

# Monitoring Urban Growth with Spatial Filtering of Satellite Image Time Series

Çaglayan Tuna  
Université Bretagne Sud, IRISA  
Vannes, France  
caglayan.tuna@irisa.fr

François Merciol  
Université Bretagne Sud, IRISA  
Vannes, France  
francois.merciol@irisa.fr

Sébastien Lefèvre  
Université Bretagne Sud, IRISA  
Vannes, France  
sebastien.lefevre@irisa.fr

**Abstract**—Monitoring urban growth and change is an important task for urban planning and disaster management. While several change detection approaches have been proposed to deal with growing urban areas, their performances are usually limited due to outliers in Satellite Image Time Series (SITS). In this study, in order to discriminate urban growth from the other changes, we exploit spatial connectivity of the changed pixels. To do so, we first stack SITS to a single synthetic image whose pixel values denote the temporal variability along the series. Then, we propose to rely on efficient and well-established spatial filtering by means of the max-tree image representation, leading to a novel approach for detecting changes in urban areas, and more precisely focusing on the spatial extent of such changes in relationship with the urban growth. Experimental results obtained on Landsat imagery of Dar es Salaam showed that our approach helps to remove outliers from the change map and provides satisfactory accuracy.

**Index Terms**—Urban growth monitoring, change detection, Attribute filtering, Satellite Image Time Series, Max-Tree.

## I. INTRODUCTION

Analysis of satellite image time series (SITS) have shown growing interest due to the availability of high temporal frequency Earth Observation (EO) data provided by the new satellite missions (e.g. Sentinel) [1]. Change detection [2] receives most of attention since it allows to monitor urban areas, and such urban growth information is required for city planners, disaster management, infrastructure, *etc.*

SITS data do not always include homogeneous combination of time series information. There are several challenges to overcome, such as cloud, haze, shadows, *etc.* [3]. Thus, outliers that are due to missing temporal samples, atmospheric contamination, viewing angle, cloud, *etc.* lead to abnormal pixel values that have then to be carefully addressed by change detection methods. As such, purely pixel-based approaches may be inadequate to deal with such outliers. Therefore, spatial connectivity appears as an important information for urban change monitoring [4]. Among recent works that embed spatial information are the multi-level change detection using stereo images from [5], and the comparison of pixel-based and object-based approaches for urban monitoring in [6]. However, these works only compare two satellite images, and clouds are not specifically addressed (but rather either filtered with a prior outlier detection method or by considering Synthetic Aperture Radar (SAR) images instead of optical ones).

Recently, morphological hierarchies through tree-based representations have appeared as an efficient way to implement

(spatially) connected filters [7] over ordered sets such as grayscale images. However, to the best of our knowledge, extension of such hierarchies to time series has only been addressed in [8], where temporal connectivity is considered between two Synthetic Aperture Radar (SAR) Images to build a Binary Partition Tree. Another strategy consists in considering SITS as multivariate data, where each pixel is assigned a vector made of its values at different dates. It thus allows to access multivariate morphological approaches [9]. But conversely to color images for which vector orderings have been extensively studied, to the best of the author’s knowledge, there is no ordering of time series able to cope with its specific challenges (outliers, irregular sampling, *etc.*) yet.

Aiming at change detection, [10] performs a pixelwise comparison between two images, comparing each pair of pixels based on respective attribute profiles (AP). AP are computed from successive applications of connected filters, and changes are measured in each pixel coordinates by summing differences between all filtered images (or scales). Extension of these approaches to longer time series is not straightforward.

In this paper, we propose to monitor changes in urban areas by combining spatial and temporal information. Temporal information is provided by Landsat SITS, that is first analyzed to derive a synthetic image embedding the temporal variability for each pixel. Then, spatial connectivity is taken into account through connected filtering, efficiently implemented with hierarchical representations. For the sake of illustration and validation, we applied our method to a dataset covering the Dar es Salaam city during the last 30 years, and compared the results with some public reference data.

## II. METHODOLOGY

As already mentioned, change detection for urban monitoring based on spatial information has accomplished great success w.r.t. pixel based approaches with long time series, and morphological hierarchies have been proved to be a good detector for changes between two images. We build upon these ideas to design a spatially-aware change detector relying on tree-based representation of long time series. Our approach is based on two steps: first we conduct a temporal analysis to find the highly possible changed areas, before filtering the resulting image using spatial information.

### A. Temporal processing

While a SITS is made of a series of satellite images, it can also be seen as a single image where each pixel is associated to a time series containing its values in the images corresponding to the different dates. In order to deal with long time series, we propose here to reduce SITS to such a single image by using pixel time series information for each location. We thus project each pixelwise time series to a single scalar value that is further used to form a unique image. By doing so, we limit the memory requirements brought by long term SITS data and also access to the efficient implementation of connected filters through max-trees that are uniquely defined for grayscale images (while there is no unique choice for multivariate images, see [11]).

Let  $I$  be a gray scale image,  $[x, y]$  the pixel coordinates, and  $I_t[x, y]$  the pixel value in the  $t^{\text{th}}$  frame of the SITS data made of  $T$  dates. In the sequel, we consider various functions to project pixel time series into scalar values, written as

$$I_{new}[x, y] = f(I_1[x, y], \dots, I_t[x, y], \dots, I_T[x, y]), \quad (1)$$

where  $I_{new}$  represent the created synthetic image.

Since we aim to find changed areas, we measure in  $I_{new}$  the spread of the time series data. Statistical dispersion functions such as range, inter-quartile range and standard deviation are well-known to calculate such variability or spread [12]. Indeed, if spreading of the data is high, possibility of the change would be high as well. Higher pixel value of the new synthetic image commits higher change for each location.

A simple spreading measure is range, i.e. difference between the maximum and minimum values in each location:

$$I_{range}[x, y] = \max_{t \in [1, T]} (I_t[x, y]) - \min_{t \in [1, T]} (I_t[x, y]) \quad (2)$$

A second measurement is the inter-quartile range (IQR) which is the difference between the  $75_{th}$  (third quartile,  $Q_3$ ) and  $25_{th}$  (first quartile,  $Q_1$ ) percentile of the data:

$$I_{IQR}[x, y] = Q_3[x, y] - Q_1[x, y]. \quad (3)$$

Quartile coefficient of dispersion is also computed by using first and third quartiles:

$$I_{Q_{coef}}[x, y] = \frac{Q_3[x, y] - Q_1[x, y]}{Q_3[x, y] + Q_1[x, y]}. \quad (4)$$

Another measure of spread is standard deviation. It is calculated in terms of how far the data differs from the average:

$$I_{std.dev}[x, y] = \sqrt{\frac{1}{T} \sum_{t=1}^T (I_t[x, y] - \bar{I}[x, y])^2}, \quad (5)$$

with

$$\bar{I}[x, y] = \frac{1}{T} \sum_{t=1}^T I_t[x, y] \quad (6)$$

Although other statistical dispersion measurements have been introduced (e.g. entropy [13]), we limit ourselves to these basic functions for our preliminary experiments. Dispersion measures were computed for every pixel location separately, and gathered into a new synthetic image as already explained.

### B. Spatial processing

We have shown how we stack pixel time series into scalar values in order to obtain a single image that gathers pixelwise temporal variability. We now analyze this image by taking into account spatial information. To do so, we rely on attribute filters that are a class of connected filters (i.e. filters operating on connected components, or connected sets of neighboring pixels sharing the same value) whose filtering criterion relates to a predefined attribute [14].

Various attributes have been used so far, such as size, shape and contrast. We use here the popular area attribute (measured as the number of pixels) since we aim to distinguish between large connected components denoting the urban extent and smaller ones that can be seen as outliers or noise.

As already stated, attribute filters can be efficiently implemented through the representation of the input image as a morphological hierarchy or a tree structure such as the max-tree. More precisely, since  $I$  is a gray scale image, it admits an ordering relation  $\leq$ . So we can define, for each threshold  $L \in \mathbb{Z}$  or  $L \in \mathbb{R}$  (for images with integer and real-valued pixels, respectively), the upper threshold set

$$[I \geq L] = \{[x, y], I[x, y] \geq L\}. \quad (7)$$

This set is made of several connected components  $C_h^L$  that are further nested into the so-called max-tree hierarchy, whose definition for integer-valued images is:

$$\forall L \in [0, 255[, \forall h, \exists j \text{ s.t. } C_h^L \subseteq C_j^{L+1} \quad (8)$$

We prefer here max-tree over other alternatives since we focus on higher values corresponding to high possible changes (see [7] for a recent survey). Such a hierarchical representation allows for very efficient attribute filtering, that only consists in computing one (or several) attribute  $A$  for each node, and filtering out the nodes that do not fulfill a given criterion (i.e. threshold)  $\lambda$ . Since we use here an increasing attribute (area), the filtering is straightforward: if  $A(C_h^L) < \lambda$ ,  $C_h^L$  is removed with all its descendants  $C_j^{L'} \subseteq C_h^L$ . The final (filtered) image is then reconstructed from the pruned tree.

Area filtering was applied with several thresholds, that were here defined empirically. We assume that urban growth areas are connected sets of changed pixels whose area is larger than those of other parts of the satellite images such as clouds. If the spatial extent of cloudy areas is lower than the threshold, they will be filtered out, as illustrated in Sec. III-B.

## III. EXPERIMENTS

### A. Data

We consider here the Dar es Salaam region in Tanzania, that is known for careful assessment of its urban cover expansion (from 12,988 to 20,607 hectares, about 47% increase [15]). We illustrate such expansion through reference maps given in Fig. 1 and taken from [16]. In these maps dating from 1992, 1998 and 2002 respectively, urban areas are identified in white areas. Furthermore, we also include a 2017 reference taken from OpenStreetMap [17].

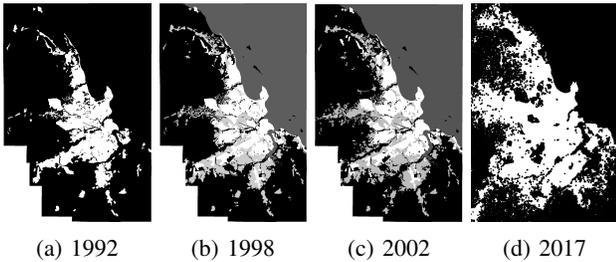


Fig. 1: Reference data for Dar es Salaam, Tanzania.

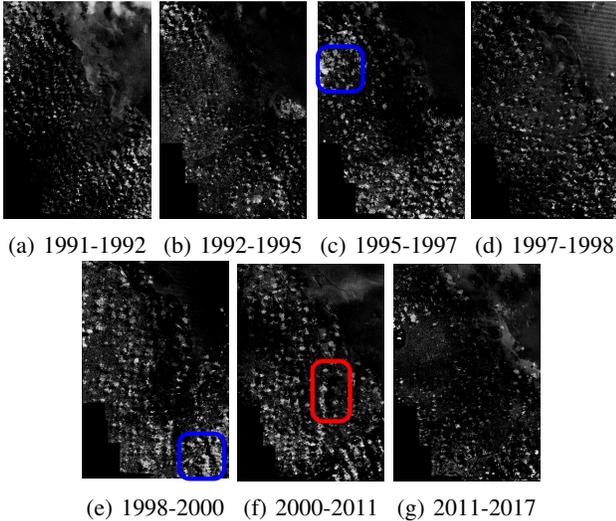


Fig. 2: Difference between the Landsat images.

The SITS is made of Landsat satellite imagery, with a 30m spatial resolution provided by the United States Geological Survey (USGS). We use in here 8 different multispectral images, for which we considered the blue band (first band) with the shortest wavelength ( $0.45\text{-}0.52\mu\text{m}$ ) and 30m spatial resolution. This choice is motivated by the higher dynamic range and its proven ability to detect clouds [18]. Let us note that using the spectral information would most probably lead to better results.

For the sake of evaluation, we clipped the Landsat data according to the reference maps shown in Fig. 1, leading to an image size of  $1595 \times 1076$  pixels. We highlight changes between successive images in Fig. 2. Quantitative evaluation is performed by comparing the obtained results with the difference map between 1992 and 2017 (see Fig. 4a).

## B. Results

Conversely to existing works that detect changes between only two dates, our approach accommodates with long time series data. Long time series data is useful for some applications such as reverted changes, vegetation dynamic [19], *etc.* As discussed in Sec. II, we stacked the SITS in a synthetic image that measures temporal spread. To do so, we consider in this paper the following functions: Range, IQR, Quartile Coefficient, and Standard Deviation. As shown in Fig. 3, these

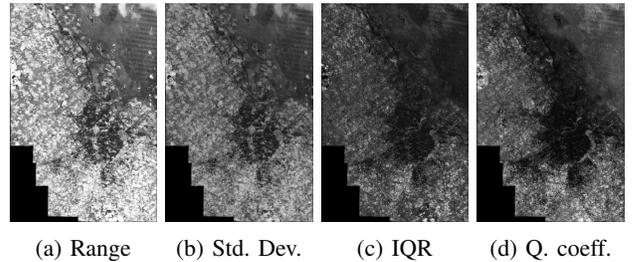


Fig. 3: Synthetic images computed from the SITS.

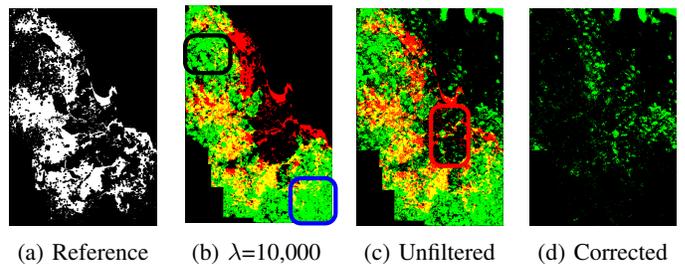


Fig. 4: Urban growth estimation: reference data, result with and without spatial filtering, and difference in between.

functions darken the stationary city area comparing to changed areas that are given higher values. Nevertheless, we can see that clouds still bring some artifacts in these images. We then apply an area filter over each of these images, with an efficient implementation based on a max-tree structure. We have considered different thresholds for the area filtering and report here the observed results.

We provide in Fig. 4 a visual comparison between the reference change map (a) made of difference between reference maps from 1992 and 2017, and the results provided by our method (b) considering range as the spread function. The color codes are as follows: red, green and yellow pixels represent false negative (i.e. wrongly classified non-urban areas), false positive (i.e. wrongly classified urban areas) and true positive (i.e. pixels correctly identified as new urban areas), respectively. The effect of spatial filtering is further assessed by comparing with the temporal stacking only (c), with corrected errors given in (d). As it can be observed, the dataset is challenging due to the presence of clouds and changes over a 25 years period. We can see the effect of using an area filtering, that allows to remove areas smaller than a given threshold such as clouds. One of them is highlighted with red box in Fig. 4b and corresponds to a cloudy area (see Fig. 2f). As emphasized by differences between Fig. 4c and 4b, spatially filtering of time series information allows to reduce artifacts due to noise or clouds. However, there are still wrongly classified parts caused by clouds connected to urban changes (see blue boxes in Fig. 4b, 2c and 2e). Such an issue might be addressed with advanced connectivity [20].

Finally, we report quantitative evaluation in Figs. 5a and 5b. We first show in Fig. 5a the stability of the synthetic image w.r.t. the chosen threshold value. We can observe that after

a severe decrease, increasing further the area has a limited filtering effect. Fig. 5b provides a comparison analysis in terms of true positive rate w.r.t. the threshold value, that confirms the stable behavior. We can observe the relevance of the spatial filtering, with significant improvement (+5-10%) over the temporal (pixel) analysis only, and better results being obtained with higher area thresholds. Also, we can notice that the presence of clouds does not affects too strongly the accuracy with only a few points lost w.r.t. a cloud-free dataset (see discussion below). The whole process takes 24 seconds with a MacBook Pro 2.6 GHz Intel Core i5 processor and Python implementation. It can be further reduced with efficient implementations of tree-based processing [21].

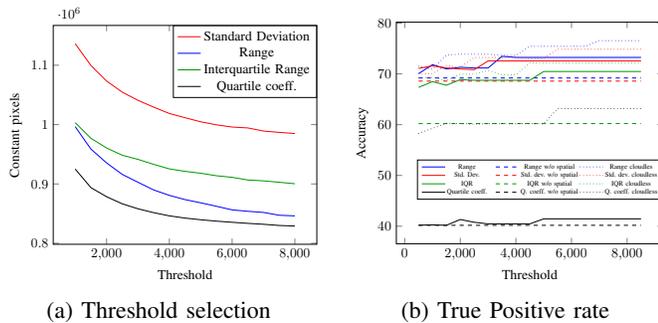


Fig. 5: Results with different dispersion functions, compared with cloudless data and without the spatial filtering.

For the sake of comparison, we compare the different synthetic images with and without the spatial filtering. Furthermore, we also compare with two variants inspired from [10], for which we set the thresholds similar to our settings. The first one computes a change map as the difference between respective AP from the first and last dates. While it can reach a high accuracy (89%) when applied to cloudless images (1992 and 2017), it remains ineffective when applied to cloudy data (30% if 2017 is compared with the cloudy 1991 image). The second variant consists in calculating the sum of AP differences of successive SITS frames. Again, we assess it on both the standard dataset made of all dates (i.e. containing clouds) and on a subset ignoring cloudy images (1991, 1997, 2000). It reaches an accuracy of 50% and 75% respectively, showing its higher sensitivity to clouds than the proposed approach.

#### IV. CONCLUSION

In this paper, we have proposed a novel method for monitoring urban growth areas, based on a spatial analysis of changed detected within a satellite image time series. We assumed that urban growth areas could be characterized by connected sets of changed pixels with a spatially significant extent. We thus proposed to stack all images of a time series into a single image where changes are highlighted, before performing spatial area filtering over this image. Area filtering allows us to remove insignificant changes and outliers from the image, and it can be efficiently implemented using hierarchical

representations such as max-tree. Preliminary experimental assessment provides promising results but also calls for further work to define optimal area threshold, for which automatic methods have been introduced recently [22]. Furthermore, we did not take into account spectral information yet, while it could lead to better results.

#### ACKNOWLEDGMENTS

This work was supported by Centre National d'Études Spatiales (CNES) and Collecte Localisation Satellites (CLS).

#### REFERENCES

- [1] C. Gómez, J. C White, and M. A Wulder. Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 116:55–72, 2016.
- [2] M. Hussain, D. Chen, Angela Cheng, H. Wei, and D. Stanley. Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, 80:91–106, 2013.
- [3] P. J. Brockwell, R. A. Davis, and M. V Calder. *Introduction to time series and forecasting*, volume 2. Springer, 2002.
- [4] Jixian Z. Zhiyong C. Xiaogang, N. Object-based city land cover classification and change analysis with multi-temporal high resolution remote sensing images in jiangyin. In *JURSE*, pages 107–110, 2013.
- [5] D. Wen, H. Liu, J. Li, and Xi. Huang. Multi-level change detection in urban areas using zy-3 multi-temporal stereo imagery. In *JURSE*, pages 1–4, 2017.
- [6] Ban Y. Yousif, O. Object-based urban change detection using high resolution sar images. In *JURSE*, pages 1–4, 2015.
- [7] P. Bosilj, E. Kijak, and S. Lefèvre. Partition and inclusion hierarchies of images: A comprehensive survey. *Journal of Imaging*, 4(2):33, 2018.
- [8] A. Alonso-González, C. López-Martínez, and P. Salembier. Polar time series processing with binary partition trees. *IEEE Transactions on Geoscience and Remote Sensing*, 52(6):3553–3567, 2014.
- [9] E. Aptoula and S. Lefèvre. A comparative study on multivariate mathematical morphology. *Patt. Recogn.*, 40(11):2914–2929, 2007.
- [10] N. Falco, M. Dalla Mura, F. Bovolo, J. A. Benediktsson, and L. Bruzzone. Change detection in vhr images based on morphological attribute profiles. *IEEE Geosc. and Rem. Sens. Lett.*, 10(3):636–640, 2013.
- [11] E. Aptoula, M. Dalla Mura, and S. Lefèvre. Vector attribute profiles for hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 54(6):3208–3220, 2016.
- [12] P. Sprent and N. C. Smeeton. *Applied nonparametric statistical methods*. Chapman and Hall/CRC, 2000.
- [13] M. Rostaghi and H. Azami. Dispersion entropy: A measure for time-series analysis. *IEEE Signal Processing Letters*, 23(5):610–614, 2016.
- [14] P. J. Salembier C. and M. Wilkinson. Connected operators: A review of region-based morphological image processing techniques. *IEEE Signal Processing Magazine*, 26(6):136–157, 2009.
- [15] C. C. Mkalawa. Analyzing dar es salaam urban change and its spatial pattern. *Int. J. of Urban Planning and Transp.*, 31(1):1138–50, 2016.
- [16] R Sliuzas, A Hill, C Lindner, and S Greiving. Dar es salaam land use and informal settlement data set. *NASA SEDAC*, 2016.
- [17] M. Haklay and P. Weber. Openstreetmap: User-generated street maps. *IEEE Pervas Comput*, 7(4):12–18, 2008.
- [18] O. Hagolle, M. Huc, D. Pascual, and G. Dedieu. A multi-temporal method for cloud detection, applied to formosat-2, venμs, landsat and sentinel-2 images. *Remote Sensing of Environment*, 114(8):1747–1755, 2010.
- [19] T. Guyet and H. Nicolas. Long term analysis of time series of satellite images. *Pattern Recognition Letters*, 70:17–23, 2016.
- [20] B. Perret, S. Lefèvre, C. Collet, and E. Slezak. Hyperconnections and hierarchical representations for grayscale and multiband image processing. *IEEE Transactions on Image Processing*, 21(1):14–27, 2012.
- [21] F. Merciol and S. Balem, T. and Lefèvre. Efficient and large-scale land cover classification using multiscale image analysis. In *ESA Conference on Big Data from Space (BiDS)*, 2017.
- [22] G. Cavallaro, N. Falco, M. Dalla Mura, L. Bruzzone, and J. A. Benediktsson. Automatic threshold selection for profiles of attribute filters based on granulometric characteristic functions. In *ISMM*, pages 169–181, 2015.