TRACKING THE URBAN CHAMELEON – TOWARDS A HYBRID CHANGE DETECTION OF GRAFFITI

Benjamin Wild[1], Geert Verhoeven[2], Norbert Pfeifer[1]

[1] Research Unit Photogrammetry – TU Wien, Vienna , Austria – (benjamin.wild, norbert.pfeifer)@geo.tuwien.ac.at
[2] Ludwig Boltzmann Gesellschaft – LBI ArchPro (Austria), Vienna, Austria – geert@projectindigo.eu

KEY WORDS: Change detection, Computer vision, Cultural heritage, Feature matching, Graffiti, Multivariate alteration detection, Photogrammetry, SIFT, Street art

ABSTRACT:

Colourful and ever-changing: Graffiti can be considered the urban chameleon skin. At the *Donaukanal* (Eng. Danube Channel), Vienna's central waterway and one of the largest and most active graffiti-scapes worldwide, this metaphor applies like hardly anywhere else. Every day a multitude of graffiti is destroyed by the creation of new works. Recently, efforts have been made to mitigate this constant loss of cultural heritage along the *Donaukanal* by systematically documenting the graffiti, mainly using photography and photogrammetry. However, keeping track of the newly added works is very time-consuming and often like finding needles in a haystack, considering the large extent and high volatility of the monitored area. Thus, an automated graffiti change detection would significantly reduce the effort and avoid overlooking graffiti.

This contribution outlines the main challenges in image-based change detection for cultural heritage and proposes a hybrid graffiti change detection method. The investigated method exploits and combines an established pixel-based change detection algorithm, the Iteratively Multivariate Alteration Detection, with a novel descriptor-based method. The latter relies on image features, rather than pixels as analysis unit and can robustly filter false alarms from the high-performing but noise-prone pixel-based approach. Overall, the results indicate that the proposed method can largely automate image-based change detection of graffiti-scapes. It can uncover graffiti-related changes and robustly distinguish them from other image differences such as shadows but tends to overlook small-scale graffiti, indicating the need for further fine-tuning.

1. INTRODUCTION

Graffiti are full of contrasts. Although short-lived, sometimes only for several hours, they are omnipresent and significantly shape the appearance of our urban environments. Their volatility and colourfulness justify drawing parallels to chameleons (Verhoeven et al., 2023), which primarily use their colour-changing skin to communicate social signals (Ligon and McGraw, 2013). Just like the pigments of chameleons give insights into their behaviour, graffiti might act as a mirror and magnifying glass to human society. Despite some metaphorical similarities, one distinct difference between them stands out: unlike the colourful lizards, "contemporary" graffiti are hardly documented for scientific purposes. Properly documenting this highly fluctuating, extensive phenomenon requires significant resources, which are often hard to come by given that graffiti are often not considered cultural heritage but vandalism by the general public. Therefore, it is not surprising that documentation of graffiti has never received the academic attention some scholars demanded (Novak, 2014, Masimiliani, 2008). In 2021, project INDIGO (INventory and DIsseminate graffiti along the dOnaukanal), a two-year academic graffiti-centred research project, set out to change that.

Launched in September 2021, project INDIGO aims to introduce more scientific rigour in graffiti research via the development of methods to optimise the systematic photographic documentation, monitoring and analysis of graffiti-scapes (Verhoeven et al., 2022). Documentation implies recording every graffito's geometrical, spectral, geographical and contentual aspects. INDIGO focuses on the graffiti along Vienna's *Donaukanal* (Eng. Danube Channel; Figure 1), a central waterway famous for its graffiti-covered walls, which constitute one of the largest

uninterrupted graffiti-scapes in the world. Tools to colour-calibrate the digital photographs (Molada-Tebar et al., submitted) and to automatically turn these into high-resolution and georeferenced graffiti orthophotos (Wild et al., 2022, 2023) have been developed alongside a graffiti-centric thesaurus which supports the categorisation and annotation of the inventoried graffiti photos (Schlegel et al., submitted). While the above developments enable graffiti analysis at a large-scale and in great detail, their full potential is not yet exploited because of a critical bottleneck: Many, primarily smaller, graffiti disappear before they are documented or even noticed. This high volatility is especially pronounced at Donaukanal's *Wienerwände* (Eng. Vienna walls), legal graffiti surfaces where often only a few hours to maximally a couple of days separate a graffito's creation and coverage.
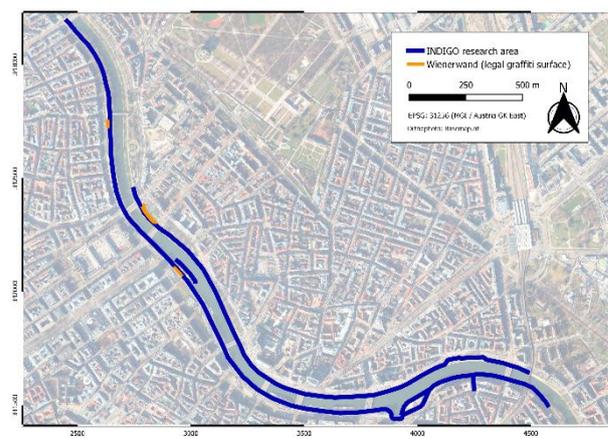

Figure 1. INDIGO's research area along Vienna's Donaukanal.

ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-M-1-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

To mitigate gaps in the documentation, project INDIGO has so far followed two systematic monitoring strategies. First, new graffiti are discovered by regularly walking or biking along INDIGO's research area, a route totalling ca. 11 km. This task solely relies on the photographer's memory. One could also consider detecting changes in the field by comparing previously acquired photos with the current graffiti cover via a mobile device. However, this approach has proven too time-consuming in practice (Verhoeven et al., submitted). A second method involves checking the social media entries of graffiti creators known to be active along the Donaukanal. Browsing their Instagram profiles allows for relatively efficient graffiti tracking. To optimise this social-media-based change detection, INDIGO's own hashtag, "#indigodonaukanal", has been promoted among graffitists. However, its uptake has been rather disappointing as it is only used by a few graffiti creators.

Once a new graffito is spotted, either online or onsite, its location gets marked in a mobile GIS system which guides the photo tours that take place at least once per week. While both methods are successfully applied in finding large graffiti, they fail when it comes to less striking creations like tags or political slogans, as they are shared less extensively via social media and are harder to memorise or see onsite. This leads to a documentation bias towards flashy and more sizeable works, counteracting the scientific approach that INDIGO envisions (Verhoeven et al., submitted). The current strategies also require significant human resources, besides being slow and tedious. Thus, it became clear that exhaustive documentation and digital safeguarding of the Donaukanal's graffiti-scape would be impossible without an automated graffiti change detection approach.

Many automated change detection techniques have emerged in recent years, mainly from satellite-based remote sensing and video surveillance (Radke et al., 2005). The latter was also the basis for existing graffiti change detection studies (Tombari et al., 2008; Di Stefano, 2008; Angiati et al., 2005). Those three studies used footage from video surveillance systems to automatically detect and identify "vandals" during their act of "vandalism". The change detection was thus more focused on the graffiti creation act rather than the actual final graffito. Despite exhibiting accurate results, these video-based methods cannot be applied along the Donaukanal. Setting up video surveillance would be extremely costly, logistically challenging and highly questionable, if not illegal, from a privacy protection perspective.

More feasible are short but frequent graffiti monitoring tours capturing the whole research area, for example, by mounting camera(s) on a bicycle or the biker's helmet. In that way, images covering the entire research area can be acquired in approximately one hour. After orienting these photos with an incremental Structure from Motion approach, creating sets of co-registered images for the entire area of interest becomes possible. The execution of this photo acquisition strategy is not trivial and requires bespoke photographic, logistical and photogrammetric solutions. Despite their importance and relevance, these solutions will only be sketched in this contribution (Section 2.1). The primary focus of this study is the development of a graffiti-aware change detection algorithm.

1.1 Challenges in image-based graffiti change detection

Change detection was defined by Singh (1989, p. 1) as "*the process of identifying differences in the state of an object or phenomenon by observing it at different times*". Following this definition, let us consider the example in Figure 2: two co-registered images of the same graffiti scene taken at different times (Figures 2a and 2b).
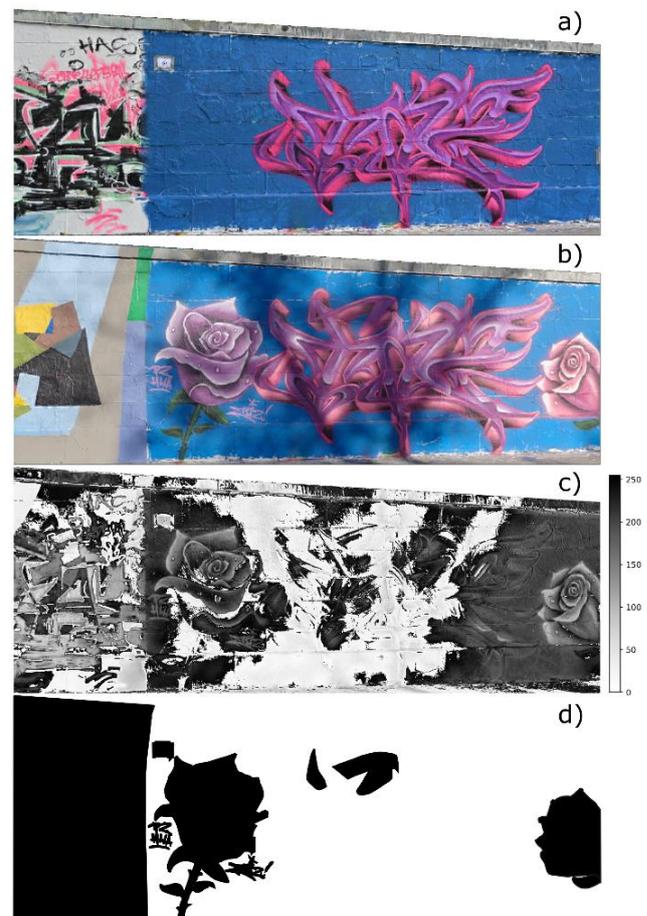

Figure 2. a) and b): co-registered images from the same graffiti scene at the Donaukanal taken on 12-10-2022 and 12-11-2022, respectively.
c): absolute differences in grey values between a) and b).
d): manually generated, ideal change map where black denotes change.

The human visual system allows for relatively quick and accurate identification of changes in the depicted graffiti cover, implicitly separating them from other differences in the image, such as shadows. However, this is a non-trivial task to be automated. The most obvious approach would probably be to subtract Figure 2b) from 2a) and classify changes between them based on the magnitude of the differences. The rationale is that the larger the difference, the larger the probability of a changed scene. While this method might produce appealing results in a laboratory setting with controlled illumination conditions and no environmental influences, it fails in real-world scenes. This is demonstrated in Figure 2c), where sunlit, mostly unchanged areas exhibit the largest differences while regions in the shade result in grey value encoded deviations close to zero. Thus, the main driver for pixel differences is not changed graffiti but the illumination conditions.

Pixel differencing is, of course, by no means considered state-of-the-art when it comes to image-based change detection. Still, it highlights that mapping radiometric differences in images is straightforward, whereas classifying them as relevant and irrelevant is the real challenge to image-based change detection (Bruzzone and Bovolo, 2013). Therefore, let us examine the reasons for the differences between the two co-registered images

ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-M-1-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

in more detail. For this specific graffiti case, a possible (non-exhaustive) list of reasons for differing (overlaying) pixel values is the following (Figure 3):

- Weather-related: Different weather conditions imply different ambient light conditions causing the texture to vary daily and even appear dissimilar throughout the day. These differences can become especially large between sunny and cloudy moments when shadows pose a significant risk of being misinterpreted as new graffiti due to substantial radiometric differences with relatively sharp transitions. Moreover, shadows generated by moving vegetation can rapidly change on windy days. Wet walls caused by rain can also cause significant alterations between two images.
- Acquisition-related: Different sensors and camera settings may produce significantly differing images. Besides camera-related differences, the camera platform plays a role. Varying incidence angles can result in spurious reflections and occlusions in the photos, and kinematic platforms such as bicycles are prone to cause image motion blur.
- Co-registration-related: Even small alignment residuals can lead to significant image differences. This is mostly relevant for change detection at pixel-level.
- Object-related: Lastly, changes can also be contentual. Although object-related changes should still be distinguished into relevant (i.e., graffiti-related) and irrelevant (i.e., non-graffiti-related) ones like by-passers, tackling this difference is not in the scope of this study.
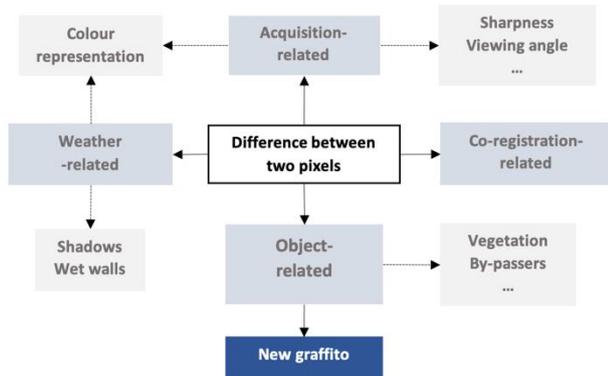

Figure 3. Sources and examples of differences between two overlapping pixels of a co-registered image.

A graffiti-aware change detection framework should be robust against all sources of change which are greyed out in Figure 3. Therefore, this contribution introduces a hybrid change detection framework which is expected to enable an automated change monitoring of the graffiti-scape along the Donaukanal.

The following sections briefly outline the image acquisition and necessary preprocessing, followed by an explanation of the proposed methodology. The performance of the developed change detection approach is analysed by applying it to various real-world examples gathered along the Donaukanal.

2. METHODS

The proposed framework exploits and combines two independent change detection methods (Figure 4):
1. The well-established, pixel-based, Iteratively Reweighted Multivariate Alteration Detection (IR-MAD; Nielsen, 2007)
2. A new descriptor-based change detection algorithm

The methods are expected to balance each other's drawbacks and provide robust and automatic identification of new graffiti in images. While accurate co-registration is not a strict requirement for the descriptor-based approach, the quality of the pixel-based IR-MAD is highly dependent on the co-registration accuracy achieved during the geometric preprocessing. Therefore the next section briefly sketches how nearly pixel-perfect co-registration is achieved in a highly automated way.
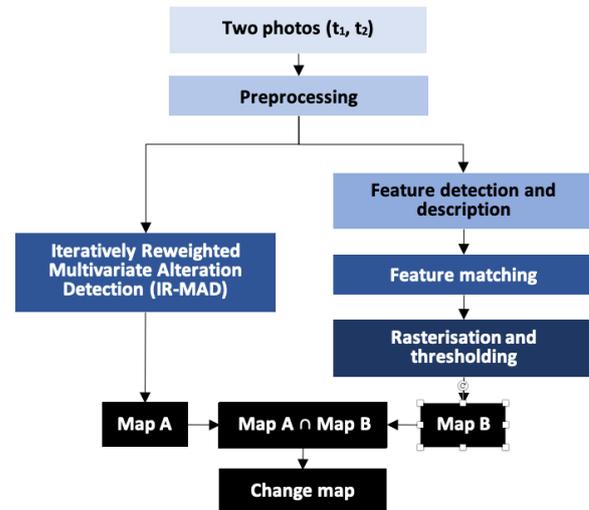

Figure 4. Workflow diagram of the hybrid graffiti change detection approach.

2.1 Image acquisition and preprocessing

Before starting this change detection study, extensive expertise in photographing this large graffiti-scape was already present. In the Spring of 2021, the entire Donaukanal graffiti-scape was photographed with a Nikon D750 and Nikon Z7ii (see Verhoeven et al. 2022 for all details). A few months later, the first tests with GoPro HERO10 Black action cameras started. Two action cameras were mounted on a camera bar, one looking to the left and another to the right. This mount was fit to a typical action camera handgrip, which allowed the dual-camera construction to be handheld while biking along the graffiti-covered walls and bridges (Verhoeven et al., submitted).

These initial tests provided much feedback on camera settings, biking speed, ideal biking route and potential post-processing issues. However, they mainly highlighted three significant issues:

1. A dual GoPro setup is insufficient to guarantee a problem-free exterior orientation of all the cameras via an SfM pipeline. When turning or biking along heavily vegetated areas, not enough robust tie points can be extracted, so those images usually fail to orient. Because the SfM-based orientation of cameras in an extended image network is very prone to drift in the estimated positions, these gaps in the oriented network also negatively influence the other cameras' estimated orientation.
2. The GoPro has various photo and video modes, each with many tuneable settings. Although only the photo interval mode was used (at two photos per second), multiple settings combinations are still possible. Combined with the different weather conditions, these settings – of which a few combinations were tried – seem to impact the SfM result more than expected. In turn, they also influence the outcome of most change detection algorithms. A more systematic approach was needed to tackle both issues.

ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-M-1-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

3. Biking with only one hand on the steering wheel while keeping an eye on the camera construction (to keep it steady) and simultaneously avoiding other bikers or walkers impacts overall traffic safety and the smoothness of photo acquisition negatively.

That is why in September 2022, new tests started with an additional GoPro HERO11 Black, which was just released back then. The three action cameras were mounted to a biking helmet (POC Crane Mips; https://www.pocsports.com/products/crane-mips) so that either side of the helmet holds a GoPro camera whose optical axis is circa 60° horizontally rotated from the front (thus sideward- but still forward-looking). A third camera is mounted on the rear helmet top and faces backwards. This setup allows better imaging of the biking surface and increases the range of observation angles for the graffiti-covered surfaces. Both outcomes are essential for tackling the image orientation problems in corners and vegetated areas mentioned above; they also substantially reduce occlusion zones, which is critical when a mesh needs to be extracted from the photos. Because the biker can hold the steering wheel with both hands and fully concentrate on other traffic, overall operation efficiency and safety have increased considerably.

However, the number of photos has now grown by 50 %. When aiming for an 80 % to 85 % longitudinal overlap between successive images, all acquired by the same camera at an approximate 4.5 m camera-to-wall distance and a rate of two photos per second, the biking speed should be between 14.9 km/h to 11.2 km/h. The lower speed is preferred as it minimises motion blur and geometrical distortions caused by GoPro's electronic rolling shutter. However, one photo tour then lasts about 70 min, yielding approximately 8400 images per camera for a total of 25 000 photographs. Even when one could likely halve the number of photos from the backwards-facing camera, orienting such a camera network would still take considerable time on high-end computer hardware. That is why – after checking that the helmet-based approach would work – testing the GoPro camera settings and acquiring the images for this study was done slightly differently to minimise processing times.

The same three GoPro cameras were mounted inside a frame composed of standard camera rig components (Figure 5). The cameras were mounted as closely as possible, with the optical axes mutually parallel. Every time the weather allowed, photos were acquired from the INDIGO test zone, a ca. 250 m stretch along the Donaukanal that includes two bridges and a legal graffiti area (as part of the Wienerwand) between them. The permitted area covers a flat horizontal area directly next to the Donaukanal and a ramp leading towards the street parallel to the waterway. Graffiti activity is very high here, and several graffiti-scape control points were determined during a total station survey at the start of the project (see Verhoeven et al. 2022).

The GoPro camera settings were changed for every photo acquisition to cover all possible camera settings-weather combinations (and study – at a later stage – their influence on the SfM and image change detection approach). Due to various reasons, only nine image acquisitions could be executed in November and December 2022, sufficient to cover circa 50 % of all possible combinations. More images will be acquired in May 2023 to cover all possibilities. Once that is done, a follow-up paper will report on the influence of these camera settings on camera orientation and change detection. This paper does not consider the influence of these image acquistion parameters.

Figure 5. The POC Crane Mips helmet onto which the three cameras can be mounted (left); Three GoPro cameras mounted inside a frame composed of standard camera rig components (right).

Each of the nine epochs thus counts three GoPro subsets. Every subset consists of two photos per second, acquired with the three-camera frame while walking at a usual 5 km/h pace. During the acquisition, the handheld frame was rotated 90° clock- and anti-clockwise as well as 180°. Most photographs featured a camera axis perpendicular to the graffiti-covered walls, but the entire ramp was also acquired with oblique or inclined optical axes. Such variation in image rotation and scale strengthens the camera network (Luhmann et al., 2016), which is especially important for such an elongated scene predominantly consisting of relatively flat walls.

All camera stations' interior and exterior orientations were estimated per epoch using an incremental SfM approach in Agisoft Metashape Professional 1.8.4 (Agisoft, 2022). A part of INDIGO's second total coverage photo survey, executed in October 2023, was used as the base image network. Once an epoch got oriented, three textures (one per GoPro subset) were generated for the mesh derived from the base image network. Given the scene's size, each texture atlas consisted of 5 tiles containing 4096 by 4096 pixels per tile. For every texture, the same UV parameterisation was used.

At this stage, the Metashape project counted nine epochs, each with three textures. Since the exterior orientations of the three cameras are almost identical at any moment, intra-epoch textures only differ due to the different GoPro settings (and to a smaller extent due to dissimilar image blending by Metashape). These textures can be used to study the robustness of the change detection algorithm to differing sharpness and contrast levels. However, real changes in the graffiti-scape must be extracted from the inter-epoch textures, which vary due to weather-related, acquisition-related and graffiti-related differences.

One could use the entire texture atlas to compute changes between intra- and inter-epoch textures. Still, such an approach is likely not scalable to extensive scenes like the whole Donaukanal, while the layout and seams of the texture patches might also negatively influence the change detection algorithm. That is why single synthetic photos are generated in Metashape. First, a camera path with more or less equally spaced camera positions is defined. Afterwards, a bespoke Python script renders a synthetic photo for each camera path station by observing one specific texture through those cameras. Assuming correctly estimated image orientations, those synthetic images should be nearly pixel-perfect co-registered. Without going further into detail, it was only possible to accomplish this by accounting for GoPro's electronic rolling shutter.

ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-M-1-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

2.2 Iteratively Reweighted Multivariate Alteration Detection (IR-MAD)

The acquired and preprocessed multi-temporal images are used as input to the Iteratively Reweighted Multivariate Alteration Detection (IR-MAD). Introduced by Nielsen in 2007, IR-MAD is an extension to the Multivariate Alteration Detection (MAD; Nielsen et al., 1998) and has been extensively used for change detection in multi-spectral remote sensing imagery. MAD allows for robust detection of uncorrelated information between the input images, which is a strong indicator for change. It is based on canonical correlation analysis (CCA). CCA finds linear combinations of the 6 (2×3) input bands, which maximise the correlation between each other (Hotelling, 1936). For bi-temporal, multi-band images, the canonical correlations are subtracted from each other. The resulting differences (i.e. MAD components) highlight potential change regions in the image (Nielsen, 2007). By setting a threshold for these MAD components, the image can be distinguished in change/no change. The threshold is determined by applying k-means-clustering (with k= 2), resulting in two clusters with minimal within-cluster variance. The great advantage of MAD is its invariance to linear scaling, making it insensitive to different illumination conditions and sensor settings.

In addition to ordinary MAD, IR-MAD iteratively assigns weights to pixels based on the magnitude of change detected during the previous iteration (i.e. pixels with minor change are assigned high weights and vice versa), making the change detection more robust (Nielsen, 2007). The iteration is terminated when the maximum difference in the canonical correlation falls below a threshold ε. Tests have shown that a threshold of $\varepsilon = 0.1$ is appropriate for the investigated graffiti use case. Larger values for ε tend to trigger a higher number of false alarms (i.e. false positives), while smaller values often lead to undetected changes (i.e. false negatives) and significantly increase the runtime. This study uses the IR-MAD implementation from ChenHongruixuan's ChangeDetectionRepository on GitHub (https://github.com/ChenHongruixuan/ChangeDetectionRepository).

Running IR-MAD on the images shown in Figures 2a and 2b yields the change map depicted in Figure 6. Changes in the graffiti cover are well highlighted, confirming the overall applicability of IR-MAD for the investigated use case. Despite the result's visually pleasing appearance, relying only on IR-MAD for the graffiti change detection will not suffice as the resulting change map is relatively noisy and contains several, albeit small, false positives. Especially at edges, change is often falsely detected (e.g. between the sandstone bricks or at the transition of different graffito layers). This susceptibility to noisy results is a well-known drawback of pixel-based approaches and is mainly related to unavoidable inaccuracies in image co-registration and strong changes in illumination between the two acquisitions (Tewkesbury et al., 2015; Figure 7). Tests have also shown that IR-MAD often fails in entirely changed/unchanged scenes (Figure 9b). This behaviour is explicable with the k-means-clustering approach which expects two classes and fails when only one is present. Therefore, an independent method which can largely compensate for the shortcomings of IR-MAD is introduced in the next section.

Figure 6. Results from MAD using the images from Figures 2a) and b) as input. Black denotes change.

2.3 Descriptor-based change detection

In his highly influential work from 1999, David G. Lowe writes, "*Object recognition in cluttered real-world scenes requires local image features that are (…) partially invariant to illumination, 3D projective transforms, and common object variations.*". Replacing "object" with "change" yields a statement very similar to the aims of this study. In particular, illumination invariance is an important feature in the context of this study. Thus, it seems logical that the key result from Lowe's work, the Scale Invariant Feature Transform (SIFT), is a promising starting point for image-based graffiti change detection. Specifically, the idea is to transform both images into a collection of local features (not restricted to SIFT features), each represented by a descriptor vector. Similar descriptors at similar positions in the co-registered images indicate no change around these points and vice versa. This principle has been tested in different variants for detecting changes in satellite imagery obtained from optical sensors (Seo et al., 2022; Liu et al., 2019) and Synthetic Aperture Radar (Wang et al., 2016; Pham et al., 2016), where its applicability was confirmed but mainly restricted to feature-rich areas like cities. The concept's transferability to conventional (terrestrial) images and applications has not been examined as far as the authors are aware.

The descriptor-based approach implemented for this study includes three main steps (Figure 7):
1. Detection of distinctive features and computation of descriptor vectors in both images using well-established feature detectors and descriptors (Figure 7a).
2. Matching of features based on their vicinity in feature and object space (Figure 7b).
3. Rasterisation of the matched feature points Figure 7b) and binary classification in changed/unchanged pixels based on the density of the matched feature points (Figures 7c and 7d).

2.3.1 Feature point detection and description: First, feature points, are detected in the input images. Those points are usually found at edges, corners or blobs (Figure 7a; Tareen and Saleem, 2018). This is beneficial as graffiti are often characterised by sharp transitions between the graffito-specific layers. In a subsequent step, the detected feature points are described on the basis of the unique patterns in their neighbouring pixels. This process is called feature description and results in a descriptor vector of fixed length for each detected interest point.

In this study, the respective OpenCV (Bradski, 2000) implementations of the following four well-established detectors and descriptors are used: SIFT (Lowe, 1999), SURF (Bay et al., 2006), AKAZE (Alcantarilla et al, 2013) and BRISK (Leutenegger et al., 2011). The rationale behind using several feature detectors and descriptors is that their different properties and sensitivities can provide partly independent and complementary information (Tareen and Saleem, 2018), thereby

ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-M-1-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

increasing the method's reliability. Every detector used the same-named algorithm for the feature description (i.e. feature points detected by SIFT were also described using SIFT).
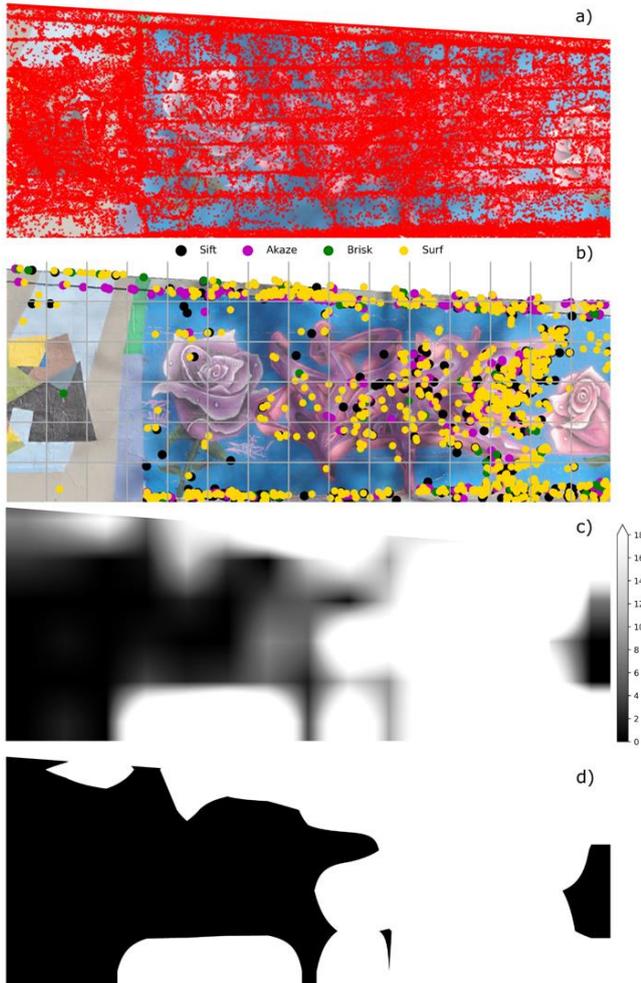

Figure 7. a) Detected feature points of all used detectors (SIFT, SURF, AKAZE, BRISK) for Figure 2a).
b) Matched and filtered feature points between the images depicted in Figure 2a) and 2b)
c) Bi-linear interpolated density map of the rasterised matched feature points
d) Descriptor-based change map. Black denotes change.

2.3.2 Feature matching: For matching the detected features in the two images, FLANN-based matching was applied to the set of detected feature points. FLANN stands for Fast Library for Approximate Nearest Neighbours. It consists of algorithms optimised for finding the nearest neighbours in high-dimensional data. For each detected feature point in image A, FLANN efficiently finds the feature point in image B for which the Euclidean distance of the corresponding descriptor vectors is minimal. This approach usually leads to a significant number of false matches as many features are only found in one image or arise from background clutter and are thus not sufficiently distinctive (Lowe, 2004).

To robustly filter false matches, we exploit the nearly pixel-perfect co-registration of the images by only considering matched points which are sufficiently close in object space. Specifically, we set a threshold of 20 px, approximately corresponding to a real-world distance of ca. 3 cm. Matched features which are farther apart are considered falsely matched (Figure 7b).

For image pairs with less accurate co-registration between the input images, one could increase the maximally allowed distance between matched features and, in return, additionally perform a Lowe Ratio Test (Lowe, 2004), which allows for filtering matches with non-discriminative, and thus likely falsely matched descriptor vectors. However, this was not necessary for the examples used in this study.

2.3.3 Rasterisation and thresholding: The matched and filtered feature points (which are collectively called tie points) are translated into changed/unchanged regions. Therefore, the image is divided into 400×400 px raster cells (Figure 7b). This raster size was empirically determined to be an appropriate compromise between achievable granularity and accuracy. Each cell is assigned the number of tie points within the boundaries of the respective cell. An increasing number of tie points indicates a decreasing likelihood for change. Bilinear interpolation was applied to achieve smoother transitions between the grid cells (Figure 7c). The resulting raster image holds information on the spatial distribution of (dis)similar features.

Thresholding transforms the density map of tie points into a binary change map (Figure 7d). A threshold of ten tie points points per raster cell was empirically determined to be an appropriate compromise between sensitivity and robustness of the results.

2.4 Derivation of final change map and postprocessing

Finally, the change map is computed by intersecting the IR-MAD (Map A) with the descriptor-based change map (Map B):

$$Change\ Map\ =\ Map\ A\ \cap\ Map\ B, \qquad (1)$$

Only pixels classified as 'changed' in both maps will be considered as change in the final map. The result is filtered by applying morphological opening with kernel sizes of 15×15 px for the erosion and dilation operations. The resulting final change map for the introduced example map can be seen in Figure 8.
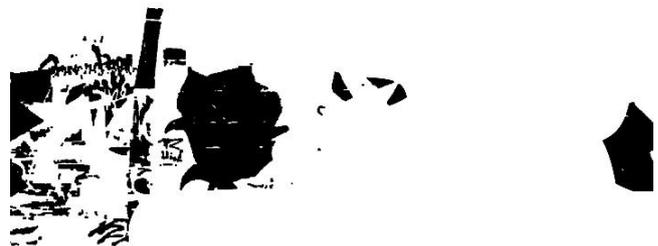

Figure 8. Final post-processed change map of the images depicted in Figure 2a) and 2b)

3. RESULTS

The proposed methodology was tested on 15 synthetic image pairs of INDIGO's test zone. While this number of tests is not sufficient to draw significant quantitative conclusions on the method's performance, it allows gaining first insights on possible advantages/disadvantages of the introduced hybrid change detection. The main results from this experiment are summarised below.

First, the proposed method leads to a very low number of false positives. Our tests consistently identified completely unchanged graffiti scenes as such (Figure 9b). This is mainly due to the high robustness of the descriptor-based method, which finds

ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-M-1-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

corresponding features even for image pairs where one or more of the potential pitfalls shown in Figure 3 are present. However, the high robustness of the descriptor-based methods comes at the cost of reduced sensitivity and spatial granularity, meaning that smaller graffiti, such as tags and stickers, are less likely to be detected. On the other hand, the IR-MAD detects nearly all new graffiti, even small tags, at the same time being noisy and resulting in a significant number of false alarms. By combining both methods, the IR-MAD noise is filtered, and the descriptor-based methods' low granularity can be partly surpassed, and change can be better condensed and highlighted. However, if the prime interest is not the exact nature of the change but only whether or not contentual change occurred between two photo acquisitions, solely the descriptor-based change method would suffice.
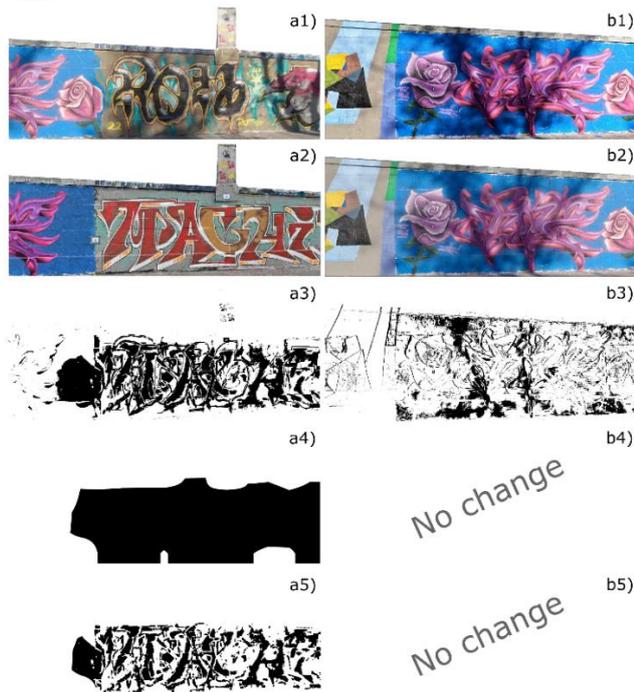

Figure 9. Two examples of a scene with (a1-a5) and without (b1-b5) graffiti-related changes. The first two rows show the co-registered input images. The third and fourth rows depict the IR-MAD and descriptor-based change maps. The respective final change maps are shown in the last row.

Considering the processing times, an average change detection on a 6000×4000 px image pair takes ca. 60 seconds as a single-threaded process. Approximately 40 seconds are needed to compute the IR-MAD change map, while the feature extraction, matching, rasterisation and classification are finished in ca. 20 seconds.

4. CONCLUSIONS

This paper presented a graffiti change detection method that detects new graffiti in two images taken at different times. The proposed method can largely distinguish between content-related changes and content-unrelated radiometric changes, such as shadows or differences in colour representation.

The proposed method should be understood as a starting point for subsequent research to develop an automated, image-based workflow for detecting graffiti-scape-related changes. Future improvements are envisioned in all parts of the current workflow. In particular, further research is needed to optimise photo acquisition and preprocessing. Other image preprocessing techniques, such as histogram equalisation or white balance correction, could further improve the accuracy and should be explored.

Although the current hybrid change detection method has proven to work reliably and efficiently, systematic testing with synthetic images that represent all possible combinations of camera settings and weather conditions, including the quantitative comparison with manually generated change maps, is needed to validate its applicability at a larger scale. In addition, some critical architectural decisions, such as the choice of different parameters (e.g., maximum tie point density or the raster cell size), need to be further investigated and possibly adapted. Also, it might be beneficial to find an alternative to the current rasterisation step because it decreases the achievable spatial granularity of the method. One could, for example, detect change by finding patterns of dissimilar feature points using advanced clustering algorithms like Density-Based Spatial Clustering (DBSCAN), thereby surpassing the rigid rasterisation approach.

Despite several possible improvements, this first proof-of-concept indicates the general applicability of the method and provides the basis for continued research on this topic. Hopefully, one day, such change detection approaches can facilitate largely automated monitoring and documentation of temporal change in cultural heritage in general and extensive graffiti-scapes in particular.

REFERENCES

Agisoft Metashape Version 1.8.3 Agisoft Metashape User Manual.Agisoft Metashape. 2022. Available online: https://www.agisoft.com/pdf/metashape-pro_1_8_en.pdf (accessed on 8 February 2023).

Alcantarilla, P.F., Nuevo, J., Bartoli, A., 2013: Fast explicit diffusion for accelerated features in nonlinear scale spaces. *Proceedings of British Machine Vision Conference 2013,* 13.1-13.11, doi.org/10.5244/C.27.13.

Angiati D., Gera, G, Piva, S., Regazzoni C.S., 2005: A novel method for graffiti detection using change detection algorithm. *Proc. IEEE Conf. Adv. Video Signal Based Surveillance*, 242-246, doi.org/10.1109/AVSS.2005.1577274.

Bay, H., Tuytelaars, T.,Van Gool, L., 2006: SURF: Speeded up robust features. *ECCV 2006. Lecture notes in computer science*, 3951, 404-417, doi.org/10.1007/11744023_32.
Bradski, G., 2000: The openCV library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*, 25(11), 120-123.

Bruzzone, L., Bovolo, F., 2013: A novel framework for the design of change-detection systems for very-high-resolution remote sensing images. *Proc. IEEE*, 101(3), 609-630, doi.org/10.1109/JPROC.2012.2197169.

Di Stefano, L., Tombari, F., Lanza, A., Mattoccia, S., Monti, S., 2008: Graffiti detection using two views. *Advanced Concepts for Intelligent Vision Systems (ACIVS 2008): 10th International Conference Proceedings*, 645-654.

Hotelling, H., 1936: Relations between two sets of variates. *Breakthroughs in statistics*, Springer, New York, NY.

ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume X-M-1-2023
29th CIPA Symposium "Documenting, Understanding, Preserving Cultural Heritage:
Humanities and Digital Technologies for Shaping the Future", 25–30 June 2023, Florence, Italy

Leutenegger, S., Chli, M., Siegwart, R.Y., 2011: BRISK: Binary Robust invariant scalable keypoints. *2011 International Conference on Computer Vision*, Barcelona, Spain, 2011, pp. 2548-2555, doi.org/10.1109/ICCV.2011.6126542.

Ligon, R.A., McGraw, K.J., 2013: Chameleons communicate with complex colour changes during contests: different body regions convey different information. *Biol. Lett.*, 9, 20130892, doi.org/10.1098/rsbl.2013.0892.

Liu, G., Gousseau, Y., Tupin, F., 2019: A Contrario Comparison of Local Descriptors for Change Detection in Very High Spatial Resolution Satellite Images of Urban Areas. *IEEE Trans. Geosci. Remote Sens.*, 57, 3904–3918, doi.org/10.1109/TGRS.2018.2888985.

Lowe, D.G., 1999: Object recognition from local scale-invariant features. *Proceedings of the seventh IEEE international conference on computer vision*, Vol. 2, 1150-1157, doi.org/10.1109/ICCV.1999.790410.

Lowe, D.G., 2004: Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vision*, 60, 91–110, doi.org/10.1023/B:VISI.0000029664.99615.94.

Luhmann, T., Fraser, C.S., Maas, H.G., 2016. Sensor modelling and camera calibration for close-range photogrammetry. *ISPRS Journal of Photogrammetry and Remote Sensing* 115, 37–46.

Masilamani, R., 2008. Documenting Illegal Art: Collaborative Software, Online Environments and New York City's 1970s and 1980s Graffiti Art Movement. *Art Doc. J. Art Libr. Soc. North Am*, 27 (2), 4-14, doi.org/10.1086/adx.27.2.27949489.

Molada-Tebar, A., Verhoeven, G., 2022. Towards colour-accurate documentation of anonymous expressions. Proceedings of the goINDIGO2022 international graffiti symposium. Document | archive | disseminate graffiti-scapes, Vienna, Austria. 11-13 May 2022. Urban Creativity, Lisbon.

Novak, D., 2014: Methodology for the Measurement of Graffiti Art Works: Focus on the Piece. *World Appl. Sci. J.*, 32, 40–46, doi.org/10.5829/idosi.wasj.2014.32.01.301.

Nielsen, A.A., Conradsen, K., Simpson, J.J., 1998: Multivariate alteration detection (MAD) in multispectral, bi-temporal image data: a new approach to change detection studies. Dept. Mathematical Modelling, Technical University of Denmark, Lyngby, Denmark, doi.org/10.1016/S0034-4257(97)00162-4.

Nielsen, A.A., 2007: The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data. *IEEE T. Image Process.*, 16(2), 463-478, doi.org/10.1109/TIP.2006.888195.

Pham, M.T., Mercier, G., Michel, J., 2016: Change Detection Between SAR Images Using a Pointwise Approach and Graph Theory. *IEEE T. Geosci. Remote.*, 54(4), 2020-2032, doi.org/10.1109/TGRS.2015.2493730.

Radke R.J., Andra, S., Al-Kofahi, O., Roysam, B., 2005: Image change detection algorithms: a systematic survey. *IEEE T. Image Process.*, 14 (3), 294-307, doi.org/10.1109/TIP.2004.838698.

Schlegel, J., Carloni, M., Wogrin, S., Graf, A., & Verhoeven, G. (2023). Making a Mark—Towards a Graffiti Thesaurus. Document | Archive | Disseminate Graffiti-Scapes,

203 - 219. https://doi.org/10.48619/indigo.v0i0.710

Seo, J., Park, W., Kim, T., 2022: Feature-Based Approach to Change Detection of Small Objects from High-Resolution Satellite Images. *Remote Sens.*, 14(3), 462,doi.org/10.3390/rs14030462.

Singh, A., 1989: Review article digital change detection techniques using remotely-sensed data. *Int. J. Rem. Sens.*, 10(6), 989-1003, doi.org/10.1080/01431168908903939.

Tareen, S.A.K., Saleem, Z., 2018: A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK. *Proc. IEEE Int. Conf. Comput. Math. Eng. Technol, 1-10*, doi.org/10.1109/ICOMET.2018.8346440.

Tombari, F., Di Stefano, L., Mattoccia, S., Zanetti, A., 2008: Graffiti Detection Using a Time-Of-Flight Camera. *Advanced Concepts for Intelligent Vision Systems (ACIVS 2008)*, Lecture Notes in Computer Science, 5259, Springer, Berlin, Heidelberg doi.org/10.1007/978-3-540-88458-3_58.

Tewkesbury, A.P., Comber, A.J., Tate, N. J., Lamb, A., Fisher, P.F., 2015: A critical synthesis of remotely sensed optical image change detection techniques. *Remote Sens. Environ.*, 160, 1-14, doi.org/10.1016/j.rse.2015.01.006.

Verhoeven, G., Wild, B., Schlegel, J., Wieser, M., Pfeifer, N., Wogrin, S., Eysn, L., Carloni, M., Koschiček-Krombholz, B., Molada-Tebar, A., Otepka-Schremmer, J., Ressl, C., Trognitz, M., Watzinger, A., 2022: Project INDIGO–Document, disseminate & analyse a graffiti-scape. *Int. Arch. Photogram. Remote Sens.*, 46, 513–520, doi.org/ 10.5194/isprs-archives-xlvi-2-w1-2022-513-2022.

Verhoeven, G., Wogrin, S., Schlegel, J., Wieser, M., & Wild, B. (2023). Facing a Chameleon—How Project INDIGO Discovers and Records New Graffiti. Document | Archive | Disseminate Graffiti-Scapes, 63 - 85. https://doi.org/10.48619/indigo.v0i0.703

Verhoeven, G., Carloni, M., Schlegel, J., Wild, B., & Wogrin, S. (2023). Finding Listeners for Walls that Speak. Document | Archive | Disseminate Graffiti-Scapes, 6 - 15. https://doi.org/10.48619/indigo.v0i0.699

Wang, Y., Du, L., Dai, H., 2016: Unsupervised SAR Image Change Detection Based on SIFT Keypoints and Region Information. *IEEE Geosci. Remote S.*, 13(7), 931-935, doi.org/10.1109/LGRS.2016.2554606.

Wild, B., Verhoeven, G., Wieser, M., Ressl, C., Schlegel, J., Wogrin, S., Otepka-Schremmer, J., Pfeifer, N., 2022: AUTOGRAF—AUTomated Orthorectification of GRAFfiti Photos. *Heritage*, 5(4), 2987-3009, doi.org/10.3390/heritage5040155.

Wild, B., Verhoeven, G., Wogrin, S., Wieser, M., Ressl, C., Otepka-Schremmer, J., & Pfeifer, N. (2023). Urban Creativity Meets Engineering. Document | Archive | Disseminate Graffiti-Scapes, 131 - 145. https://doi.org/10.48619/indigo.v0i0.705