

Title	Mismatched image identification using histogram of loop closure error for feature-based optical mapping
Author(s)	Elibol, Armagan; Chong, Nak-Young; Shim, Hyunjung; Kim, Jinwhan; Gracias, Nuno; Garcia, Rafael
Citation	International Journal of Intelligent Robotics and Applications, 3(2): 196-206
Issue Date	2019-03-26
Type	Journal Article
Text version	author
URL	<a href="http://hdl.handle.net/10119/16228">http://hdl.handle.net/10119/16228</a>
Rights	This is the author-created version of Springer, Armagan Elibol, Nak-Young Chong, Hyunjung Shim, Jinwhan Kim, Nuno Gracias, Rafael Garcia, International Journal of Intelligent Robotics and Applications, 3(2), 2019, 196-206. The original publication is available at <a href="http://www.springerlink.com">www.springerlink.com</a> , <a href="http://dx.doi.org/10.1007/s41315-019-00089-0">http://dx.doi.org/10.1007/s41315-019-00089-0</a>
Description	

## Mismatched image identification using histogram of loop closure error for feature-based optical mapping

Armagan Elibol · Nak-Young Chong ·  
Hyunjung Shim · Jinwhan Kim · Nuno Gracias ·  
Rafael Garcia

Received: date / Accepted: date

**Abstract** Image registration is one of the most fundamental steps in optical mapping from mobile platforms. Lately, image registration is performed by detecting salient points in two images and matching their descriptors. Robust methods (such as Random Sample Consensus (RANSAC)) are employed to eliminate outliers and compute the geometric transformation between the coordinate frames of images, typically a homography when the images contain views of a flat area. However, the image registration pipeline can sometimes provide a sufficient number of wrong inliers within the error bounds even when images do not overlap at all. Such mismatches occur especially when the scene has repetitive texture and shows structural similarity. Such pairs prevent the trajectory (thus, a mosaic) from being estimated accurately. In this paper, we propose to utilize closed-loop constraints for identifying mismatches. Cycles appear when the camera revisits an area that was imaged before, which is a common practice especially for mapping purposes. The proposed method exploits the fact that images forming a cycle should have an identity mapping when all the homographies between images in the cycle are multiplied. Our proposal obtains error statistics for each matched image pair extracting several cycle bases. Then, by using a previously trained classifier, it identifies image pairs by comparing error histograms. We present experimental results with different image sequences.

---

A. Elibol and N.Y. Chong  
School of Information Science  
Japan Advanced Institute of Science and Technology, Ishikawa-Japan E-mail: aelibol@jaist.ac.jp

H. Shim  
School of Integrated Technology  
Yonsei University, Incheon, Rep. of Korea

J. Kim  
Dept. of Mechanical Engineering  
Korea Advanced Institute of Science and Technology, Daejeon-Rep. of Korea

N. Gracias and R. Garcia  
Computer Vision and Robotics Group  
University of Girona, Girona-Spain

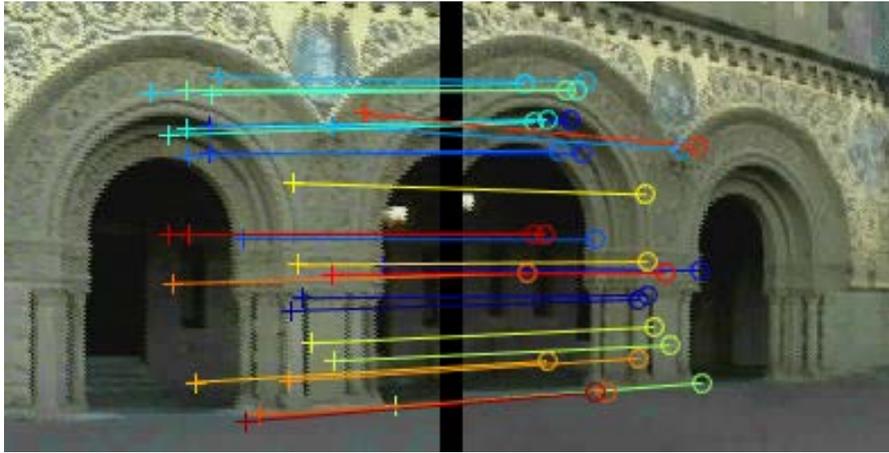
**Keywords** Mismatched identification · image matching · supervised classification.

## 1 Introduction

Obtaining image sequences for mapping has become an easier task over the last years, thanks to the rapid progress on optical sensors and robotic platforms. One of the steps in the processing pipeline of image sequences (e.g., mosaicing [3, 17] and 3D reconstruction [8]) is usually 2D image registration, which is the process of stitching two or more views of the same scene taken from different viewpoints. It is one of the crucial steps and plays a very important role in a vast amount of computer vision processes. Image registration is mainly accomplished by using either feature or featureless methods. Over the last decade, impressive progress (such as Scale Invariant Feature Transform (SIFT) [10], Speeded Up Robust Features (SURF) [1], etc.) has been made on detecting and extracting distinctive salient points (also known as features) in the image, which leads to foster and promote the usage of feature-based methods more than the featureless intensity based methods. Feature-based methods follow a pipeline that is composed of (1) feature detection, (2) feature description and (3) descriptor matching. The matching of features is commonly done by computing a function that depends on the Euclidean distance between their descriptors. The initial matching frequently produces some incorrect correspondences, which are called *outliers*. Outliers are typically identified and removed with a robust estimation algorithm (e.g., RANSAC [4]). These probabilistic methods might fail when the scene presents *Visual Aliasing*, that is highly repetitive textures and structural similarities. This failure often leads to mismatched image pairs (e.g., Figure 1), although the resulting homography accuracy is within the error bounds handled by RANSAC. While processing such a sequence of images with repetitive texture, the probability of occurrence of such cases becomes high and these mismatched image pairs provide misleading information about the camera trajectory and prevents an accurate final outcome.

In this paper, we present a method to identify such image pairs, using loop constraints. Our method relies on the fact that images forming a loop (a cycle in graph-based notation) should have identity mapping when all the homographies are multiplied (see Figure 3). If there is a mismatched image pair or pairs in a cycle, then the error (computed as a deviation from identity) will grow drastically. Errors computed by using a different number of cycles involving the same image pair can be used to indicate a false match positive for that image pair.

Our algorithm starts by finding cycles (loops) in the sequence. Next, errors are computed for each pair inside the cycles (see Table 1). To have meaningful descriptive statistics, this process is repeated several times with different cycles. Cycles are obtained by generating different cycles bases. After obtaining a series of errors for each image pair, we compute error histograms and try to identify the different ones using a previously trained classifier. To validate our proposal, we present experimental results with artificially added mismatched image pairs on underwater image sequences as well as real sequences from panoramic image sequences.



**Fig. 1** An example of mismatched image pair due to visual aliasing. Twenty-eight inliers are detected and the corresponding homography is computed using DLT [6] with an average error of 3.41 pixels in one image. Note that the image hardly overlap

Finding such mismatched pairs has been studied in the global alignment literature by iteratively removing the image pair with the highest residual error [11]. Although this might be feasible for small and/or sparser datasets, repeating the global alignment may lead to a high computational cost. The concept of "Missing correspondences" for an image pair using a third image was proposed in [18] to identify mismatched image pairs. A similar solution was proposed using this missing correspondences analysis and timestamps of images in [14]. This approach is only feasible for sequentially acquired image sequences due to the use of timestamps. Our method shares some similarities and builds upon the method proposed by Zach et al. [19]. The main differences are that we do not set any limit on the length (number of vertices) of a cycle and we do not use any empirical values for the parameters of the error distributions of image pairs. Instead, we employ a machine learning technique, operating on error histograms to identify mismatched image pairs.

The rest of the paper is structured as follows. The following section provides a brief background about graph-based topology representation and details our proposal. Experimental results are illustrated and discussed in Section 3. Finally, we present our conclusions in the last section.

## 2 Mismatch Identification using Cycles in Graph-based Topology Representation

A graph  $G = (V, E)$  consists of vertices (nodes) and edges (links) between vertices.  $V$  is the set of vertices while  $E \subset V \times V$  represents the set of edges. The total number of vertices  $n = |V|$  defines the order of the graph while the total number of edges  $m = |E|$  is the size of the graph [16]. A cycle in a graph is defined as a subgraph

in which every vertex has even degree, which is defined as the total number of edges incident to the vertex. A cycle basis is a list of cycles in the graph, with each cycle expressed as a list of vertices [7]. These cycles define a basis for the cycle space of the graph so that every other cycle in the graph can be obtained from the cycle basis using only symmetric differences, which is defined as the set of elements that are in either of the sets but not in their intersection. Commonly used methods to compute cycle basis are based on spanning trees. The spanning tree of a connected graph is a tree that connects all the vertices together [16]. One graph can have various different spanning trees. The Minimum Spanning Tree (MST) is a spanning tree whose edges have a total weight less than or equal to the total weight of every other spanning tree of the graph. The cycle basis obtained by using the MST is called the minimum weight spanning tree basis, which is computed by adding an edge to the MST and removing the path connecting the endpoints of the edge. The total number of cycles in the basis for a connected graph is  $t_c = (m - (n - 1))$ . If the graph is not connected<sup>1</sup> then total number of cycles in the basis is computed as  $(m - (n - 1) + (p - 1))$ , where  $p$  is the number of connected components. Every cycle that is not in the basis can be written as a linear combination of two or more cycles in the basis.

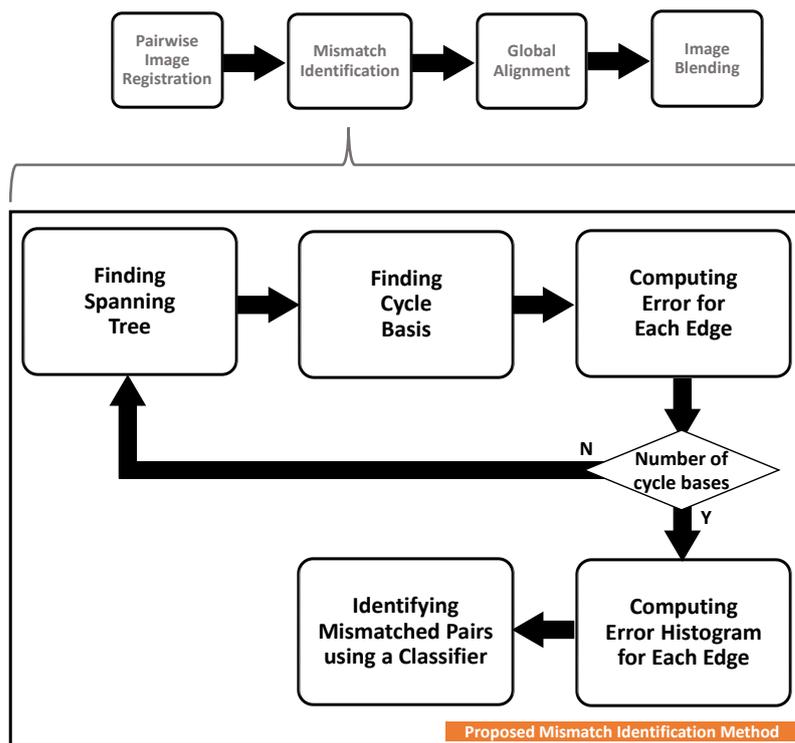
We use graph-based topology representation [15] for our problem where images are denoted as vertices and successfully registered images are connected with edges as this representation enables us to use some existing graph theory algorithms for finding cycles. In this work, having at least 20 correspondences after rejecting the outliers using RANSAC [4] is considered as a “successful registration”. As the outcome of the image registration process are two sets of coordinates of correspondences in the images, a planar transformation matrix  $\mathbf{H}$  (homography) can be computed for each edge in the graph.

Cycles represent closed-loops in our context. When the relative homographies for each edge in the cycle are multiplied consecutively, one could expect to obtain the identity mapping. In practice this does not hold due to error accumulation. For each edge in the cycle, an equation can be written by applying circular shift operations. Eq. 1 shows the possible equalities for the example cycle in Figure 3. These equations allow for computing errors for individual edges. If there is no wrong edge in the cycle, then all individual errors should be relatively small. However, if there is a wrong edge in the cycle, the error will increase drastically for all the edges in the cycle. This might be sufficient to identify cycles with wrong edges but might not be fully suitable to pinpoint which one of the edges is wrong in the cycle. On the other hand, one edge can appear in more than one cycle, e.g., the ones in the MST. And also, cycles can be composed of different number of edges. The longer the cycle, the more error it accumulates. This might cause some inaccurate interpretations of the magnitude of the errors on cycles and edges.

To address all aforementioned issues, we propose to use error statistics on edges. We generate several MSTs using randomly-generated edge weights. For each MST, we find the corresponding cycle basis. Then, for each edge in cycles in the basis, the error is computed by using circular shift as in the Eq. 1. The outcome of this process

---

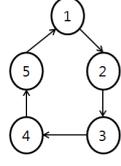
<sup>1</sup> In an undirected graph  $G$ , two vertices  $u$  and  $v$  are called connected if  $G$  contains a path from  $u$  to  $v$ . A graph is said to be connected if every pair of vertices in the graph is connected.



**Fig. 2** Pipeline of the proposed mismatched image pairs identification method embedded into a traditional FIM framework.

is a list of errors for each edge in the graph. We generate an error histogram for each edge using the list of errors. If there is no wrong edge in a cycle, the error is not likely to be as high as in the cases where there is a wrong edge. Therefore, correct edges might have small errors mostly and their errors must be varying in a relatively large interval due to the appearance of wrong edges. This leads to the conclusion that the shape of the error histogram for a correct edge might be seen similar to an exponential distribution. A real example of error histograms can be seen in Figure 5. Since the wrong edges are the main sources of error, regardless of the properties of the cycles they are involved in, they are more likely to have bigger errors. This results in the shape of the histogram being different than that of the exponential distribution. By comparing the shape of the histograms, wrong edges can be identified.

In order to automate this process, we use supervised classification using the AdaboostM1 algorithm [5] operating on error histograms. Our pipeline for identifying mismatched image pairs embedded in FIM framework is illustrated in Figure 2.



**Fig. 3** An illustrative example of a cycle showing a closed-loop. The loop should the following equality:  $\mathbf{I} = {}^1\mathbf{H}_2 \cdot {}^2\mathbf{H}_3 \cdot {}^3\mathbf{H}_4 \cdot {}^4\mathbf{H}_5 \cdot {}^5\mathbf{H}_1$ . However, this equality does not often hold due to error accumulation. The longer the loop, the more error it accumulates. When there is a wrong edge in a cycle (resulting from an image matching wrongly accepted as valid), then effect on the error is much large than the effect of accumulation.

$\mathbf{I} = {}^1\mathbf{H}_2 \cdot {}^2\mathbf{H}_3 \cdot {}^3\mathbf{H}_4 \cdot {}^4\mathbf{H}_5 \cdot {}^5\mathbf{H}_1$	$E(1, 2) = \ \mathbf{H}_1 - {}^2\mathbf{H}_3 \cdot {}^3\mathbf{H}_4 \cdot {}^4\mathbf{H}_5 \cdot {}^5\mathbf{H}_1\ _F$
$\mathbf{I} = {}^2\mathbf{H}_3 \cdot {}^3\mathbf{H}_4 \cdot {}^4\mathbf{H}_5 \cdot {}^5\mathbf{H}_1 \cdot {}^1\mathbf{H}_2$	$E(2, 3) = \ \mathbf{H}_2 - {}^3\mathbf{H}_4 \cdot {}^4\mathbf{H}_5 \cdot {}^5\mathbf{H}_1 \cdot {}^1\mathbf{H}_2\ _F$
$\mathbf{I} = {}^3\mathbf{H}_4 \cdot {}^4\mathbf{H}_5 \cdot {}^5\mathbf{H}_1 \cdot {}^1\mathbf{H}_2 \cdot {}^2\mathbf{H}_3$	$E(3, 4) = \ \mathbf{H}_3 - {}^4\mathbf{H}_5 \cdot {}^5\mathbf{H}_1 \cdot {}^1\mathbf{H}_2 \cdot {}^2\mathbf{H}_3\ _F$
$\mathbf{I} = {}^4\mathbf{H}_5 \cdot {}^5\mathbf{H}_1 \cdot {}^1\mathbf{H}_2 \cdot {}^2\mathbf{H}_3 \cdot {}^3\mathbf{H}_4$	$E(4, 5) = \ \mathbf{H}_4 - {}^5\mathbf{H}_1 \cdot {}^1\mathbf{H}_2 \cdot {}^2\mathbf{H}_3 \cdot {}^3\mathbf{H}_4\ _F$
$\mathbf{I} = {}^5\mathbf{H}_1 \cdot {}^1\mathbf{H}_2 \cdot {}^2\mathbf{H}_3 \cdot {}^3\mathbf{H}_4 \cdot {}^4\mathbf{H}_5$	$E(5, 1) = \ \mathbf{H}_5 - {}^1\mathbf{H}_2 \cdot {}^2\mathbf{H}_3 \cdot {}^3\mathbf{H}_4 \cdot {}^4\mathbf{H}_5\ _F$
where $\ A\ _F = \sqrt{\text{trace}(A^T A)}$	

(1)

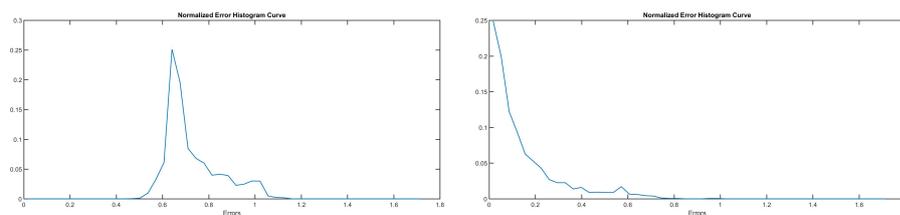
### 3 Experimental Results

To validate our proposal, we present experiments with different datasets from two main application domains of image mosaicing, namely panoramic imaging and optical mapping with unmanned platforms. For all experiments, we used 4-Degree-of-Freedom (DOF) similarity type planar transformations as they generally contain enough DOFs for optical mapping with robotic platforms. SIFT is used for feature detection and matching while RANSAC is employed for outlier rejection and transformation computation. Although we did some experiments with a different set of parameters, we used 250 randomly-generated cycle bases and 50 bins for error histograms. As a Global Alignment (GA) method, we use Symmetric Transfer Error Minimization (STEMin) over absolute homographies. The Symmetric Transfer Error (STE) is defined in Eq. 2:

$$\varepsilon = \sum_{k,t} \sum_{j=1}^s \left( \| {}^k\mathbf{x}_j - {}^1\mathbf{H}_k^{-1} \cdot {}^1\mathbf{H}_t \cdot {}^t\mathbf{x}_j \|_2 + \| {}^t\mathbf{x}_j - {}^1\mathbf{H}_t^{-1} \cdot {}^1\mathbf{H}_k \cdot {}^k\mathbf{x}_j \|_2 \right) \quad (2)$$

where  $k$  and  $t$  are image indices that were successfully matched,  $s$  is the total number of correspondences between the overlapping image pairs, and  $({}^1\mathbf{H}_k, {}^1\mathbf{H}_t)$  are the absolute transformation parameters for images  $k$  and  $t$ , respectively. This error term is minimizing the Euclidean distance between correspondences in each image frame. The first image frame is chosen as a global frame and fixed. To train a classifier, we used four different datasets obtained with unmanned underwater vehicles carrying a



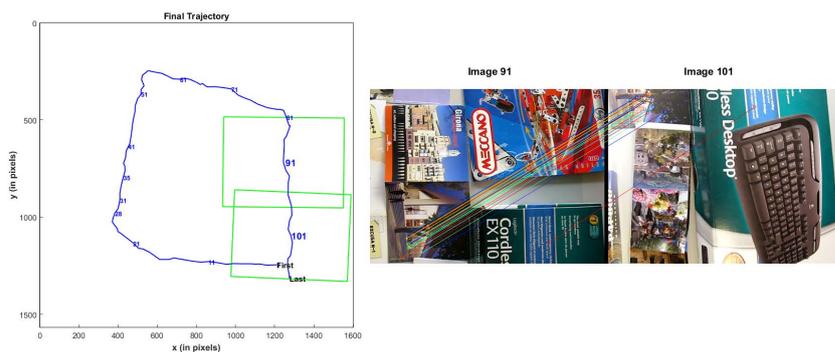


**Fig. 5** Sample Normalized Error Histogram Curves for Memorial Church Dataset. [Left] Error Histogram Curve for the edge between images 28 and 35. From the curve, it can be said that this pair is mismatched as its error histogram is mostly composed of large errors. This means error for this edge was large for all the cycles it involved in. [Right] Error Histogram Curve for the edge between images 1 and 2. From the curve, it can be concluded that this pair is a correct edge as it has mostly smaller errors.

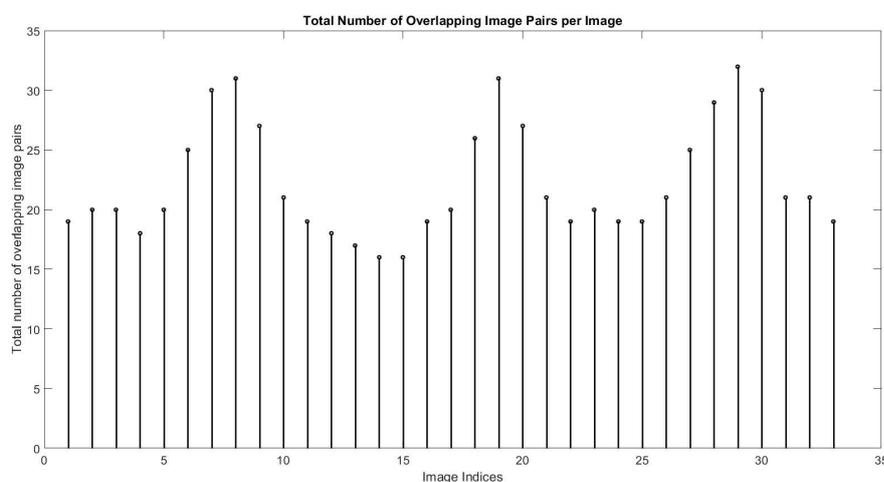
there might be some mismatched image pairs and/or some outliers among the correspondences. We run our proposal and four image pairs were identified as wrong edges. After removing these image pairs, we run STEMin again and mean error is 2.89, standard deviation is 1.08 and maximum error is found as 25.34 computed using 261,678 correspondences. Furthermore, we examined these 4 image pairs; one of them is a mismatched pair (illustrated in Figure 4) and the others have an overlapping area but they contain some outliers. The second dataset was acquired with a camera translating over a close range of a scene mostly composed of high color and texture containing books and boxes. The trajectory is a single loop-closing, drawing a smoothed rectangular shape. The dataset has 108 images of  $640 \times 480$  pixels with 1,827 overlapping image pairs. Mean STE computed over 638,357 is 8.14 pixels. Standard deviation is 4.55 while the maximum error is 1,263.67 pixels. Again such high maximum error suggests some outliers and/or mismatched image pairs. We run our proposal and eight overlapping image pairs were identified as mismatched. After removing those image pairs, we run STEMin. Mean STE computed over 638,093 was 8.10 pixels. Standard deviation has reduced to 3.67 pixels while maximum error was reduced to 69.46 pixels. We can infer that our proposal was able to remove the mismatched pairs and/or image pairs with outliers. For this dataset, we examined the removed 8 overlapping image pairs and there were no mismatched pairs but they were all containing some outlier(s). One of the removed image pairs is illustrated in Figure 6.

The third dataset is a video of the S. Zeno Cathedral in Verona<sup>3</sup> [12], similar to the first dataset. It is composed of 33 images of  $240 \times 282$  pixels having 368 overlapping image pairs. We run STEMin and mean error, standard deviation and maximum error are computed as 1.83, 0.71 and 7.08 respectively. We run our method to see whether there was any problematic image pairs or not. The classifier returned an empty set, which could be predicted by looking at the error values. Then we randomly added a few wrong edges to the dataset and run again our proposal. We repeated this procedure 100 times. Although we have seen from the previous experiments that the total number of problematic image pairs is usually low considering the total number of overlapping image pairs, we added wrong edges up to 8-10% of the total number of

<sup>3</sup> The dataset is obtained from [www.diegm.uniud.it/fusiello/demo/hrm/](http://www.diegm.uniud.it/fusiello/demo/hrm/)



**Fig. 6** One of the identified image pair for the second dataset. [Left] Figure shows the final trajectory and footprints of images 101 and 91 as rectangles. [Right] Images and correspondences. There is a single outlier in this case.



**Fig. 7** Overlapping image pair per image for the Szeno dataset. Please note that overlapping image pairs are counted for both images (e.g., The overlapping pair between image  $a$  and  $b$  is counted for both images.).

overlapping image pairs during random trials. Table 2 summarizes the obtained results in a confusion matrix form. The total number of added wrong edges is 1,365 while the total number of correct edges is  $100 \times 368$ . The total number of overlapping image pairs per image is illustrated in Figure 7.

From the confusion matrix, it can be seen that only one wrong edge was classified as correct by our proposal and 99 correct edges were also identified as wrong. In order to see the effects of this classification errors on trajectory, we run STEMIn on the remaining image pairs after removing the ones identified as wrong and computed STE on final trajectories obtained. The results are summarized in Table 3. Wrong edges were completely removed in the 99 trials out of 100. From the numbers in the

**Table 2** Confusion matrix for 100 random trials on the tested datasets.

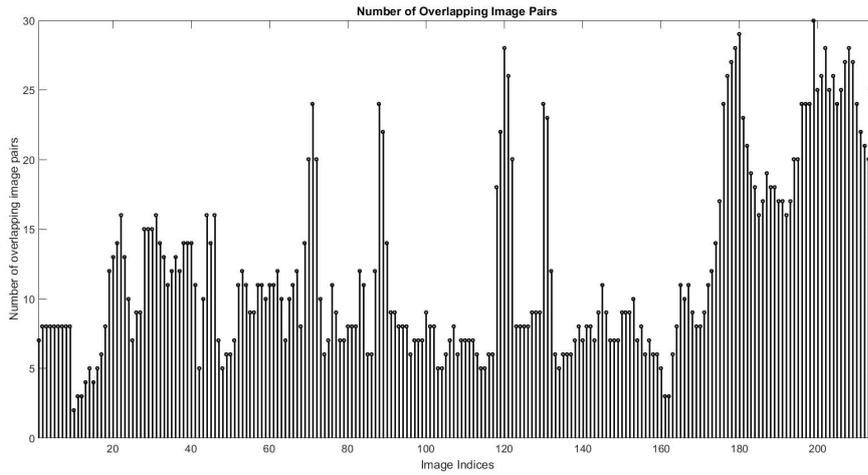
Dataset			Predicted		
			Wrong	Correct	Total
Szeno 33 Images 368 Overlapping pairs	Real	Wrong	1,364	1	1,365
		Correct	99	36,701	36,800
		Total	1,463	36,702	38,165
Underwater I 215 Images 1,297 Overlapping pairs	Real	Wrong	5,376	15	5,391
		Correct	245	129,455	129,700
		Total	5,621	129,470	135,091
Underwater II 486 Images 3,225 Overlapping pairs	Real	Wrong	12,487	15	12,502
		Correct	2,157	320,343	322,500
		Total	14,644	320,358	335,002

table, it can be inferred that removing a small amount of correct edges did not have much impact, as opposed to not removing the wrong edge(s). This is also related with the dataset as it can be regarded as pretty "dense", taking into account the total number of overlapping image pairs and images. Removing some correct edges did not have a notable effect. Our approach did not remove any correct edges on 45 trials out of 100, It removed only one correct edge on 32 trials, two edges on 9 trials, three edges on 10 trials, four edges on 2 trials, five and six edges on 1 trial.

**Table 3** Error measures (in pixels) on final trajectories over 100 trials for the tested datasets. Numbers given in the second and the third row are statistically computed over all trials.

Dataset		Mean Err.	Std. Dev.	Max.Err.
Szeno (33 images) (37,976 correspondences)	Initial Dataset	1.83	0.71	7.08
	Wrong Edges Added	40.41	15.06	192.58
	Wrong Edges Removed	2.02	0.81	25.67
Underwater I (215 images) (208,346 correspondences)	Initial Dataset	6.69	2.78	31.06
	Wrong Edges Added	44.35	27.73	2,302.02
	Wrong Edges Removed	6.84	3.61	1,724.44
Underwater II (486 images) (360,262 correspondences)	Initial Dataset	8.05	3.21	41.84
	Wrong Edges Added	68.24	37.94	975.87
	Wrong Edges Removed	8.08	3.44	387.34

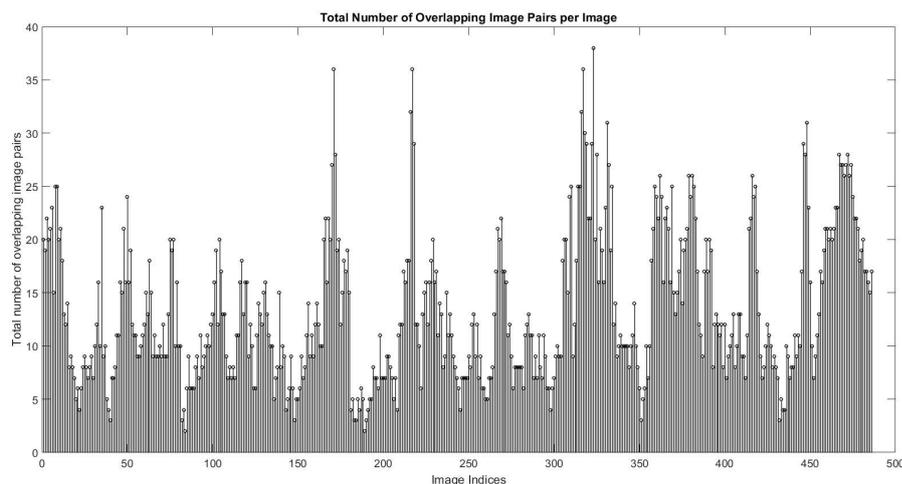
The fourth dataset (named as Underwater I) is a subset of an image sequence gathered by the *ICTINEU*<sup>AUV</sup> [13] underwater robot. The dataset is composed of 215 images ( $384 \times 288$  pixels) and covers approximately  $400m^2$ . Time-consecutive images have overlap and the total number of registered overlapping image pairs number is 1,297. This dataset is sparser and has low overlap between image pairs, when comparing to the Szeno Dataset. The confusion matrix is summarized for 100 ran-



**Fig. 8** Overlapping image pair per image in the Underwater I dataset. Please note that overlapping image pairs are counted for both images (*e.g.*, the overlapping pair between image  $a$  and  $b$  is counted for both images.). Some of the images have very few overlapping image pairs. This reduces the total number of cycles in which they could be involved. Depending on the wrong edges, these images might always be in the cycles along with wrong edges. This could lead them to be identified as wrong edges.

dom trials in Table 2 while the errors on the obtained trajectories are given in Table 3. Since the time-consecutive images have overlap in this dataset, we did not remove them even if our proposal suggested to do so. The approach proposed in this paper successfully removed the wrong edges in 92 out of 100 trials. The total number of wrong edges added for the remaining 8 cases (where our proposal was not able to identify all wrong edges) is 543, and 523 of them were successfully removed.

Our proposal relies on having cycles with and without wrong edges so that the different error histograms can be observed between correct and wrong edges. In other words, if a correct edge always involves cycles where there is at least one wrong edge, then this correct edge is most likely to be classified as wrong since its histogram will be very similar to those of wrong edges. This may not be a problem for dense datasets as the number of overlapping image pairs per image is relatively high (*e.g.*, Szeno dataset, Figure 7) and this increases the probability of having cycles without wrong edges. In the case of sparse datasets, this assumption of having cycles without wrong edges may not be realistic. In such cases, the use of additional information on image positions coming from different sensors could be useful. Increasing the total number of cycle bases could also be a solution, as it increases the probability of having cycles without wrong edges. However, the data characteristics may still not allow this assumption to be hold. For the Underwater I dataset, the total number of overlapping image pairs per image is illustrated in Figure 8. As it can be seen, some images overlap with a reduced number of neighboring images. This may rule out the initial assumption depending on the randomly added wrong edges.



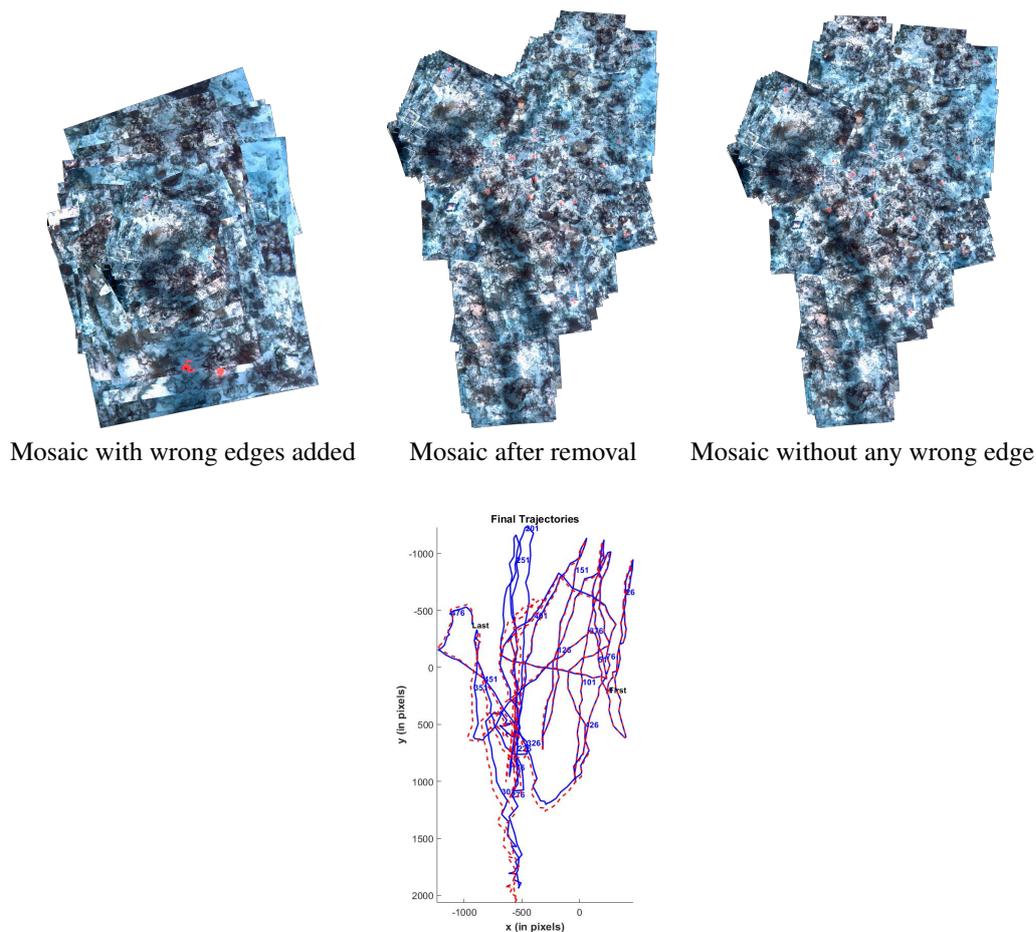
**Fig. 9** Overlapping image pair per image in the Underwater II Dataset. Please note that overlapping image pairs are counted for both images (e.g., the overlapping pair between image  $a$  and  $b$  is counted for both images.).

The fifth dataset (named as Underwater II) was obtained using a digital camera carried by a Phantom XTL Unmanned Underwater Vehicle (UUV) during a survey of a patch of reef located in the Florida Reef Tract [9]. It consists of 486 images of  $512 \times 384$  pixels. The total number of overlapping image pairs is 3,225. There are 6 time-consecutive image pairs that do not have any overlap. We kept the overlapping time-consecutive images even if they were identified as a wrong edge.

The confusion matrix is summarized for 100 random trials in Table 2 while the errors on the obtained trajectories are given in Table 3. The proposed method was not able to successfully identify all wrong edges in 5 trials out of 100. The total number of wrong edges added for those 5 trials were 898 and the proposed method was able to remove 890. From the confusion matrix, it can be seen that the total number of correct edges identified as wrong by our proposal can be considered as relatively high. This is mainly due to the way in which the classifier is trained. We opt to put high penalization for classifying wrong edges as correct since even one of them is usually enough to disturb the final result. To illustrate the performance of the method, the final mosaics and trajectories for one of the random trials (with the maximum STE error of 387.34) are illustrated in Figure 10. In this case our proposal was not able to remove all wrong edges. The total number of overlapping image pairs per image is illustrated in Figure 9.

## 4 Conclusions

Image registration is generally done by extracting and matching some salient points (features) in the images. Although some robust methods are employed, some mismatched image pairs can occur due to the repeated textures and high structural simi-



**Fig. 10** [Top] Final Mosaics for Underwater II dataset rendered using last-on-top strategy. [Bottom] Final Trajectories: Solid line denotes the trajectory without any wrong edge while the dashed one shows the trajectory after removing wrong edges. A total number of 214 wrong edges were added and 212 of them were successfully removed. The remaining two (between images 201 – 168 and 202 – 131) of them has caused some degradation on the trajectory and mosaic, respectively.

larity in images. These mismatched pairs can cause significant degradation on the final result (e.g., trajectory, mosaic, 3D Reconstruction, etc.). In this paper, we propose a method to identify mismatched image pairs using close-loop constraints (cycles) in camera trajectory. When the motion between images in a closed-loop is accumulated, the combined motion should be close to identity.

Our proposal relies on generating different cycles and computing error histograms for each overlapping image pairs. Mismatched image pairs have different error histogram compared to the those of correct edges. To identify mismatched image pairs, we employ a machine learning technique operating on error histograms. We presented

experimental results with both real and synthetic datasets under different scenarios. Our proposal was able to remove wrong edges successfully and also able to identify edges not fully mismatched but including some outliers. Our method can be embedded in any application that requires processing of image sequence.

**Acknowledgements** This research was partially supported by the Spanish National Project UDRONE (Robot submarino inteligente para la exploracion omnidireccional e inmersiva del bentos) CTM2017-83075-R, the EU VII Framework Programme as part of the Eurofleets2 (grant number FP7-INF-2012-312762) and by the MIST(Ministry of Science and ICT), Rep.of Korea, under the ICT Consilience Creative Program (IITP-2018-2017-0-01015) supervised by the IITP(Institute of Information & Communications Technology Planning & Evaluation)

## References

1. Bay, H., Tuytelaars, T., Van Gool, L.J.: SURF: Speeded up robust features. In: European Conference on Computer Vision, pp. 404–417. Graz, Austria (2006)
2. Davis, J.: Mosaics of scenes with moving objects. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. I, pp. 354–360. Santa Barbara, CA, USA (1998)
3. Fang, X., Zhu, J., Luo, B.: Image mosaic with relaxed motion. *Signal, Image and Video Processing* **6**(4), 647–667 (2010)
4. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM* **24**(6), 381–395 (1981)
5. Freund, Y., Schapire, R., et al.: Experiments with a new boosting algorithm. In: ICML, vol. 96, pp. 148–156 (1996)
6. Hartley, R., Zisserman, A.: *Multiple View Geometry in Computer Vision*, second edn. Cambridge University Press, Harlow, UK (2004)
7. Kavitha, T., Liebchen, C., Mehlhorn, K., Michail, D., Rizzi, R., Ueckerdt, T., Zweig, K.: Cycle bases in graphs characterization, algorithms, complexity, and applications. *Computer Science Review* **3**(4), 199–243 (2009)
8. Koutsoudis, A., Vidmar, B., Ioannakis, G., Arnaoutoglou, F., Pavlidis, G., Chamzas, C.: Multi-image 3d reconstruction data evaluation. *Journal of Cultural Heritage* **15**(1), 73 – 79 (2014)
9. Lirman, D., Gracias, N., Gintert, B., Gleason, A., Reid, R.P., Negahdaripour, S., Kramer, P.: Development and application of a video–mosaic survey technology to document the status of coral reef communities. *Environmental Monitoring and Assessment* **159**, 59–73 (2007)
10. Lowe, D.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* **60**(2), 91–110 (2004)
11. Martinec, D., Pajdla, T.: Robust rotation and translation estimation in multiview reconstruction. In: IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR’07., pp. 1–8 (2007)
12. Marzotto, R., Fusiello, A., Murino, V.: High resolution video mosaicing with global alignment. In: IEEE Conference on Computer Vision and Pattern Recognition, vol. I, pp. 692–698. Washington, DC, USA (2004)
13. Ribas, D., Palomeras, N., Ridao, P., Carreras, M., Hernandez, E.: ICTINEU AUV wins the first SAUCE competition. In: IEEE International Conference on Robotics and Automation. Roma, Italy (2007)
14. Roberts, R., Sinha, S., Szeliski, R., Steedly, D.: Structure from motion for scenes with large duplicate structures. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3137–3144. IEEE (2011)
15. Sawhney, H., Hsu, S., Kumar, R.: Robust video mosaicing through topology inference and local to global alignment. In: European Conference on Computer Vision, vol. II, pp. 103–119. Freiburg, Germany (1998)
16. Skiena, S.: *Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica*. Addison-Wesley, Reading, MA (1990)
17. Xu, Z.: Consistent image alignment for video mosaicing. *Signal, Image and Video Processing* **7**(1), 129–135 (2013). DOI 10.1007/s11760-011-0212-1. URL <http://dx.doi.org/10.1007/s11760-011-0212-1>

- 
18. Zach, C., Irschara, A., Bischof, H.: What can missing correspondences tell us about 3d structure and motion? In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–8 (2008)
  19. Zach, C., Klopschitz, M., Pollefeys, M.: Disambiguating visual relations using loop constraints. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1426–1433 (2010)