

Evaluation of clustering algorithms for unsupervised change detection in VHR remote sensing imagery

Tobias Leichtle^{1,2}, Christian Geiß², Michael Wurm², Tobia Lakes³, Hannes Taubenböck²

¹ Company for Remote Sensing and Environmental Research (SLU), 81243 Munich, Germany, tobias.leichtle@slu-web.de

² German Remote Sensing Data Center (DFD), German Aerospace Center (DLR), 82234 Wessling, Germany

³ Geography Department, Humboldt-Universität zu Berlin, 12489 Berlin, Germany

Abstract—Remote sensing has proven to be an adequate tool for observation of changes to the Earth’s surface. Especially modern space-borne sensors with very-high spatial resolution offer new capabilities for monitoring of dynamic urban environments. In this context, clustering is a well suited technique for unsupervised and thus highly automatic detection of changes. In this study, seven partitioning clustering algorithms from different methodological categories are evaluated regarding their suitability for unsupervised change detection. In addition, object-based feature sets of different characteristics are included in the analysis assessing their discriminative power for classification of changed against unchanged buildings. In general, the most important property of favorable algorithms is that they do not require additional arbitrary input parameters except the number of clusters. Best results were achieved based on the clustering algorithms k-means, partitioning around medoids, genetic k-means and self-organizing map clustering with accuracies in terms of κ statistics of 0.8 to 0.9 and beyond.

Keywords—change detection; clustering; object-based image analysis; very-high resolution (VHR) remote sensing;

I. INTRODUCTION

Monitoring of changes is one of the intrinsic capabilities of remote sensing. With respect to increasingly available remote sensing images with very-high spatial resolution (VHR), highly automatic techniques for detection of changes become more and more important [1]. Especially in urban environments, where different surface materials and image objects are concentrated in a spatially complex and highly dynamic manner, robust and simultaneously flexible change detection is a challenging task. Particularly the high level of detail in VHR images within urban environments favors object-based methods over rather traditional pixel-based methods for automatic detection of changes [2]. Regarding rapidly changing urban areas in emerging and developing countries of Africa and Asia, monitoring of urban growth in terms of expanding building inventory is a common task [3]. For these reasons, an unsupervised object-based change detection approach with focus on the high detail of individual buildings is intended in this study.

In this context, a promising and recently emerging methodology for unsupervised change detection is partitioning clustering [4]. For example, [5] and [6] apply k-means

clustering to difference images of optical and SAR imagery, respectively. Kernel k-means is used for change detection in QuickBird images by [7], while the same group of authors extended their approach and applied it to SPOT and Landsat data in [8]. Ghosh et al. [9] employ two fuzzy clustering algorithms for binary change detection in Landsat difference images. A technique from a different clustering domain - sparse hierarchical clustering - is used by [10] for change detection in VHR imagery. It has to be noted that all of the mentioned studies either work directly on pixel-level or incorporate the spatial domain on a local neighborhood level (i.e. directly neighboring pixels) only. Furthermore, only a small fraction of these studies employs VHR imagery.

The objective of this study is to evaluate and discuss the capability of different partitioning clustering algorithms for unsupervised two-class change detection of individual buildings using VHR imagery. Seven clustering algorithms that basically satisfy the needs of unsupervised change detection are evaluated since the most qualified algorithm for any specific application cannot be determined a priori. Furthermore, comprehensive feature sets of different characteristics are included in the evaluation in order to indicate the basic suitability of distinct types of features for change detection of buildings regarding different algorithms.

II. STUDY AREA AND DATA SETS

The evaluation of clustering algorithms for unsupervised change detection is conducted in the city of Dongying, China, which is located in the Yellow River Delta south of Beijing. Dongying is a typical example of emerging Chinese cities and exhibits dynamic urban growth. The experimental site within the study area covers a characteristic excerpt of the city comprising about 550 buildings.

Building geometries are available from segmentation of a normalized Digital Surface Model (nDSM) with 4m spatial resolution [11], which was acquired on October 18, 2013 (t_0). The building geometries correspond to building footprints due to their derivation from the nDSM. Optical imagery with very-high spatial resolution is available by means of QuickBird data acquired on February 25, 2007 (t_{0-1}) and WorldView data recorded on January 17, 2013 (t_0), respectively. Both image products possess a spatial resolution of 50cm and are available

This work was supported by means of the DELIGHT project (www.delight.eoc.dlr.de, grant no. 02WCL1249A and 02WCL1249I) as well as the project EO-CITI (grant no. 033LK001A and 033LK001B), both funded by the German Federal Ministry of Education and Research, BMBF.

with four spectral bands (blue, green, red, NIR). Although the individual bands of the two sensors cover slightly different wavelengths, the deviation of their respective mean values is marginal. The optical images were geometrically adjusted with a root mean square error (RMSE) of co-registration of 1.5 pixels (i.e. 75cm). In addition, the imagery was atmospherically corrected using the Atmospheric and Topographic Correction (ATCOR) toolbox [12] and radiometrically adjusted by means of band-by-band histogram matching.

III. METHODOLOGY

The workflow of this study is structured according to [13] and conducted as follows: First, different sets of object-based features are calculated based on multi-temporal VHR imagery and building geometries. Second, change detection in terms of unsupervised discrimination of changed and unchanged buildings is implemented by means of clustering. Finally, seven partitioning clustering algorithms are opposed and validated for evaluation of their applicability for change detection in the context of urban growth at the building level.

A. Generation of Object-Based Features

Different sets of object-based features (Table 1) are calculated based on each VHR data set (i.e. QuickBird at (t_0-1) and WorldView at t_0) and common object geometries (i.e. building footprints). The multi-temporal features of each building are subtracted in terms of $t_0 - (t_0-1)$ in order to directly incorporate the temporal relationship (i.e. information on possible changes between the two images). The employed features comprise a comprehensive selection as listed in Table 1. In addition to the three feature sets of multispectral (MS), ratio and grayscale texture features, a merged feature set consisting of all 33 features is considered.

B. Change Detection Workflow

First, the object-based features were normalized using minimum-maximum normalization to a common range of values of $[-1, 1]$. Second, a principal component analysis (PCA) is conducted in order to eliminate redundancies that are present in the comprehensive input data set. Since PCA rearranges the data by transformation of redundant information to principal components (PCs) of higher order, the number of relevant PCs is determined applying a tripartite procedure as presented in [13].

The initial analysis for selection of relevant PCs is a Bartlett's test, which determines if the first PC significantly deviates from the remaining PCs (i.e. if structure with respect

to the PCs is present in the data). This is followed by investigation of cumulative proportion of variance, whereas the summarized variance of the considered PCs must exceed the common threshold of 0.9. Subsequently, an averaged parallel analysis is conducted, which returns the number of PCs associated to their respective eigenvalue exceeding the eigenvalue of random data of equal dimensions. The second and third conditions are wrapped up by means of a maximum rule in order to consider both methods. More details on the selection procedure for relevant PCs can be found in [13].

Finally, the relevant PCs are classified into the two target groups (i.e. changed and unchanged buildings) by means of clustering. For this purpose, partitioning clustering algorithms are employed and evaluated in this study.

C. Evaluation of Clustering Algorithms

Seven relevant and suitable partitioning clustering algorithms from different methodological categories are evaluated since the most qualified algorithm cannot be chosen a priori [4]. With respect to the scope of unsupervised change detection, the prerequisite to relevant techniques is the potential of fully automatic implementation, i.e. no additional arbitrary input parameters for processing besides the number of clusters are required (or otherwise a reliable technique for the estimation of these parameters is available). In addition, the algorithms under study are intended to cover different methodological categories and represent well-established approaches in partitioning clustering. Thus, the following techniques are chosen in this study (with utilized R packages):

1. k-means $\{stats\}$
2. partitioning around medoids (PAM) $\{cluster\}$
3. genetic k-means $\{skmeans\}$
4. kernel k-means $\{kernelab\}$
5. spectral clustering $\{kernelab\}$
6. self-organizing map (SOM) clustering $\{som\}$
7. expectation-maximization (EM) clustering $\{mclust\}$

Most clustering algorithms under study can be implemented and initialized in a completely unsupervised way when the number of clusters is set to "two" (referring to changed and unchanged buildings, respectively) a priori. With respect to arbitrary input parameters, the EM clustering algorithm encompasses selection of an appropriate mixture model, which can be initialized using a Bayesian model selection procedure as shown in [15]. For kernel k-means and spectral clustering,

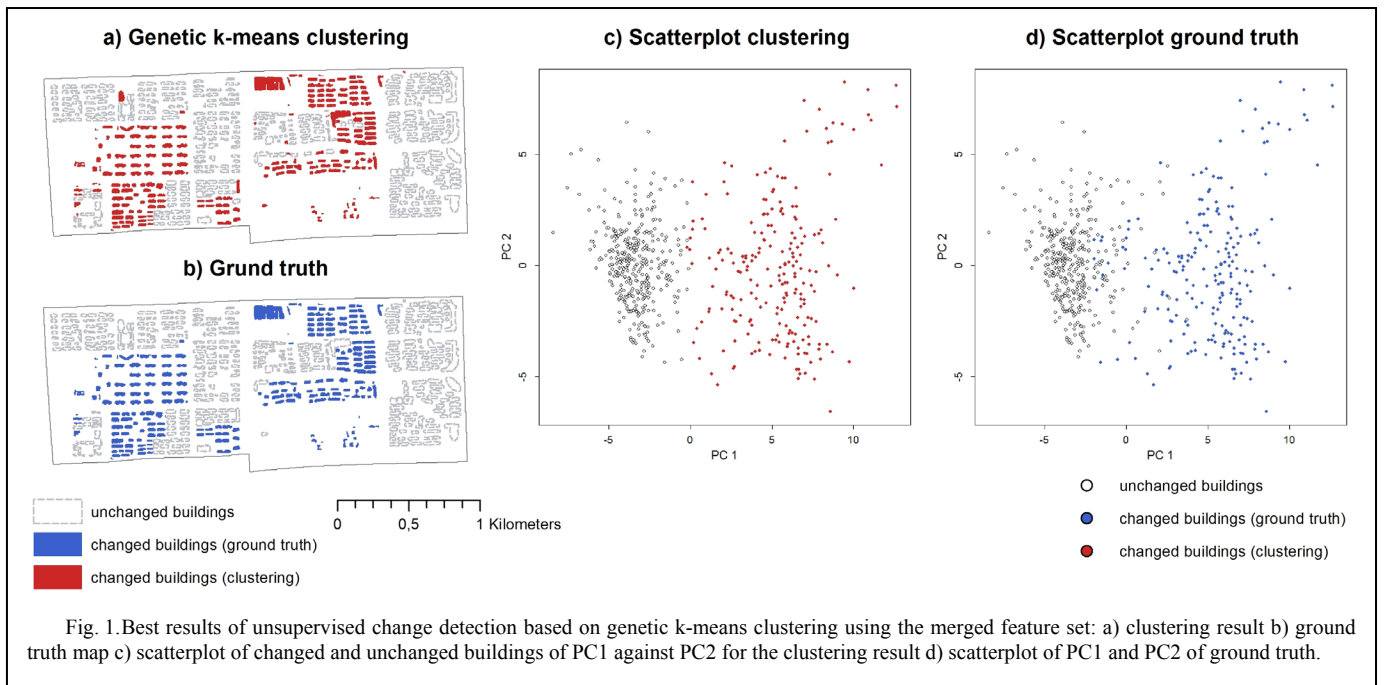
TABLE I. LIST OF OBJECT-BASED FEATURES

| Multispectral (MS) Features (10) | Ratio Features (11) | Grayscale Texture Features (12) |
|----------------------------------|-------------------------------|-------------------------------------|
| mean intensity blue band | blue / green | GLCM angular 2 nd moment |
| mean intensity green band | blue / red | GLCM contrast |
| mean intensity red band | blue / NIR | GLCM correlation |
| mean intensity NIR band | green / red | GLCM dissimilarity |
| mean intensity all bands | green / NIR | GLCM entropy |
| standard deviation blue band | red / NIR | GLCM homogeneity |
| standard deviation green band | normalized green ^a | GLCM mean |
| standard deviation red band | normalized red ^b | GLCM standard deviation |
| standard deviation NIR band | normalized NIR ^c | GLDV angular 2 nd moment |
| maximum difference between bands | NDVI ^d | GLDV contrast |
| | SAVI ^e | GLDV entropy |
| | | GLDV mean |

$$\text{grayscale image} = (\text{blue} + \text{green} + \text{red}) / 3$$

^a normalized green = green / (NIR + red + green), ^b normalized red = red / (NIR + red + green), ^c normalized NIR = NIR / (NIR + red + green)

^d NDVI = (NIR - red) / (NIR + red), ^e SAVI = [(NIR - red) / (NIR + red + L)] * (1 + L), where L = 0.5



the Gaussian RBF kernel is utilized due to its versatility [14]. The Gaussian RBF kernel transformation comprises the arbitrary parameter of inverse kernel width within the kernel function [14]. This parameter must be initialized and is estimated by a common heuristic using the median value within the 0.1 and 0.9 quantile of the mean squared error of the input data [14]. Since the estimation of inverse kernel width is based on a subset of the original data and thus causes varying results, the change detection procedure is iterated 1.000 times for all feature sets in case of kernel k-means and spectral clustering. Concluding, the median value of accuracy in terms of Cohen's Kappa (κ) is used for final evaluation of change detection accuracy.

IV. RESULTS

After calculation and normalization of the object-based features, PCA was conducted. The number of relevant PCs was determined by means of a tripartite selection procedure. The Bartlett's test was highly significant for the PCs of all feature sets, i.e. relevant structure with respect to PCA is present in the data. The investigation of the cumulative proportion of variance resulted in two relevant PCs for each set of MS, ratio and grayscale texture features, respectively, while four PCs were needed in case of the merged features to exceed the defined threshold of 90 percent of variance. Finally, the averaged parallel analysis returned two relevant PCs for the MS features, three relevant PCs for both ratio and grayscale

texture features and four relevant PCs in case of the merged feature set. According to the maximum rule of the second and third condition, two PCs were retained for the MS features, three PCs for ratio and grayscale texture features and four relevant PCs in case of the merged feature set.

Based on these relevant PCs, change detection was conducted by means of seven partitioning clustering algorithms and summarized values of resulting κ statistics are displayed in Table 2. The best overall result was achieved based on the relevant PCs of the merged feature set in combination with the genetic k-means clustering algorithm. This combination shows very good classification agreement with only few errors of commission and omission, which were present at approximately equal shares (Figure 1). In general, the MS and merged features showed good classification results, i.e. mean κ around 0.8 for all clustering algorithms. Moderate results of mean κ around 0.6 were returned for ratio and grayscale texture features due to less favorable distributions of the classes of interest (i.e. changed and unchanged buildings) in the PC feature space.

With respect to different clustering algorithms under study, superior results were achieved using k-means, PAM, genetic k-means and SOM clustering, whereas the remaining algorithms, i.e. kernel k-means, spectral clustering and the EM algorithm, returned worse results across all feature sets (Table 2). Especially spectral clustering and the EM algorithm showed

TABLE II. SUMMARIZED VALUES OF KAPPA (κ)

| Feature set | <i>k-means</i> | <i>PAM</i> | <i>genetic k-means</i> | <i>kernel k-means*</i> | <i>spectral clustering*</i> | <i>SOM</i> | <i>EM</i> | Mean of clustering algorithms |
|-----------------------------|----------------|-------------|------------------------|------------------------|-----------------------------|-------------|-------------|-------------------------------|
| Multispectral (MS) | 0.87 | 0.86 | 0.84 | 0.67 | 0.85 | 0.84 | 0.80 | 0.82 |
| Ratio | 0.73 | 0.73 | 0.66 | 0.41 | 0.57 | 0.67 | 0.06 | 0.62 |
| Grayscale texture | 0.81 | 0.83 | 0.83 | 0.56 | 0.11 | 0.82 | 0.22 | 0.57 |
| Merged | 0.90 | 0.89 | 0.91 | 0.78 | 0.82 | 0.90 | 0.78 | 0.76 |
| Mean of feature sets | 0.83 | 0.83 | 0.81 | 0.61 | 0.59 | 0.81 | 0.47 | |

* median value of 1.000 iterations

highly variable accuracy with respect to different feature sets with lowest accuracies of κ around 0.1. The accuracies of kernel k-means were slightly higher throughout, but showed high variation among the 1,000 iterations. In contrast, k-means, PAM, genetic k-means, SOM clustering and the EM algorithm exhibited stable change detection results without any variations since they do not require estimation and initialization of crucial input parameters. As mentioned above, κ was substantially lower for ratio features across all seven clustering algorithms due to overlapping of classes in the PC feature space. Nevertheless, k-means as well as PAM achieved best results for ratio features in terms of $\kappa > 0.7$. In case of grayscale texture features, spectral clustering and the EM approach were sensitive against outliers in the PC feature space. Outliers were only present in the grayscale texture features due to few (ca. 20 objects) extraordinary strong changes from very homogeneous undeveloped surfaces at t_0-1 to highly textured buildings at t_0 . The clustering algorithms PAM and genetic k-means were capable of properly handling these outliers and returned good classification accuracy of $\kappa = 0.83$.

V. CONCLUSION

This study presents an unsupervised change detection approach for VHR remote sensing imagery investigating the change (i.e. growth) of the building inventory in urban areas. More specific, the aim of this study was to evaluate the capability of seven partitioning clustering algorithms under consideration of different sets of object-based input features.

The most important prerequisite for clustering algorithms employed in unsupervised change detection is that no additional input parameters besides the number of clusters are required. Although there are commonly utilized heuristics for estimation of arbitrary initialization parameters, this study showed that these unsupervised estimation procedures may cause varying classification results that are outperformed by other, more robust clustering algorithms. Furthermore, relatively simple methods like the well-established k-means algorithm provided among the best results in the presented unsupervised change detection context. Other favorable techniques were PAM, genetic k-means and the SOM clustering approach. With respect to different feature sets, the comprehensive merged feature set consisting of MS, ratio and grayscale texture features provided best results. An eligible alternative is the exclusive use of MS features, which performed similarly well in the presented unsupervised change detection analysis.

ACKNOWLEDGMENT

The authors would like to thank Matthias Boes (European Space Imaging, EUSI) for providing the WorldView imagery and Dr. Klaus Martin (Company for Remote Sensing and Environmental Research, SLU) for his support.

REFERENCES

- [1] L. Bruzzone and F. Bovolo, "A novel framework for the design of change-detection-systems for very-high-resolution remote sensing images," *Proceedings of the IEEE*, vol. 101, pp. 609-630, 2013.
- [2] M. Hussain, D. Chen, A. 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*, vol. 80, pp. 91-106, 2013.
- [3] H. Taubenböck, T. Esch, A. Felbier, M. Wiesner, A. Roth, and S. Dech, "Monitoring urbanization in mega cities from space," *Remote Sensing of Environment*, vol. 117, pp. 162-176, 2/15/ 2012.
- [4] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognition Letters*, vol. 31, pp. 651-666, 2010.
- [5] T. Celik, "Unsupervised change detection in satellite images using principal component analysis and k-means clustering," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, pp. 772-776, 2009.
- [6] Y. Zheng, X. Zhang, B. Hou, and G. Liu, "Using combined difference image and k-means clustering for SAR image change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, pp. 691-695, 2014.
- [7] M. Volpi, D. Tuia, G. Camps-Valls, and M. Kanevski, "Unsupervised change detection by kernel clustering," in *Proc. SPIE 7830, Image and Signal Processing for Remote Sensing XVI*, Toulouse, 2010, pp. 78300V-78300V-8.
- [8] M. Volpi, D. Tuia, G. Camps-Valls, and M. Kanevski, "Unsupervised change detection with kernels," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, pp. 1026-1030, 2012.
- [9] A. Ghosh, N. S. Mishra, and S. Ghosh, "Fuzzy clustering algorithms for unsupervised change detection in remote sensing images," *Information Sciences*, vol. 181, pp. 699-715, 2011.
- [10] K. Ding, C. Huo, Y. Xu, Z. Zhong, and C. Pan, "Sparse hierarchical clustering for VHR image change detection," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, pp. 577-581, 2015.
- [11] M. Wurm, H. Taubenböck, M. Schardt, T. Esch, and S. Dech, "Object-based image information fusion using multisensor earth observation data over urban areas," *International Journal of Image and Data Fusion*, vol. 2, pp. 121-147, 2011.
- [12] R. Richter, "A spatially adaptive fast atmospheric correction algorithm," *International Journal of Remote Sensing*, vol. 17, pp. 1201-1214, 1996.
- [13] T. Leichtle, C. Geiß, M. Wurm, T. Lakes, and H. Taubenböck, "Unsupervised change detection in VHR remote sensing imagery – an object-based clustering approach in a dynamic urban environment," *International Journal of Applied Earth Observation and Geoinformation*, vol. 54, pp. 15-27, 2017.
- [14] A. Karatzoglou, A. Smola, K. Hornik, and A. Zeileis, "kernlab - An S4 package for kernel methods in R," *Journal of Statistical Software*, vol. 11, pp. 1-20, 2004.
- [15] C. Fraley and A. E. Raftery, "Model-based clustering, discriminant analysis, and density estimation," *Journal of the American Statistical Association*, vol. 97, pp. 611-631, 2002.