

CHANGE DETECTION BASED ON FEATURES INVARIANT TO MONOTONIC TRANSFORMS AND SPATIAL CONSTRAINED MATCHING

Marco Túlio A. N. Rodrigues, Luciano O. Milen, Erickson R. Nascimento, William Robson Schwartz

Computer Science Department
Universidade Federal de Minas Gerais, Belo Horizonte - MG, Brazil
Email: {tulio.rodrigues, lucianomilen, erickson, william} @dcc.ufmg.br

ABSTRACT

Discovering regions that have changed in a set of images acquired from a scene at different times and possibly from different view points and cameras is a crucial step for many image processing applications. Remote sensing, visual surveillance, medical diagnosis and treatment, civil infrastructure, and underwater sensing are some examples of such applications. This work proposes a novel approach to detect changes automatically without a learning step by using image analysis techniques and segmentation based on super-pixels. Unlike most common approaches, which are pixel-based, we present an approach that combines super-pixel extraction, hierarchical clustering and segment matching. The experimental results show the effectiveness of the proposed approach comparing it a background subtraction technique, demonstrating the robustness of our algorithm to illumination variations, non-uniform attenuations, atmospheric absorption and swaying trees.

Index Terms— change detection, super-pixel, hierarchical clustering, region growing.

1. INTRODUCTION

Finding changes in a scene plays a central role in a myriad of applications. Disaster management [1], urban growth [2], security [3], burned areas [4] and surveillance [5] are few of examples of problems which the knowledge of local changes is essential. Detecting structural changes is useful to gather information from the environment which may have social and economic impact.

One instance of the importance of change detection is the monitoring of oil pipelines (Figure 1), since there are many factors used to determine oil prices and one of them is the shutdown of a pipeline. In this application, changes in the environment such as the construction of new buildings, natural disasters like tsunamis and earthquakes are some of the causes of hazards around pipelines. Therefore, a timely and precise detection of activities affecting oil pipeline integrity ensures a faster problem resolution, helps preventing disasters, reduces maintenance costs and payout for damages. However, the usual procedure to detect problems in the neighborhood of a pipeline is to use human operators watching a video from a monitoring camera and identifying changes in images from the scene.

In general, video monitoring performed by humans is error prone due to the known lack of attention in repetitive tasks that take many hours [7]. Therefore, the automation of the monitoring is a direct and intuitive solution to reduce such errors. Besides the capability of reducing errors and speeding the monitoring process up, the employment of an automatic system can be used as a filter to provide a selection of important video segments that should be



Fig. 1: Trans-Alaska Pipeline on slider supports where it crosses the Denali Fault [6]. This pipe suffered some damage, but no oil spills occurred. A large arc in the pipe can be seen in image (b), due to shortening of the zigzag-shaped pipeline trace in the fault zone.

further analyzed by the operators, which would help them in their decision making process regarding actions to be executed.

The basic task of change detection in images is to find pixels (or regions) from a *reference image* that are different from other images [8]. Significant changes may include appearance or disappearance of objects, motion of objects, and shape changes in objects. On the other hand, variations in the image caused by camera motion, noise, illumination changes, nonuniform attenuation, atmospheric absorption, swaying trees, rippling water or flickering monitors, must be ignored by the change detection method.

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. Mostly are pixel-based methods and assume that pixels are independent. In remote sensing, the image change detection performance depends on a trade-off between the spatial resolution and the spectral resolution [9]. In addition to the particular drawbacks of background subtraction and remote sensing techniques, a learning step might be necessary to identify the characteristics of the images, which is an unfeasible task when only a single reference image is available.

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. Unlike the background subtraction and the remote sensing approaches, our technique does not require a learning stage and it is capable of detecting changes by using only two images (the reference image and a second image, referred to as *test image*, possibly containing changes).

Our technique is based on a computationally simple and effective change detection. It first extracts the super-pixels of the image, then performs a hierarchical clustering to obtain segments with simi-

lar characteristics, and finally matches segments of the reference image to segments in the test image based both on features (extracted using the Local Binary Patterns [10] to be invariant to monotonic pixel intensity changes and easy to compute) and spatial location constraint (considering that a change only takes place when corresponding segments present different characteristics). As a result, the test image receives a score proportional to the likelihood of a change takes place when compared to the reference image.

To evaluate our approach, we perform several experiments using datasets acquired under controlled and uncontrolled conditions. We also compare the proposed method to the widely used method for background subtraction MoG (Mixture of Gaussians) proposed by Zivkovic [11]. According to the experimental results, our approach outperforms MoG based approach in the change detection task.

2. RELATED WORK

Over the years, the image processing and computer vision community have been proposing a large number of change detection techniques, many of them based on background subtraction and remote sensing techniques, which will be briefly reviewed. For a comprehensive survey of different approaches for change detection the reader is referred to [8].

Change detection Based on Background Subtraction. In general, change detection approaches based on background subtraction compare the pixels of a test image to a background model learned using several reference images that do not contain any object. Then, pixels that do not fit the background model are classified as foreground, which indicates changes [12, 13, 14, 15]. Background subtraction techniques can divided in parametric or non-parametric.

In the parametric methods it is assumed that each pixel in the scene can be modeled as a statistical process that can be approximated by a given parametric distribution or a mixture of distributions. A popular method that use Mixture of Gaussians (MoG) was proposed by Stauffer and Grimson [16]. Another parametric method is presented by Zivkovic [11], that extend [16]. Zivkovic developed an efficient adaptive algorithm using Gaussian mixture probability density. By choosing the number of mixtures components for each pixel in an on-line procedure the algorithm can automatically fully adapt to the scene. Our proposed method, however, does not need the learning step and is not pixel-based.

In non-parametric approaches, the background model is not modeled by a parametric distribution [17]. The values of each pixel in the image are used to estimate the probability density function (PDF) of each pixel. So, the main characteristic is the dependence heavily data depend (many of them require the pixels to be stored), but they are capable of adapting to rapid changes in the background, such as a light is switched on or off, which is desirable.

Although the background subtraction approaches are usually simple and easy to implement, virtually all these approaches have a training step. Thus, they will no work properly whenever the background scene suddenly changes or not have enough samples to estimate the background model. Therefore, they are unfeasible when it is provided only a single reference image.

Remote Sensing Techniques in Change Detection. Similar to the background subtraction methods, the remote sensing techniques in change detection also need a learning stage to model the reference image [18]. However, unlike the former, the remote sensing approaches are based on feature extraction, which avoids errors due to pixels noise issues and small changes in the reference image. Benedek et al. [19] proposed a model that eliminates the registration errors and noises caused by the static objects having consider-

able height as buildings, trees and walls from the difference between the test and reference images. Similarly to Benedek et al. [19], we propose an approach based on features descriptors but the match is constrained to segments located in similar regions in both images.

An approach for change detection in multespectral remote sensing images was proposed by Bovolo et al. [20]. The technique is based on a compressed representation, in which the multidimensional feature space is compressed into a 2-D subspace without neglecting any available spectral band, returning possible information about changes. However, differently from our approach, which is based on a single and simple feature descriptor, Bovolo et al. [20] employed two complex feature descriptors. The first is the magnitude of the spectral channel vector and the second is a direction variable that models the information about different kinds of change without rejecting any spectral channel. In addition, a learning step was necessary, which is not performed by our approach.

All approaches described in this section (based on background subtraction and remote sensing techniques) require a learning stage and present high computational evaluation cost. In addition, they both are pixel-based, indicating the specific location of the changes in the image, which is susceptible to noise, such as illumination changes, and not always required since an estimation of changes able to rank the frames according to change likelihood would be enough to many application.

3. PROPOSED APPROACH

The approach proposed in this work is composed of the three main steps: (1) Super-pixel extraction; (2) hierarchical clustering, and (3) matching of segments. These steps are illustrated in Figure 2 and will be described in details in the next sections.

Our method receives two images as input, a reference image and a test image. Its output is a score that indicates if there is a change between the test image and reference image. Therefore, the goal of our algorithm is to decide whether there is a significant change between two images.

3.1. Super-pixel

In general, pixel-based methods are very susceptible to noise and do not provide semantic information. A common approach to overcome these issues is the use of super-pixel techniques. Originally developed by Ren and Malik [21], super-pixel consists in extracting a higher level representation for images by partitioning an image into meaningful regions. This method is more robust to noise than pixels and is useful to represent objects in images.

In our methodology, we have used the super-pixel approach developed by Radhakrishna et al. [22]. Their approach is built on the Simple Linear Iterative Clustering (SLIC), which generates super-pixels by clustering pixels based on their color similarity and proximity in the image. This technique is capable of producing regular super-pixel shapes, and it also can be employed on both grayscale images and color images. The parameter regularizer reflects the clustering appearance.

3.2. Hierarchical Clustering

Our approach combines the widely employed Hierarchical Clustering technique with the super-pixel methodology to create meaningful segments in the images. This avoids, for instance, an object to be represented by two different super-pixels. In such cases, these super-pixels need to be merged into a new segment. The Hierarchical Clustering step has two main stages: i) Extraction of a binary

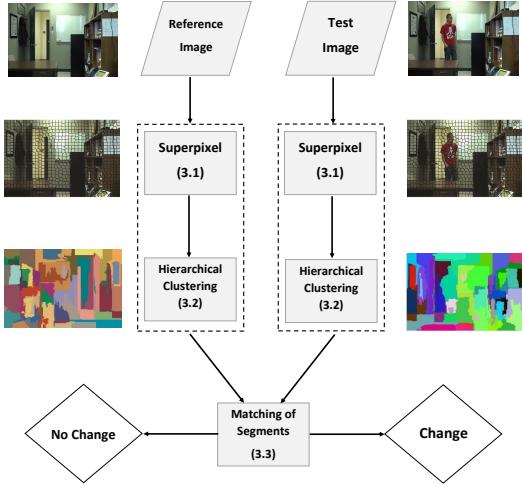


Fig. 2: Diagram showing the change detection approach employed in this paper. Each block is detailed in the Section with similar name.

hierarchical tree and ii) region growing using a similarity measure and spatial constraints. The first step groups the segments generated by the super-pixel extraction using a binary hierarchical cluster tree. Then, by using a threshold, we define the height in the tree from which all leafs will be merged.

Binary Hierarchical Tree. To start the construction of the tree, each super-pixel represents a leaf node. Therefore, at the beginning there are k leaves for the k super-pixels or segments. Then, the closest pairs of segments (according to a similarity measure) are merged iteratively into a new internal node, parent of the merged nodes, composed by the union of these segments. The process ends when all the super-pixels have been merged to a single node, the root node. This structure is then transformed in a dendrogram [23] which will be used to obtain the final clustering by performing a cut according to an experimentally defined threshold λ to produce segments can be merged into larger segments.

Region Growing. To avoid merging of non-adjacent segments, we use a region growing algorithm and the cutoff value, λ , of the dendrogram aforementioned. An initial set of small regions is iteratively grouped according to similarity constraints by the cutoff value. We begin by choosing an arbitrary seed pixel and comparing it with the neighboring pixels. Segments grow from the seed pixel by adding similar pixels in its vicinity. When the growth of one segment stops, we simply choose another seed pixel which does not belong to some segment and restart the iteration. This process is continuous until all pixels belong to some segment.

To provide robustness to lighting changes, since this is one of the major factors that produce false positive change detection in image, each segment in our methodology is described by a feature vector.

Due to its invariance to monotonic intensity changes, we extract features using the Local Binary Patterns (LBP) [10] which has also been employed in a variety of applications and is extremely fast to compute. The LBP value of a pixel $C = (x_c, y_c)$ and radius R is calculated by thresholding the gray value of N sampling pixels defined by the indicator function $s(x_1, x_2)$. This function returns 1 when the intensity value of pixel x_1 is greater than x_2 and 0 otherwise.

By considering the g_c as the intensity of the center pixel and g_p ($p = 0, \dots, N-1$) corresponding to the 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:

$$LBP_{N,R}(C) = \sum_{p=0}^{N-1} s(g_p - g_c) 2^p. \quad (1)$$

Thereafter, the LBP code given by Equation 1 is calculated for each image I using a N neighborhood (in this work we have used $N = 8$). Given a super-pixel or segment in an image I , the LBP descriptor is summarized in 256-bin normalized histogram. This way, a histogram is calculated for each image segment.

Similarity Measure. In order to characterize similarity among segments in Hierarchical Clustering and Segments Matching, we calculate the distance between LBP histograms extracted from each segment. In this work we evaluate the usage of three different distance functions: Euclidean distance, correlation coefficient and the cosine distance. Thus, the closest neighbor segment t of a segment r is given by

$$cn(r) = \arg \max_t \mathcal{D}(f, r, t), \quad (2)$$

where $\mathcal{D}(f, r, t)$ measures the similarity between segments r and t according to the similarity distance f . For Euclidean distance small values of $f(r, t)$ means high similarity, while for correlation and cosine distance large values of $f(r, t)$ means high similarity.

3.3. Matching of Segments

To be able to detect whether a change happened in a segment of the test image, this final stage of the proposed method searches for matching segments between the reference and test image based on the following idea. When two segmented images of the same scene present the same segments, it is likely that no changes have taken place. However, if different segments are present (e.g., a segment in the reference image correspond to two segments in the test image), a change may have happened. Therefore, to detect changes, we search for segments that should be spatially corresponding.

To detect changes, we first locate corresponding segments in both images. For a given segment in the test image, we locate which segment in the reference image corresponding to it by computing the intersection over the union ratio ($\theta = 0.8\%$) of pixel coordinates in both segments (this assumes that both images are fairly registered). The pairs of segments presenting the highest ratios are set as corresponding segments. This is basically performing a spatial constrained match. Such match still need to be validated by the region characteristics – the corresponding segments should present similar features, otherwise a change happened. To do that, for each pair of segments, the LBP feature descriptors are extracted and a similarity measure is computed. Then, a segment in the test image is declared as unchanged when the similarity measure presents a high value.

As the output of our method, we create a list containing one similarity value associated with each segment in the test image, in which segments associated with similarity values smaller than a threshold are said to present changes. Finally, the ROC curves are obtained by changing the threshold value.

4. EXPERIMENTAL RESULTS

To demonstrate the effectiveness and viability of the proposed approach, in this section we evaluate several of its aspects and compare it to a well-known background subtraction approach [11].

Datasets. The first, referred to as *controlled environment*, is performed to evaluate the our method in controlled conditions, such as lack on illumination changes, in which background subtraction

techniques are usually applied. The second experiment, referred to as *uncontrolled environment*, is employed to evaluate our method in more challenging conditions, containing illumination changes and small movements of the camera.

To perform the controlled environment experiment we use the Change Detection 2012 (OFFICE) dataset [24] and to perform the uncontrolled environment experiment we use a self collected dataset called *UFMG Change Detection Dataset*¹, which was captured from a parking lot containing waving trees and the camera suffers small movements. Samples of both datasets are shown in Figure 3.

It is important to note that for each group of experiments we considered one image as reference and the remaining as test images. Therefore, we want to detect changes according to the reference image. To evaluate the results we use the area under the curve (AUC) obtained from the Receiver Operating Characteristic (ROC) curve. The ROC is computed based on the true positive rate and the false positive rate, in which a test frame is considered a false positive (a frame without changes compared to the reference image that was classified as changed) when the amount of changes is larger than a threshold (used to generate the ROC curve).

4.1. Experimental Setup

In this section we evaluate the parameters that will be used for the super-pixel segmentation algorithm, choose the similarity measure and the color space that the proposed method will consider.

Super-pixel parameters. In Section 3.1, we explain the motivation to use super-pixels instead of pixels. The techniques employed on the experiments are [22]. For the used implementation, two parameters have to be set: super-pixel or region size τ_1 and regular size τ_2 . Changing these parameters may affect the accuracy of the technique, which may produce worthless semantic regions. High values of τ_1 increases the computation time of the algorithm and high values of τ_2 produces more regular regions. These values will be set experimentally to $\tau_1 = 50$ and $\tau_2 = 0.01$.

Similarity measure. Three different measures were considered: Euclidean distance, correlation normalized coefficient and cosine distance. According to the results obtained in the OFFICE dataset, the cosine distance and the correlation coefficient achieved the highest values, AUC = 0.9985 and 0.9982, respectively, and the Euclidean distance achieved a AUC = 0.6959. As cosine and correlation are very close and most applications use correlation to work with histograms, we selected correlation as the measure similarity to be used.

Color space. We also used the OFFICE dataset to choose the color space. We tested in four setups: RGB, grayscale and LAB and the component H of the HSV, which obtained UAC = 0.9920, 0.7105, 0.9982 and 0.9935, respectively. We then, selected the LAB as the color space to be used.

Hierarchical tree. The parameter λ , the cutting value for the dendrogram, was set experimentally to 0.3, because it produces good visual segmentation. Therefore, all segments that are below this value are merged during the region growing stage.

4.2. Comparisons

In this section, we compare the proposed approach to the background subtraction technique proposed by Zivkovic [11] in the controlled and uncontrolled environments.

Controlled environment. In this experiment we evaluate our approach in controlled conditions (OFFICE dataset) considering 200

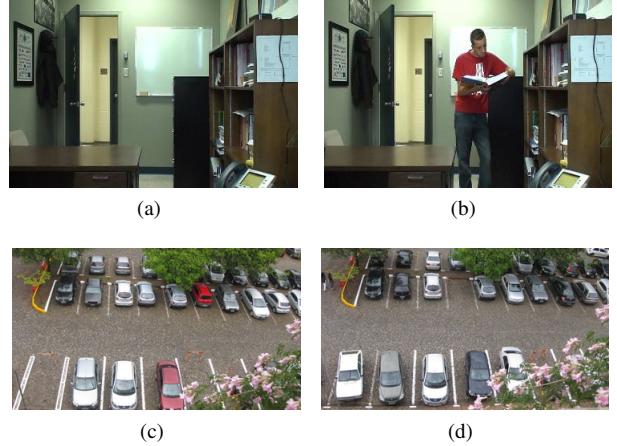


Fig. 3: Samples of the datasets. OFFICE dataset: (a) reference image and (b) example of a test image. UFMG Change Detection Dataset: (c) reference image and (d) example of a test image.

frames (100 with changes and 100 without changes), the first as reference and the remaining as test.

To compare to the background subtraction proposed in [11], we execute his method considering the reference image to build the background model and the remaining for testing. We then use the resulting foreground masks to decide whether an image should be considered as false positive according to the number of pixels classified as foreground that actually should be set as background pixels.

The resulting AUC for both, our method and the background subtraction technique were the same (AUC = 0.99). This results was expected since no illumination changes happened and the camera was static, condition favorable to background subtraction techniques which model the background for each pixel.

Uncontrolled environment. To demonstrate the effectiveness of our approach we also tested in a more realistic scenario (UFMG Change Detection Dataset). The images were taken during a day on influences of lighting, shadows and waving trees. When we compare our method to the background subtraction in this scenario, we obtain significant differences in the results, while the method proposed in [11] obtained AUC = 0.85, our approach achieved AUC = 0.95.

The different results can be explained by the influence of the illumination variation, changes in the background, camera movement, conditions that our approach handles well, differently from the background subtraction technique. Even though the image samples contain the influence of others characteristics, such as shadow, and the people passing, the performance of the approach did not degraded. This may be explained by the fact that we used regions and LBP features instead of single pixels to detect changes.

5. CONCLUSIONS

In this paper, we described a combination of existing techniques applied to the problem of change detection. The proposed approach uses super-pixels and a region growing strategy instead of single pixels. When compared to a background subtraction technique, the proposed approach achieved the best results in uncontrolled environment, demonstrating to be a promising technique to employed to change detection when a reference image is given. Future works include the addition of an automatic selection of the regions in the hierarchical tree.

¹The UFMG Change Detection Dataset will be released to public after the acceptance of this paper.

6. REFERENCES

- [1] T. Tadono, S. Hashimoto, M. Onosato, and M. Hori, "Applications of the automatic change detection for disaster monitoring by the knowledge-based framework," in *Spie Digital Library*, 2012, vol. 8528, pp. 85280T–85280T-9.
- [2] RA Alagu Raja, V Anand, A Senthil Kumar, Sandeep Maithani, and V Abhai Kumar, "Wavelet Based Post Classification Change Detection Technique for Urban Growth Monitoring," *Journal of the Indian Society of Remote Sensing*, pp. 1–9, 2013.
- [3] Moeness G Amin and Fauzia Ahmad, "Change detection analysis of humans moving behind walls," *IEEE Transactions on Aerospace and Electronic Systems*, 2013.
- [4] Daniel Capella, Hiran Zani, and Yosio Edemir Shimabukuro, "Automatic Detection of Burned Areas in Wetlands By Remote Sensing Multi-temporal Images," in *IEEE International Geoscience & Remote Sensing Symposium*, 2013.
- [5] Stefan Huwer and Heinrich Niemann, "Adaptive change detection for real-time surveillance applications," in *International Workshop on Visual Surveillance*. IEEE, 2000, pp. 37–46.
- [6] Wikipédia, "Construction of the Trans-Alaska Pipeline System," 2013.
- [7] Hina Keval, "CCTV Control Room Collaboration and Communication: Does it Work?," in *Human Centred Technology Workshop*, 2006.
- [8] R.J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam, "Image change detection algorithms: a systematic survey," *IEEE Transactions on Image Processing*, vol. 14, pp. 294–307, 2005.
- [9] Paul Aplin, Peter M Atkinson, and Paul J Curran, "Fine spatial resolution simulated satellite sensor imagery for land cover mapping in the United Kingdom," *Remote Sensing of Environment*, vol. 68, no. 3, pp. 206–216, 1999.
- [10] Timo Ojala, Matti Pietikäinen, and Topi Mäenpää, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [11] Zoran Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction," in *International Conference on Pattern Recognition*. IEEE, 2004, vol. 2, pp. 28–31.
- [12] T. Celik, "Unsupervised Change Detection in Satellite Images Using Principal Component Analysis and k-Means Clustering," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 4, pp. 772–776, 2009.
- [13] T. Celik, "Multiscale Change Detection in Multitemporal Satellite Images," *IEEE Geoscience and Remote Sensing Letters*, vol. 6, no. 4, pp. 820–824, 2009.
- [14] Jeisung Lee, Minkyu Cheon, Chang-Ho Hyun, Hyukmin Eum, and Mignon Park, "Adaptive background model for non-static background subtraction by estimation of the color change ratio," *Electronic Materials Letters*, vol. 9, pp. 33–38, 2013.
- [15] LingFeng Wang and Chunhong Pan, "Effective multi-resolution background subtraction," in *International Conference on Acoustics, Speech, and Signal Processing*. 2011, pp. 909–912, IEEE.
- [16] Chris Stauffer and W. Eric L. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, 2000.
- [17] Ahmed Elgammal, David Harwood, and Larry Davis, "Non-parametric model for background subtraction," in *European Conference on Computer Vision*, pp. 751–767. Springer, 2000.
- [18] Qian Xu and Lina J Karam, "Change detection on SAR images by a parametric estimation of the KL-divergence between Gaussian Mixture Models," in *International Conference on Acoustics, Speech and Signal Processing*. IEEE, 2013, pp. 2109–2113.
- [19] C. Benedek, T. Sziranyi, Z. Kato, and J. Zerubia, "A Multi-Layer MRF Model for Object-Motion Detection in Unregistered Airborne Image-Pairs," in *IEEE International Conference on Image Processing*, 2007, vol. 6, pp. VI-141–VI-144.
- [20] F. Bovolo, S. Marchesi, and L. Bruzzone, "A Framework for Automatic and Unsupervised Detection of Multiple Changes in Multitemporal Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 6, pp. 2196–2212, 2012.
- [21] Xiaofeng Ren and Jitendra Malik, "Learning a classification model for segmentation," in *IEEE Intl. Conf. on Computer Vision*. IEEE, 2003, pp. 10–17.
- [22] Radhakrishna Achanta, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk, "SLIC Superpixels," .
- [23] A. Webb, *Statistical Pattern Recognition*, A Hodder Arnold Publication, 2 edition, 1999.
- [24] N. Goyette, P. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, "Changedetection.net: A new change detection benchmark dataset," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2012, pp. 1–8.