Identification of spatially corresponding imagery using content-based image retrieval in the context of UAS video exploitation

Stefan Brüstle, Daniel Manger, Klaus Mück, Norbert Heinze
Fraunhofer Institute of Optronics, System Technologies and Image Exploitation (IOSB),
Fraunhoferstr. 1, 76131 Karlsruhe, Germany

ABSTRACT

For many tasks in the fields of reconnaissance and surveillance it is important to know the spatial location represented by the imagery to be exploited. A task involving the assessment of changes, e.g. the appearance or disappearance of an object of interest at a certain location, can typically not be accomplished without spatial location information associated with the imagery. Often, such georeferenced imagery is stored in an archive enabling the user to query for the data with respect to its spatial location. Thus, the user is able to effectively find spatially corresponding imagery to be used for change detection tasks.

In the field of exploitation of video taken from unmanned aerial systems (UAS), spatial location data is usually acquired using a GPS receiver, together with an INS device providing the sensor orientation, both integrated in the UAS. If during a flight valid GPS data becomes unavailable for a period of time, e.g. due to sensor malfunction, transmission problems or jamming, the imagery gathered during that time is not applicable for change detection tasks based merely on its georeference. Furthermore, GPS and INS inaccuracy together with a potentially poor knowledge of ground elevation can also render location information inapplicable.

On the other hand, change detection tasks can be hard to accomplish even if imagery is well georeferenced as a result of occlusions within the imagery, due to e.g. clouds or fog, or image artefacts, due to e.g. transmission problems. In these cases a merely georeference based approach to find spatially corresponding imagery can also be inapplicable.

In this paper, we present a search method based on the content of the images to find imagery spatially corresponding to given imagery independent from georeference quality. Using methods from content-based image retrieval, we build an image database which allows for querying even large imagery archives efficiently. We further evaluate the benefits of this method in the context of a video exploitation workflow on the basis of its integration into our video archive system.

Keywords: Airborne Intelligence, Surveillance, Reconnaissance, ISR, Content-based Image Retrieval, UAS.

1. INTRODUCTION

Imagery gathered from sensors mounted on aerial platforms is vital for many fields of application. In recent years operational costs of unmanned aircraft systems (UAS) have been decreasing and new sensors satisfying weight and size restrictions of even small UAS have been developed. These sensors provide high quality image sequences in all kinds of resolutions, spectral bands and frame rates. To be able to exploit these images effectively and efficiently, a human analyst is typically assisted today by some kind of image exploitation system. Many tasks such as e.g. activity recognition can be performed by such systems by automatically or interactively exploiting a single image or a number of subsequent frames from an image sequence. However, other tasks involving the evaluation of any kind of change that occurred within a certain time period require reference imagery to be included in the exploitation step. Imagery spatially corresponding to the imagery to be exploited, i.e. imagery covering the same spatial location but taken at a different time, fulfills this requirement, if different time refers to a time before the beginning of this time period. In order to be able to verify if imagery is spatially corresponding, it is necessary for this imagery to be stored with respect to its spatial location which in turn requires the image to be associated with some kind of spatial information. Most UAS have GPS and INS devices onboard which allow for automatic georeferencing of the gathered imagery. There are many approaches to retrieve georeferenced imagery from an archive, varying in their suitability for the archive size given and both in effectiveness and efficiency. Besides the implementation and evaluation of several other approaches, Fraunhofer IOSB has been developing a video archive system based on a spatial relational database which enables the image analyst to query for archived georeferenced imagery by specifying among other things a geographical search area. The corresponding query results are retrieved in an efficient way even with large archives. This video archive system integrates well with Fraunhofer IOSB's video exploitation system providing a complete workflow for change detection tasks. Both Fraunhofer IOSB's video archive system and video exploitation system have been described in depth in other papers, namely [9] and [8] and will be only introduced briefly in this paper in chapter 3. While this workflow usually works fine in the case of georeferenced reference imagery, sometimes the occurrence of occlusions within the imagery, due to e.g. clouds or fog, or image artefacts, due to e.g. transmission problems render the retrieved query results less well applicable for change detection tasks. In addition to that we also have to consider situations where imagery comes with a very poor georeference or even lacks georeference at least partially. These can occur due to poor telemetry data reception or complete (temporary) GPS or INS failure. To address these situations, other approaches have to be evaluated. In this paper we focus on a content based image retrieval approach to identify reference imagery.

2. CONTENT BASED IMAGE RETRIEVAL

The aim of content-based image retrieval systems is to compare images with respect to their content. To this end, local image regions are compared using local features such as the popular Scale-Invariant Feature Transform (SIFT) [2] which are used in many different topics of computer vision. They typically detect repeatable salient regions in an image and subsequently encode their local image appearance in a descriptor. Given the two sets of descriptors of two images, similar regions in both images can be searched by determining descriptors which are similar in descriptor space, which is for SIFT usually 128 dimensional. Typically, distances are calculated by L2 norm and a threshold is applied on the distance or on the ratio of closest to second closest distance. The similarity of two images is then often calculated as the number of matching features.

For matching sets of descriptors, various heuristic algorithms have been proposed which can lead to an impressive speedup while sacrificing not too much of the descriptors discriminance [3],[4]. Nevertheless, in large-scale content-based image retrieval (CBIR) systems with thousands or millions of images, a pair-wise image comparison of the query image with every image of the database becomes infeasible. Besides, the memory consumption of the image features and their processing during one query prohibit a direct matching of descriptors sets. To solve this, the bag-of-words (BOW) representation has been proposed [9], which quantizes the features by assigning every feature to one element of a set of feature representatives called visual words. Thus, the image matching can be performed with text retrieval methods analyzing the common visual words of images. The set of visual words termed codebook or visual vocabulary is commonly obtained by clustering an independent set of features. Using large codebooks, the representation of an image becomes a very sparse vector indicating the occurring visual words. This sparsity can be exploited by inverted files which store for every visual word a list of references to the images containing at least one feature corresponding to that visual word.

While enabling the construction of fast and efficient systems, the quantization of features also comes with the drawback of loss of information which leads to a reduced accuracy of the overall system. We use two popular extensions to circumvent the loss of accuracy namely Hamming Embedding and Weak Geometry Consistency [1]. Both techniques have shown a significant improvement of performance in large scale image retrieval. As rare visual words are assumed to be more discriminative, the similarity of two images given the two BOW vectors is commonly calculated using the tf-idf scheme [6]. It weights the BOW vectors according to both the local frequency (within the image) and the global frequency (within the entire database). In all experiments in this paper, we use the similarity function of [5] which is the cosine angle between the weighted BOW vectors which equals the L2 normalized dot product of the vectors. To further increase the performance, we make use of a subsequent re-ranking step, which performs a matching based on the original features. The images are re-ranked according to the number of direct matching features with the query image incorporating the local consistency of the matches. This means, that the similarity score is further increased if for two matching features there are still more consistent matches in the two local neighborhoods of the involved features. See [7] for a detailed description of the CBIR setup used in this work.

Sequence	**No. of images**	**Sequence length**	**No. of images filtered**	**Remaining images after filtering (compaction rate)**
Database sequences				
A	210k	2,3 hours	40k	19%
B	192k	2,1 hours	34k	18%
C	90k	1,0 hour	17k	19%
D	271k	3,0 hours	64k	24%
E	183k	2,0 hours	39k	21%
F	150k	1,7 hours	39k	26%
G	280k	3,1 hours	10k	4%
H	200k	2,2 hours	49k	25%
I	220k	2,4 hours	31k	14%
Sum DB	1,796k	20 hours	323k	18%
Query sequences				
AI	743	30 seconds	169	23%
BI	675	27 seconds	146	22%
BII	498	20 seconds	70	14%

Table 1: The nine chosen flights in the database, their durations and compaction rates.

3. FRAUNHOFER IOSB'S IMAGERY EXPLOITATION WORKFLOW: VIDEO ARCHIVE SYSTEM AND VIDEO EXPLOITATION SYSTEM

This chapter will give a brief overview of the capabilities of Fraunhofer IOSB's Video Archive System (VABUL) and their Video Exploitation System (ABUL). Internally, both systems work on single images, but as in the context of UAS, usually one individual image is taken per certain time interval, those single images constitute image sequences whose subsequent frames are spatially and temporarily related. Those image sequences are referred to as video in the context of VABUL and ABUL. ABUL was started as an integration frame for image enhancement and exploitation procedures and grew to a point where it can be used as a productive system. Besides its image enhancement capabilities it incorporates today many exploitation procedures working in real time on the video stream such as image stabilization, image carpet or mosaic formation, object detection and tracking and also change detection. Thereby real time working complies with the temporal limits given by the video streams. A detailed description of the system can be found in [8]. ABUL can work on live imagery as well as on archived imagery. Loading a certain image sequence from the archive into ABUL or selecting reference imagery for a task requiring reference imagery such as change detection tasks can be accomplished with VABUL. VABUL is a system consisting of four components: a spatial relational database, a server application representing an interface to that database, a search client and an import application. A spatial relational database is a relational database that can store geometries in addition to standard data types and offers a set of spatial methods and operations that can be executed upon the geometries. While the imagery itself is stored in the file system, a reference to the respective image file together with its spatial location information and a set of attributes is stored into the spatial database. The server application provides access to the database for the search client and the import application. The latter is used to import new imagery into the database, enabling the user to set values for attributes represented in the database such as mission name or platform and sensor information. The search client is used for composing queries onto the database. It is possible to define a geographical *search area* and again set values of represented attributes to define *geographical* and *non-geographical* restrictions respectively to limit the amount of query results. The query is carried out performing two steps. The first step consists of the retrieval of all single images which satisfy the defined restrictions. These single images represent the so called *raw query results*. The second so called *aggregation* step transforms these raw query results into result sequences by grouping several single result images of the same flight/mission together with respect to a set of aggregation *parameters*. Three parameters of this set, namely *separation*, *duration* and *maximum sequence overlap* (MSO) are specifically important in the context of this work. The objective of the aggregation step is to transform the single result images into as few as possible result sequences with respect to the aggregation parameters whose duration is accordingly as long as possible. The criterion which decides if and when to break one result sequence into two is the separation parameter: it defines the maximum acceptable temporal gap in which the given restrictions are allowed to be violated. In practice this approach results in longer and completer instead of short and many sequences. To illustrate how that parameter works, let's assume we have a sequence of 20 seconds in length. Furthermore, we assume that the frames of this sequence are only satisfying the geographical restrictions within the first 10 seconds and within the last 8 seconds. No we set the separation parameter for example to 1 sec. In our example this setting results in two sequences one of a length of 10 seconds and the other of a length of 8 seconds. Setting the separation parameter to 3 seconds will result in only one sequence of 20 seconds. In consequence this sequence will contain frames which violate the given restrictions. But such sequences with many and short gaps violating the given restrictions benefit from a better representation when visualized. Another important parameter, the duration parameter, acts as a filter: a result sequence only constitutes a valid result sequence if its length exceeds the value of the duration parameter.

The third parameter to be briefly introduced is the maximum sequence overlap (MSO). When the query is carried out, for each raw query result image, two areas are calculated: the geographical area it represents and its represented area lying within the search area. The quotient of the latter divided by the former constitutes the so called *sequence overlap*. During the aggregation step several raw query result images are grouped together to form a result sequence whose maximum sequence overlap is set to the value of the largest sequence overlap among the grouped query result images.

The resulting sequences are finally displayed in the result table of VABUL's search client application. A more detailed description of the system can be found in [9].

Figure 1: Four examples of the image retrieval results. The left images show the query frames, the corresponding images on the right show the respective most similar images from the database of other flights.

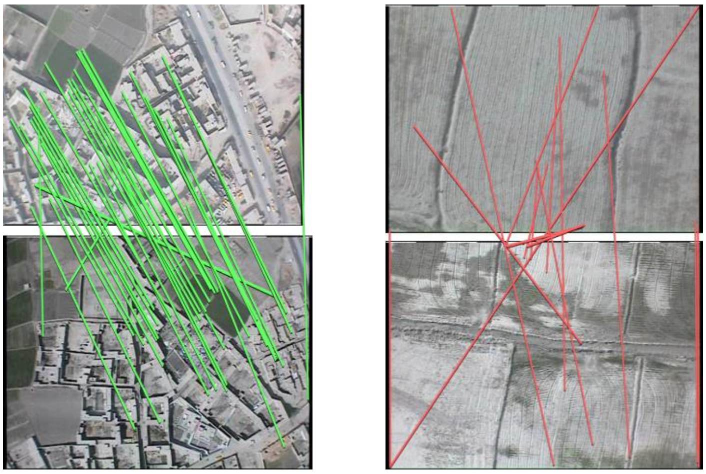

Figure 2: Visualization of the matching based on local features in the re-ranking step of the CBIR system. Left: the matching of the first image pair of figure 1 (query image is rotated for a better understanding of the visualization). Right: Example of a failure of the retrieval process: due to the repetitive structures in the fields, some false matches occurred.

4. CBIR APPROACH TO IDENTIFY REFERENCE IMAGERY

To address the problem, that images with poor or lacking georeference (in the following referred to as unreferenced images) can not be retrieved in a query applying a geographical restriction, and to benefit from the ability to evaluate a reference image according to its content, it seems promising to integrate CBIR into the VABUL search workflow. The two steps described in chapter 3, the retrieval of raw query results and the subsequent aggregation step can be extended by an additional CBIR step which identifies those images among all unreferenced images which correspond to a query image with respect to their contents. The higher such an image (in the following referred to as retrieved image) scores, the greater the probability is that it represents the same spatial location as its corresponding query image. Furthermore, if the threshold and re-ranking parameters are chosen suitably, a retrieved image can be assumed to satisfy the same geographical restriction as its corresponding query image without any actual knowledge about its spatial location. If performance is not an issue, all raw query result images can constitute query images and the CBIR step can be carried out between step one and three. However if performance is an issue, the amount of query images has to be reduced as CBIR processing time of course increases proportionally with the amount of query images. A reduction can be achieved either directly by only considering a subset of raw query result images or indirectly by performing the aggregation step twice. In the latter case initially only raw query result images are aggregated and after that the CBIR step is either performed on all or again only on a subset of images originating from either one or several selected result sequences. In both cases the (final) aggregation is performed both on raw query result images and retrieved images and thus following the above assumption on *all* imagery satisfying the geographical restriction independent of availability of georeference. Subsets of images can be determined in various ways, the one we chose is described in detail in chapter 5.1.

5. EVALUATION OF THE CBIR APPROACH

In order to be able to assess both the suitability and efficiency of the proposed CBIR approach we designed an experiment. It is described in depth in chapter 5.1, followed by the presentation of the assessment outcome in chapter 5.2.

5.1 Experiment description

Foremost we had to identify imagery from recorded surveillance flights satisfying two requirements: it had to be suitable for change detection tasks, which means images spatially overlapping other images had to be present, and there had to be a substantial amount of images to evaluate the efficiency of the CBIR approach. We identified imagery of 3 complete flights satisfying the requirements with several subsequences each and imported it into a test database using VABUL's import application. In addition to that we imported imagery of 6 more flights into the test database to increase the amount of data to be queried on. This resulted in a test database with 9 flights with durations between 1 and 4 hours. All flights were recorded with a standard definition sensor operating in the visual spectrum at a frame rate of 25 fps mounted to a small UAS flying in an altitude of approximately 600 meters. Due to frequent circling over particular locations within each flight several sequences overlap other sequences of the same flight. In addition to this inter-flight overlap but limited to a much smaller number in each flight there are sequences which overlap sequences of other flights. To ensure the availability of ground truth for the assessment, we only chose flights with accurate telemetry information resulting in well georeferenced imagery in the database. Due to the flight speed and altitude and the frame rate of 25 fps, two consecutive video frames often show nearly the same area on the ground. To keep the CBIR index compact, we therefore filter this redundant information. As the flight speed varies, instead of simply using every n^{th} frame we perform a filtering based on image contents. After calculating the 2D shift of images, we leave out images if they show an overlap in image area of more than 75% with any filtered image of the last second.

The CBIR system is created with all 9 flights listed in table 1 resulting (after filtering) in a database of some 323,000 images which led to 600 million features (on average, 1860 features per image). In the index, after quantization, these features are represented with 12 bytes each, thus, 7 GBs of main memory are used. We perform a re-ranking of the 200 most similar images retrieved in the first stage of the CBIR system which is a trade-of between performance and runtime. On a laptop with a i7-3720QM CPU, the first step of the CBIR search takes (without feature calculation) about 1.5 seconds, the re-ranking another 10 seconds for every query image since it involves accessing 200 feature files on the hard disk. After querying all images of a query sequence, the result similarity scores are aggregated for all images in the database. By applying a threshold on the aggregated scores, the list of results for this query sequence is generated. From all results of the query sequences, only three query images led to result images with scores above the threshold which do not show the same location in the image. These false positives arose from repetitive structures which confuse the local feature principle. See figure 2 for one of the three false positives.

To evaluate the suitability of the CBIR approach, the objective is to compare its capability to find spatially corresponding images with the capability of VABUL's search client to do so. This is done by comparing the result sequences retrieved with VABUL's search client and the result sequences retrieved by applying the CBIR approach. To make sure that the results would be comparable, we chose three evaluation sequences, AI, BI and BII within 2 out of the 9 flights between 20 and 30 seconds in length (view table 1). For the retrieval of spatially corresponding sequences using VABUL the minimum bounding rectangle for each evaluation sequence had to be calculated and set as search area for the respective query. For the retrieval of spatially corresponding sequences applying the CBIR approach the three evaluation sequences were filtered in the same way the 1.8 million images in the test database were filtered to build a compact CBIR index. This resulted in 743, 675 and 498 query images for the query sequences AI, BI and BII respectively that were in the following used for a query on the CBIR index. Finally the retrieved images were aggregated in the sense of VABUL's aggregation step (view chapter 3) to produce result sequences.

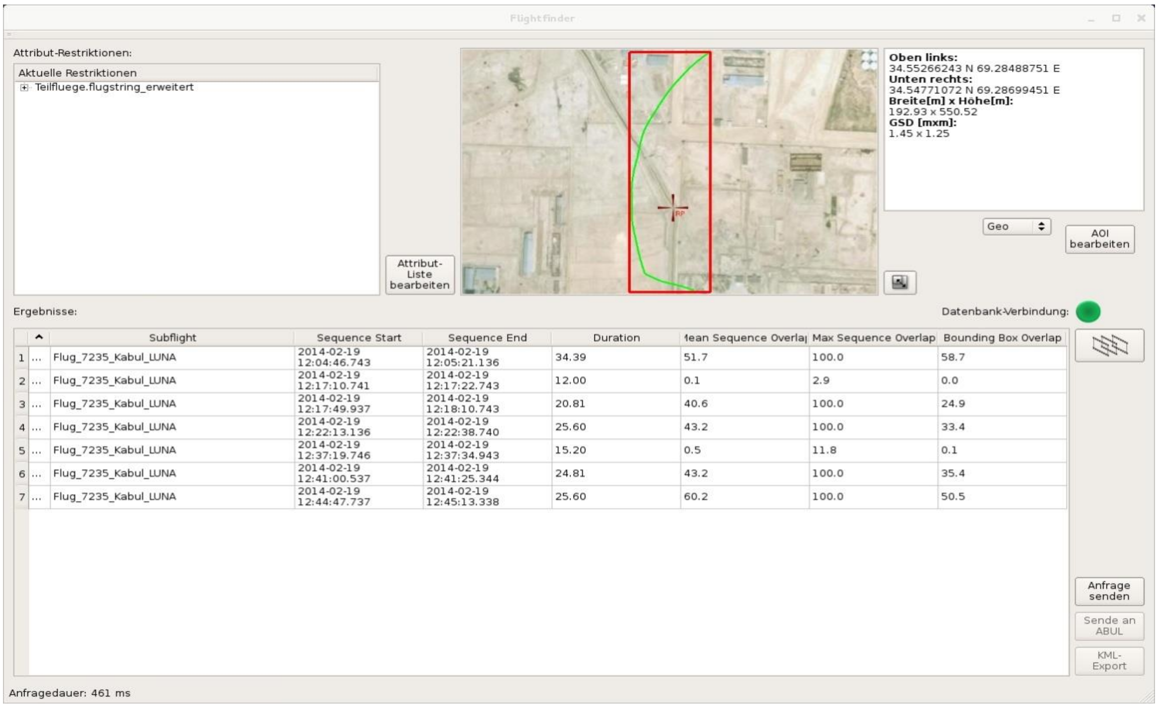

Figure 3: VABUL's search client with minimum bounding rectangle as search area.

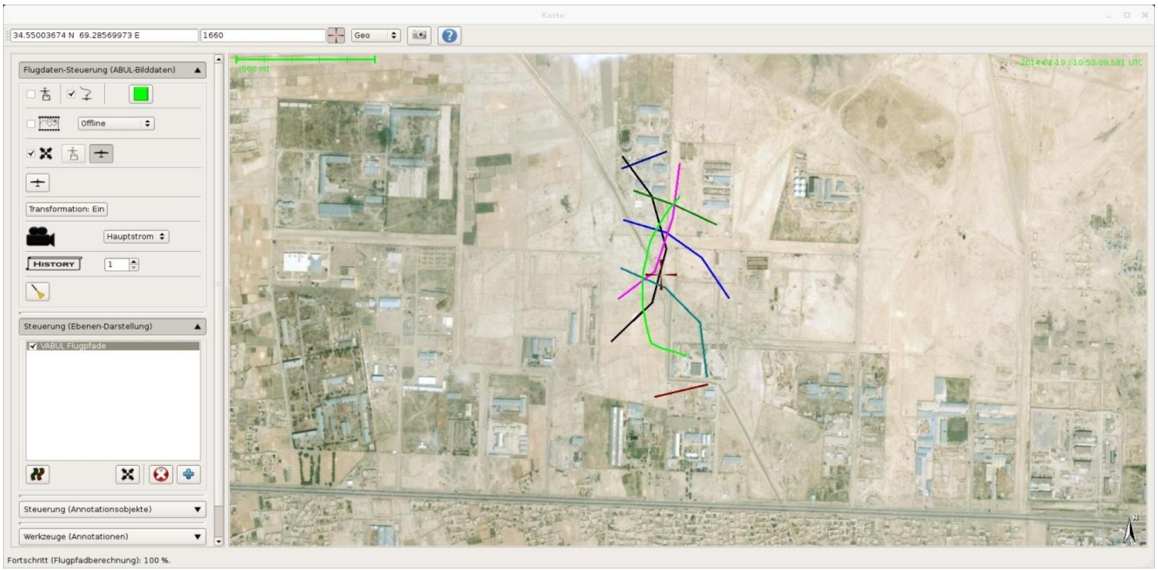

Figure 4: Query sequence (light green) and result sequences from figure 3 visualized on a reference map.

5.2 Results and evaluation

The tables 2 to 4 show the comparison results.

Query sequence	VABUL result sequence	Sequence length [sec]	MSO in VABUL result sequence [%]	Temporal overlap CBIR result sequence - VABUL result sequence [%]	Additional sequence lenth [%]
AI	B1	55.9	10.1	-	-
AI	B2	38.8	>50	43	-
AI	B3	8.3	2.2	-	-
AI	B4	33.3	>50	65	-
AI	B5	25.0	19.7 the first 6 sec, then >50	-	
AI	B6	55.1	2.0 the last 10 sec, else >50	17	-
AI	B7	10.0	>50	-	-
AI	B8	37.6	>50	56	23
AI	B9	15.8	>50	-	-

Table 2: Results for query sequence AI.

Query sequence	VABUL result sequence	Sequence length [sec]	MSO in VABUL result sequence [%]	Temporal overlap CBIR result sequence - VABUL result sequence [%]	Additional sequence lenth [%]
BI	A1	34.4	>50	14	-
BI	A2	12.0	2.9	-	-
BI	A3	20.8	>50	-	-
BI	A4	25.6	>50	-	-
BI	A5	15.2	11.8	-	-
BI	A6	24.0	>50	1	2
BI	A7	25.6	>50	58	33

Table 3: Results for query sequence BI.

Query sequence	VABUL result sequence	Sequence length [sec]	MSO in VABUL result sequence [%]	Temporal overlap CBIR result sequence - VABUL result sequence [%]	Additional sequence lenth [%]
BII	C1	50.2	33.8	35	-
BII	C2	11.1	1.1	-	-

Table 4: Results for query sequence BII.

The 3 query sequences AI, BI and BII were used to query the database both with VABUL and using the CBIR approach. This produced several result sequences for each approach. The *VABUL result sequence* column represents the result sequences retrieved with VABUL and the *Sequence Length [sec]* column their lengths. For each VABUL result sequence the corresponding result sequence for the CBIR approach had to be identified. If that was possible, the latter was compared to the former. The case where the CBIR result sequence constituted a subset of its corresponding VABUL result sequence resulted in an entry into the column *Temporal overlap CBIR result sequence – VABUL result sequence [%]* indicating the length of the CBIR result sequence as fraction of its corresponding VABUL result sequence. The other case, where the CBIR result sequence not only temporarily overlapped its corresponding VABUL result sequence but also temporarily exceeded and/or preceded it resulted in an entry into the column *Additional sequence length [%]* where the additional sequence length is again indicated as fraction of the corresponding VABUL result sequence length. If no corresponding result sequence for the CBIR approach could be identified this was indicated by the entry "-" in the column Temporal overlap rate CBIR *result sequence – VABUL result sequence [%]*. The column *MSO in VABUL result sequence [%]* with MSO standing for *Maximum Sequence Overlap* indicates the maximum sequence overlap as described in chapter 3 for each VABUL result sequence. As this value represents the spatial overlap between the geographical search area as defined in VABUL's search client and the particular image within the VABUL result sequence where this value has its maximum, a small value (<<50%) means that within the whole VABUL result sequence, there is no image exceeding this small spatial overlap. As the CBIR approach requires large parts of two images to be similar, the less two images spatially overlap the smaller the probability gets that the CBIR approach works on them. This explains why no corresponding CBIR result sequences could have been identified for the VABUL result sequences B1 and B3 in table 2, A2 and A5 in table 3 and C2 in table 4 with a maximum sequence overlap of 10.1%, 2.2%, 2.9%, 11.8% and 1.1% respectively. Therefore these VABUL result sequences weren't considered for further evaluation. In the case of query sequence AI for 4 out of 7 VABUL result sequences a corresponding CBIR result sequence could be identified accordingly. This equals an identification rate of 57%. The mean temporal overlap for those 4 CBIR result sequences is 42%, calculated by averaging their temporal overlaps with their corresponding VABUL result sequences weighted with the corresponding VABUL result sequences' lengths. For the query sequences BI and BII, identification was possible for 3 out of 5 and 1 out of 1 case equaling an identification rate of 60% and 100% respectively. The mean temporal overlaps where 24% and 35% respectively. Table 5 summarizes these numbers.

Query sequence	**Number of VABUL result sequences with a MSO >33%**	**Number of identified CBIR result sequences**	**Identification rate [%]**	**Mean temporal overlap CBIR result sequence - VABUL result sequence [%]**
AI	7	4	57	42
BI	5	3	60	24
BII	1	1	100	35

Table 5: Overall results of the three query sequences.

Identification rates of over 50% for all three query sequences demonstrate that the CBIR approach is generally capable of the identification of spatially corresponding imagery. Both those cases, where identification was still not possible and the fact, that the rates for mean temporal overlap didn't exceed 50% in cases where identification was possible are most likely a result of the way VABUL retrieves images for a given geographical search area as defined in its search client. While the CBIR approach requires an image to largely overlap the query image to be retrieved from a query with that query image, the criterion for VABUL for retrieving an image is intersection. Two images intersect if they overlap somewhere with the overlap being allowed to be as little as possible. That means that already a very small overlap, e.g. only a couple of pixels in the edge of the image, results in an image to be retrieved. Such an image however is not likely to be retrieved by a CBIR approach. If one wanted to bring the identification and the mean temporal overlap rate close to 100%, VABUL's intersection criterion would have to be changed. This however has not been the focus of this work.

Another effect that has an influence can be seen when the additional sequence lengths that occur in three cases (see table 2 and table 3) are considered. Additional sequence lengths are most likely the result of inaccurate georeference information associated with both the query sequence and the VABUL result sequence. This inaccuracy present in flights in our test database can result in the spatial location of particular images being as far as 50 meters away from their actual location. If a subsequence of a continuous flight is retrieved with respect to a search area using VABUL, it is therefore possible that certain parts of the sequence in fact lie outside the search area or that certain parts of the flight don't contribute to the result sequence as they seem to lie outside the search area due to their georeference whereas in fact they lie within the search area or at least intersect with it. The former case results in a false positive which in turn results in a worse temporal overlap, the latter case results in a false negative which in turn results in an additional sequence length.

To summarize, a medium value temporal overlap or even not identifying a particular CBIR result sequence doesn't mean that the CBIR approach isn't applicable. On the other hand identification of a CBIR result sequence and a high temporal overlap with the corresponding VABUL result sequence provide an additional quality criterion for deciding how suitable a VABUL result sequence might be for performing exploitation tasks involving change detection subtasks. In the case of many or even too many VABUL result sequences, this criterion could also be used to filter and rank these sequences.

SUMMARY/CONCLUSION

CBIR has proven to be applicable for reference imagery identification in the context of image sequence exploitation tasks such as change detection. If integrated into an archive system that provides methods to query imagery with respect to its geographical location, the system can benefit from CBIR in two ways: (1) its capability to make imagery with inaccurate or missing georeference information available for such queries and (2) its capability to provide an additional quality criterion to efficiently decide on the suitability of a query result for change detection tasks based on the content of the imagery.

REFERENCES

[1] Jegou, H., Douze, M., & Schmid, C. ,“ Hamming embedding and weak geometric consistency for large scale image search”, In European Conference on Computer Vision. ECCV 2008 (pp. 304-317),2008

[2] Lowe, D. G., “Distinctive image features from scale-invariant keypoints”, International journal of computer vision, 60(2), 91-110. Springer,2004

[3] Muja, M., & Lowe, D. G., “Fast approximate nearest neighbors with automatic algorithm configuration.”, In International Conference on Computer Vision Theory and Applications (VISSAPP’09),2009.

[4] Nister, D., & Stewenius, H., “Scalable Recognition with a Vocabulary Tree”, 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2 (CVPR'06), 2161-2168. doi: 10.1109/CVPR,2006

[5] Schmid, C., “Improving bag-of-features for large scale image search”, In International Journal of Computer Vision , 316-336,2011

[6] Sivic, J., & Zisserman, A., “Video Google: A text retrieval approach to object matching in videos.”, In Computer Vision,. Proceedings. Ninth IEEE International Conference on (pp. 1470-1477), 2003.

[7] D. Manger, “Large-Scale Tattoo Image Retrieval”, Canadian Conference on Computer and Robot Vision ,(2012).

[8] Heinze, N., Esswein, M., Krüger, W., Saur, G., “Image exploitation algorithms for reconnaissance and surveillance with UAV,” Proc. SPIE 7668, 76680 (2010).

[9] Brüstle, St., Heinze, N., “Archiving image sequence considering associated geographical and non-geographical attributes,” Proc. SPIE 8359, 83590Y (2012).