performance which can be achieved. In our case we computed the magnitude of the Fourier components as features, which under the assumption that all frequencies contained within temporal sliding window are multiple harmonics of the fundamental frequency and all are orthogonal to each other, allows us to show equation 8 derived in [2]. The error variance in our features reduces the ability of the machine learning method to separate different land cover classes. In addition it is also important to ensure proper meaningful analysis of time series features [3].

Our experimental results and lower bound shown in equation 9 indicates that increasing the length of the temporal sliding window improves the estimates of the features. The bound also states that improving the SNR also improves the estimation of the feature. The classification and change detection accuracy are improved by increasing the length of the temporal sliding window. Increasing the length of the sliding window also has the potential chance of increasing the error factor ϵ. This is because the features estimated between different years might vary and increasing the length of the temporal sliding window adversely affect the variation. This can be seen in the results shown in table 1 where the classification and change detection performance slowly plateau. This objective is thus to maximize the length of the temporal sliding window to improve change detection and classification accuracy.

6. REFERENCES

[1] P. Coppin et al., "Digital change detection methods in ecosystem monitoring: A review,," *International Journal of Remote Sensing*, vol. 24, pp. 1565–1596, 2004.

[2] S.M. Kay, *Fundamentals of Statistical Signal Processing*, Prentice Hall, Upper Saddle River, New Jersey, 07458, first edition, 1993.

[3] B.P. Salmon et al., "Unsupervised land cover change detection: Meaningful sequential time series analysis," *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing*, vol. 4, pp. 327–335, 2011.

[4] R. Juarez and W. Lui, "FFT analysis on NDVI annual cycle and climatic regionality in northeast brazil," *Int. J. Climatol.*, vol. 21, pp. 1803–1820, 2001.

[5] M. Jakubauskas, D. Legates, and J. Kastens, "Crop identification using harmonic analysis of the time-series AVHRR NDVI data," *Computers and electronics in agriculture*, vol. 37, pp. 127–139, 2002.

[6] B.P. Salmon et al., "The use of a multilayer perceptron for detecting new human settlements from a time series of MODIS images," *International Journal of Applied Earth Observation and Geoinformation*, vol. 13, pp. 873–883, 2011.