A Non-Parametric Approach to Detect Changes in Aerial Images

Marco Túlio Alves Rodrigues, Daniel Balbino, Erickson R. Nascimento, William Robson Schwartz

Department of Computer Science, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, Pampulha, Belo Horizonte, Brazil
{tulio.rodrigues,balbino,erickson,william}@dcc.ufmg.br

Abstract. Detecting changes in aerial images acquired from a scene at different times, possibly with different cameras and at different view points, is a crucial step for many image processing and computer vision applications, such as remote sensing, visual surveillance and civil infrastructure. In this paper, we propose a novel approach to automatically detect changes based on local descriptors and a non-parametric image block modeling. Differently from most approaches, which are pixel-based, our approach combines contextual information and kernel density estimation to model the image regions to identify changes. The experimental results show the effectiveness of the proposed approach compared to other methods in the literature, demonstrating the robustness of our algorithm. The results also demonstrate that the approach can be employed to generate a summary containing mostly frames presenting significant changes.

Keywords: Change detection, non-parametric modeling, aerial images.

1 Introduction

The detection of changes in a scene plays a central rule in a myriad of applications, such as disaster management, urban growth, security, burned areas and surveillance to name a few. In addition, detecting structural changes is useful to gather information from the environment, which might present economic impact.

The most common procedure applied to detect changes is to use human operators to watch videos from a monitoring camera and identify changes in images from the scene. However, in general, video monitoring performed by humans is error prone due to lack of attention in repetitive long tasks. Therefore, automated monitoring is a solution to reduce human error and can be used as a filter to locate video segments that should be further analyzed by operators.

The basic task of change detection in images is to locate the pixels from a *reference image* that are different from other images [9], referred to as *test images*. Significant changes may include object removal, movement of objects and shape variation of the scene structures. Changes in image pixels can also be generated by viewpoint change, noise, illumination changes, nonuniform attenuation, atmospheric absorption, swaying trees, rippling water or flickering monitors. Although such effects produce changes in the pixel intensity, they must be ignored by the change detection method, which makes it a hard problem to handle.

In general, two main goals are pursued by change detection systems: (i) location of changes through a mask, referred to as *change mask* and (ii) identification of frames where changes occur without the creation of change masks. While in former, the exact location of the changes is of interest, in the latter, the user is interested in finding the frames that the changes happened, in which a further analysis will be performed. This work focuses on the latter goal to being able to create a shortlist containing frames with high likelihood of presenting changes.

Several approaches focusing on background subtraction and remote sensing techniques have been applied to change detection. In general, the background subtraction performance is highly dependent on building and maintaining a background model. Most of these techniques are pixel-based and they assume independence among pixels. In remote sensing methods, the change detection performance depends on the trade-off between the spatial and spectral resolution [1].

In this paper, we propose a novel approach to identify, from a set of images, which ones present significant differences when compared to a reference image. Our method extracts local feature descriptors from image blocks and estimates the likelihood of a change by using a non-parametric modeling based on Kernel Density Estimation (KDE) [5]. KDE is a statistically-sound method that estimates a continuous distribution from a finite set of points. Unlike background subtraction and remote sensing methods, our technique does not require a complex learning stage (i.e., it just stores samples). Additionally, our method requires a few number of samples (single example from the reference image), being therefore, capable of detecting changes by using only two images (the reference image and a test images), which is hard to perform with parametric approaches due to the lack of samples to estimate parameters.

In the experiments, we compare our method with techniques widely used for change detection. According to the results, the proposed approach outperforms several other methods found in the literature [2, 3, 10, 11, 18], mainly due to the fact that our approach is more robust to illumination changes, frequent on aerial images taken at different times. The results also demonstrate that our method is able to filter the video segments generating a video summary containing mostly frames presenting significant changes.

2 Related Work

Over the past years, a large number of change detection techniques have been proposed, mostly based on background subtraction and remote sensing techniques [9]. However, there are still several limitations in change detection techniques since it is hard to separate significant (e.g., object removal, structural changes) from insignificant changes (e.g., noise, illumination changes).

Background subtraction methods consist in learning the background model by using several reference images. In the test phase, all pixels are classified as foreground or background. The foreground pixels indicate changes [8]. Such methods can divided into parametric and non-parametric.

Parametric methods assume that each pixel can be modeled as a random process that can be approximated by some parametric distribution [12]. Regarding non-parametric approaches, the background is modeled by a probability density function (PDF) estimated for each pixel. The main characteristic of these methods is the strong dependence on the data. Although such approaches are capable of adapting to sudden changes, they require to store the pixels. Elgammal et al. [4] presented a non-parametric kernel density estimation to perform the background subtraction. St-Charles et al. [11] proposed an approach based on the adaptation and integration of Local Binary Similarity Pattern (LBSP) features in a non-parametric background model that is then automatically tuned using pixel feedback. The major improvement is related to internal threshold that makes the binary descriptors much more robust to illumination variation.

Similar to background subtraction methods, remote sensing techniques for change detection also requires a learning stage to model the reference image [15]. However, unlike the background subtraction methods, the remote sensing approaches are based on feature extraction, which reduces errors due to pixels noise and small changes in the reference image.

Celik [2] computes the difference between the test image and the reference image by combining PCA and K-means. The change is detected by partitioning the feature vector space into two clusters. The algorithm assigns each pixel to one of the two clusters based on the Euclidean distance between the pixel feature vector and mean feature vector of clusters. To reduce the effect caused by noise in [2], Cheng et al. [3] proposed the use of the fraction Fourier transform (FRFT). Zheng et al. [17] also employ a technique based on a clustering. They apply image subtraction the log ratio operator to generate two types of change maps. Then, a simple combination uses the maps obtained by the mean filter and the median filter to improve the final change map.

Rodrigues et al. [10] investigate the sensibility of pixel to noise and the influence of monotonic transformation in change detection methods. They proposed a solution that is neither based on background subtraction nor on remote sensing techniques, but a combination of super-pixel extraction, hierarchical clustering and segment matching. The drawback of their approach is its sensibility to variations in lighting and camera displacement.

In spite of the significant progress in solving the change detection problem, the aforementioned techniques are highly limited by the large variability of irrelevant changes. Virtually, all described approaches require a learning stage and present high computational cost. Moreover, these techniques do not work properly whenever the background scene suddenly changes or there are not enough samples to estimate the background model. Therefore, they are unfeasible when it is provided only a single reference image.

3 Proposed Approach

Let the *reference image* and the *test image* be two registered aerial images acquired from the same geographical area at times t_1 and t_2, respectively, we detect changes in the scene by analyzing the image characteristics and output a score indicating whether there are changes between the two images. As illustrated in

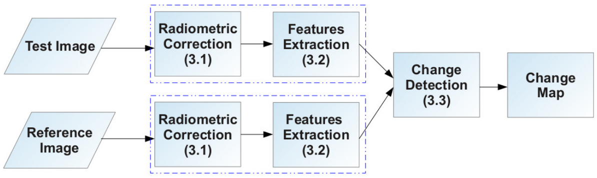


Fig. 1: Steps of the proposed change detection method. When a reference and a test image are presented to the system, we first perform the radiometric normalization and extract the features to estimate the likelihood of changes.

Figure 1, the proposed approach in this work is composed of the three main steps: (i) *radiometric correction*, (ii) *feature extraction*, and (iii) *change detection*. These steps are described in details in the following sections.

3.1 Radiometric Correction

The goal of radiometric correction is to remove or compensate for illumination changes. We employ the Self-Quotiente Image (SQI) [14], a popular method used for synthesizing an illumination normalized image from a single image. The SQI normalization is defined by a smoothed image $S(x, y)$ of a image $I(x, y)$ as

$$Q(x, y) = \frac{I(x, y)}{S(x, y)} = \frac{I(x, y)}{F(x, y) * I(x, y)}, \tag{1}$$

where $F(x, y)$ is a low-pass filter and $*$ is the convolution operator.

3.2 Feature Extraction

We use in our method two different local feature descriptors to provide high robustness to lighting changes: the Local Binary Patterns (LBP) [7] and the Local Ternary Patterns (LTP) [13]. These features are invariant to monotonic changes and are extremely fast to compute.

The LBP descriptor for a pixel $C = (x_c, y_c)$ is computed by thresholding the gray value of N sampling pixels defined by the indicator function $s(x_1, x_2)$. The indicator function returns 1 when the intensity value of pixel x_1 is greater than x_2 and 0 otherwise. By considering g_c the intensity of the center pixel and g_p ($p = 0, \ldots, N-1$) the corresponding intensity of a pixel value of N sampling points, the final feature vector is given by summing the thresholded values weighted by powers of two.

The major drawback of LBP descriptor is its sensibility to noise, since the operator thresholds at exactly the central pixel g_c. To overcome this limitation, Tan and Triggs [13] proposed to relax the intensity of the central pixel by using a slack of width equals to $\pm t$. The feature is computed as

$$LTP(g_p, g_c, t) = \begin{cases} 1, & g_p \geq g_c + t \\ 0, & |g_p - g_c| < t \\ -1, & g_p \geq g_c - t \end{cases} . \tag{2}$$

A coding scheme is used to split the ternary pattern in negative and positive LBP. Here, t is a parameter that makes the LBP more resistant to noise. In this work we use $t = 5$, defined experimentally.

After computing the LBP or LTP (positive and negative) codes for each pixel, the vector descriptor is represented by a 256-bin normalized histogram (in our tests we use $N = 8$ and $t = 5$).

3.3 Change Detection

We use a non-parametric approach to estimate the probability density function, since only one sample is available (the reference image) to learn the background model. Given a set reference image blocks $B = \{b_1, b_2, \ldots, b_N\}$, the density estimate at a new test image block b_t is given by

$$\hat{f}(b_t) = \frac{1}{N} \sum_{i=1}^{N} K_h(b_t, b_i) = \frac{1}{Nh} \sum_{i=1}^{N} K(\frac{b_t - b_i}{h}), \tag{3}$$

where h is the bandwidth, N is the number of samples in the reference image, b_i is a block in reference image and b_t is a block in test image. We used the Gaussian kernel for density estimation. Then, for each block histogram, the density estimate is found.

To evaluate the kernel summations in linear time, we use FigTree [6], which provides an efficient computation of probabilities by KDE by combining the Improved Fast Gaussian Transform (IFGT) [16] with a kd-tree on cluster centers for neighbor searching in multiple dimensions. This method accelerates the Gaussian kernel summation and reduces the computational complexity of the evaluation of the sum of N Gaussians at M points in d dimensions from $O(MN)$ to $O(M+N)$.

4 Experimental Results

To evaluate the effectiveness of the proposed approach, we conduct several experiments considering synthetic and non-synthetic datasets. First, we evaluate the feature extraction using a synthetic dataset (Section 4.1). We then evaluate different setups for the proposed method (Section 4.2). Finally, in Section 4.3, we compare the results with other methods regarding the accuracy and the ability of generating a summary with relevant frames, i.e., frames presenting changes.

We evaluate our approach using the area under the curve (AUC) obtained from the operating characteristic (ROC) curve computed based on the true positive rate and the false positive rate, in which a test image is considered a false positive when the amount of changes is larger than a threshold (used to generate the ROC curve). Due to the lack of space, only the AUC values are shown.

To evaluate parameters of our method, we consider a *synthetic* dataset, in which from a single aerial image, we generated a set of 60 new images by applying several transformations simulating common effects in a capture system: Gaussian noise (with $\sigma = 0.01$), small translations (up to 5 pixels), blur, contrast and brightness changes. Moreover, on 30 of those images, we manually inserted artificial changes. The main purpose of this test set is to validate our algorithm on a controlled environment and also for tunning some of the main parameters. It is important to note that this dataset considers only one image as reference and the remaining as test images.

To compare our approach to other methods in the literature, we use a *non-synthetic* dataset composed of 26 aerial images acquired from PETROBRAS over oil pipes in the southern part of Brazil. Each with their own registered reference image, where in 13 of those have some change and on the other 13 do not contain changes.

4.1 Feature Extraction

To detect changes, our method divides the image into m block regions, from which local feature descriptors are extracted. Moreover, we add the coordinates of the blocks to the descriptors. This includes spatial information to the feature histograms. Six different block size were considered (8×8, 16×16, 32×32, 64×64, 92×92 and 128×128). According to the results obtained in the *synthetic* dataset, the block size 64×64 achieved the best results. Therefore, this block size will be used in the comparisons using non-synthetic dataset.

Three different descriptors were considered: LBP, LTP negative and LTP positive. We use LTP negative as the feature descriptor for the evaluation since it achieved the best results in the *synthetic* dataset (AUC=0.96), compared to the LTP positive (AUC=0.95) and LBP (AUC=0.90).

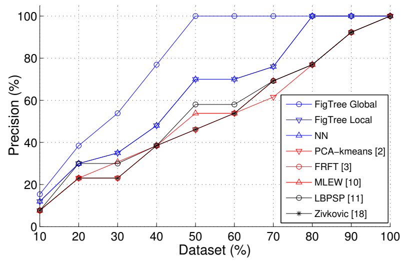
4.2 Evaluation of the Proposed Approach

In this section, we evaluate three setups for the proposed method. The first employs a simple nearest neighbor (*NN*) search, which is perhaps the most intuitive approach to change detection; the second employs a local KDE (*FigTree Local*), in which each image block is modeled by a single KDE; and the third setup employs a global KDE (*FigTree Global*) considering a single KDE model for the entire image but adding the block coordinates to the feature vector. The last two setups use the FigTree to optimize the search.

The last three rows of Table in Figure 2(a) depict the results achieved by each setup. According to the results, all three approaches achieved good results compared to the remaining methods. The small difference between the *NN* and *FigTree Local* is due the normalization performed by the KDE. They have nearly the same results, however, the complexity computational of *NN* is greater than the *FigTree Local* as discussed in Section 1. This makes the computation for large scale problems prohibitively expensive for *NN*. Finally, when a single KDE model is considered (*FigTree Global*), the results improved significantly, achieving AUC=0.88, compared with AUC=0.72 using *FigTree Local*. The global approach allows the contribution of nearby blocks, different from the *FigTree Local* and *NN*, which consider each block independently. The contribution of nearby blocks is essential to suppress acquisition noise and small errors in image registration, achieving therefore, more accurate results.

Approach	AUC values
Zivkovic [18]	0.50
PCA-kmeans [2]	0.50
FRFT [3]	0.52
MLEW [10]	0.52
LBPSP [11]	0.55
NN	0.70
FigTree Local	0.72
FigTree Global	**0.88**

(a)



(b)

Fig. 2: Results comparing the proposed approach to other methods. (a) AUC achieved considering the *non-synthetic* dataset (the best results are closer to one); (b) precision as a function of the number of frames selected to show that our approach is able to create a summary containing relevant frames (i.e., frames presenting changes).

We also evaluated the multivariate Gaussian distribution to model the blocks. However, due to the nature of our problem, in which only a reference is available, the employment of a parametric model is not appropriate since the lack of samples does not allow a proper parameter estimation.

4.3 Comparisons

In this section, we compare the proposed approach using the *non-synthetic* dataset with five change detection techniques: Zivkovic [18], LBPSP [11], MLEW [10], PCA-Kmeans [2] and FRFT [3]. According to the results shown in Table of Figure 2(a), the background subtraction approaches (e.g, *Zivkovic* and *LBPSP*) do not work properly due to the lack of samples to estimate the background model. Therefore, they are unfeasible when it is provided only a single reference image. In addition, *PCA-Kmeans*, *FRFT* and *MLEW* achieved poor results due to their lack of robustness to the strong influence of illumination variation. On the other hand, the proposed approach (*FigTree Global*) achieved very accurate results due to its robustness to illumination changes and noises.

The achieved results demonstrate the importance of developing a method robust to illumination changes by performing radiometric correction and using robust feature descriptors. Otherwise, changes due to illumination, common in aerial taken at different times, might be considered as relevant changes, which should not be the case.

Figure 2(b) shows the accuracy measured by AUC as a function of the percentage of frames selected. According to the results achieved by the *FigTree Global*, it is possible to capture 58% of the changes, showing only for 30% of the frames (by displaying 40%, we were able to show 78% of the changes), which are the best results compared to the other approaches. Therefore, the employment of an automatic system can be used as a filter to provide a shortlist with frames that should be further analyzed. This would help the operators in the decision making process regarding actions to be executed.

5 Conclusions

In this paper, we described a combination of radiometric correction and a non-parametric strategy to estimate a probability density function by kernel approaches. The proposed approach uses radiometric correction, features description and Figtree. When compared to baseline techniques, the proposed approach achieved the highest AUC values, demonstrating to be a promising technique to be employed in change detection tasks when a single reference image is provided.

Acknowledgments

The authors would like to thank CNPq, CAPES, FAPEMIG and Petrobras.

References

1. Aplin, P., Atkinson, P.M., Curran, P.J.: Fine spatial resolution simulated satellite sensor imagery for land cover mapping in the United Kingdom. RSE (1999)
2. Celik, T.: Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and k-Means Clustering. IGRSL (2009)
3. Cheng, Y.Q., Li, H.C., Celik, T., Zhang, F.: FRFT-based improved algorithm of unsupervised change detection in SAR images via PCA and K-means clustering. In: IGARSS (2013)
4. Elgammal, A., Harwood, D., Davis, L.: Non-parametric model for background subtraction. In: ECCV (2000)
5. Morariu, V.I., Srinivasan, B.V., Raykar, V.C., Duraiswami, R., Davis, L.: Automatic online tuning for fast Gaussian summation. In: ANIPRS (2009)
6. Morariu, V.I., Srinivasan, B.V., Raykar, V.C., Duraiswami, R., Davis, L.S.: Automatic online tuning for fast Gaussian summation. In: NIPS (2008)
7. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. PAMI (2002)
8. Piccardi, M.: Background subtraction techniques: a review. In: ICSMC (2004)
9. Radke, R., Andra, S., Al-Kofahi, O., Roysam, B.: Image change detection algorithms: a systematic survey. IEEE TIP (2005)
10. Rodrigues, M., Milen, L.O., Nascimento, E.R., Schwartz, W.R.: Change detection based on features invariant to monotonic transforms and spatial constrained matching. In: ICASSP (2014)
11. St-Charles, P.L., Bilodeau, G.A., Bergevin, R.: Subsense: A universal change detection method with local adaptive sensitivity (2014)
12. Stauffer, C., Grimson, W.E.L.: Learning patterns of activity using real-time tracking. PAMI (2000)
13. Tan, X., Triggs, B.: Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions. In: PICAMFG (2007)
14. Wang, H., Li, S.Z., Wang, Y.: Face recognition under varying lighting conditions using self quotient image. In: ICAFGR (2004)
15. Xu, Q., Karam, L.J.: Change detection on SAR images by a parametric estimation of the KL-divergence between Gaussian Mixture Models. In: ICASSP (2013)
16. Yang, C., Duraiswami, R., Davis, L.S.: Efficient kernel machines using the improved fast Gauss transform. In: Advances in neural information processing systems (2004)
17. Zheng, Y., Zhang, X., Hou, B., Liu, G.: Using Combined Difference Image and-Means Clustering for SAR Image Change Detection. GRSS (2014)
18. Zivkovic, Z.: Improved adaptive Gaussian mixture model for background subtraction. In: ICPR. IEEE (2004)