

# PHOTOMATCH: AN OPEN-SOURCE MULTI-VIEW AND MULTI-MODAL FEATURE MATCHING TOOL FOR PHOTOGRAMMETRIC APPLICATIONS

D. González-Aguilera <sup>1</sup>, E. Ruiz de Oña <sup>1</sup>, L. López-Fernandez <sup>1</sup>, E. M. Farella <sup>2</sup>, E.K. Stathopoulou <sup>2</sup>, I. Toschi <sup>2</sup>,  
F. Remondino <sup>2</sup>, P. Rodríguez-Gonzálvez <sup>3</sup>, D. Hernández-López <sup>4</sup>, A. Fusiello <sup>5</sup>, F. Nex <sup>6</sup>

<sup>1</sup> Cartographic and Terrain Engineering Department, Higher Polytechnic School of Avila, University of Salamanca, Avila, Spain - (daguilera, estebanrdo, luisloez89)@usal.es

<sup>2</sup> 3D Optical Metrology (3DOM) unit, Bruno Kessler Foundation (FBK), Trento, Italy - Email: (elifarella, estathopoulou, toshi, remondino)@fbk.eu

<sup>3</sup> Department of Mining Technology, Topography and Structures, Universidad de León, Ponferrada, Spain - p.rodriguez@unileon.es

<sup>4</sup> Institute for Regional Development (IDR), University of Castilla-La Mancha, Albacete, Spain - david.hernandez@uclm.es

<sup>5</sup> DPIA, University of Udine, Udine, Italy - andrea.fusiello@uniud.it

<sup>6</sup> Department of Earth Observation Science, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands - f.nex@utwente.nl

## Commission V, WG V/8

**KEY WORDS:** Photogrammetry, Tie points, Matching, Detectors, Descriptors, Open-source, Software development

### ABSTRACT:

Automatic feature matching is a crucial step in Structure-from-Motion (SfM) applications for 3D reconstruction purposes. From an historical perspective we can say now that SIFT was the enabling technology that made SfM a successful and fully automated pipeline. SIFT was the ancestor of a wealth of detector/descriptor methods that are now available. Various research activities have tried to benchmark detector/descriptors operators, but a clear outcome is difficult to be drawn.

This paper presents an ISPRS Scientific Initiative aimed at providing the community with an educational open-source tool (called PhotoMatch) for tie point extractions and image matching. Several enhancement and decolorization methods can be initially applied to an image dataset in order to improve the successive feature extraction steps. Then different detector/descriptor combinations are possible, coupled with different matching strategies and quality control metrics. Examples and results show the implemented functionality of PhotoMatch which has also a tutorial for shortly explaining the implemented methods.

## 1. INTRODUCTION

The photogrammetric problem of 3D reconstruction from multiple images has received a lot of attention in the last decade, especially focused on its two main pillars: (i) image orientation and self-calibration and (ii) dense matching reconstruction. However, the overall performance of both steps strongly depends on the quality of the initial feature (*keypoints*) extraction and matching stage. Therefore, determining which feature detectors and descriptors offer the most discriminative power and the best matching performance is of significant interest to a large part of the photogrammetry and computer vision communities. Methods for performing these tasks are usually based on representing an image using some global or local image properties and comparing them using a similarity measure or some machine/deep learning approaches. Nevertheless, most of the existing methods are designed for matching images within the same modality and under similar geometric conditions.

### 1.1 Aims of the work

The contributions are multifold:

- Develop an open-source educational tool, named PhotoMatch, that encloses different state-of-the-art algorithms for tie point extraction, including different detectors and descriptors as well as matching strategies;
- Improve the computational cost exploiting GPU and parallel computing, including CUDA programming capabilities;
- Assess the results of tie point extraction from a quantitative point of view using some statistical and robust parameters;

- Release the tool under GitHub in C++ and QT languages to allow people for further contributions;
- Prove the applicability of the developed tool with various datasets (aerial oblique, terrestrial, drone, multi-modal);
- Provide a tutorial and manual to describe the implemented methods.

### 1.2 The PhotoMatch project

With the aim of providing a contribution in the context of tie point extraction, an open-source feature extraction and matching tool, called PhotoMatch, has been developed. PhotoMatch encloses and combines different state-of-the-art detectors and descriptors, together with different matching strategies. PhotoMatch allows to solve feature extraction and matching steps with special focus on precision, reliability and flexibility. PhotoMatch is also an educational tool that allows the user to test and combine different detectors and descriptors, as well as to assess the precision and reliability of the results obtained.

The project, supported as an ISPRS Scientific Initiative, was led and managed by USAL in collaboration with UCLM, UNILEON, FBK, TWENTE and UDINE universities, aimed to develop an open-source tool for the image pre-processing, feature extraction and matching and system evaluation, including also an educational tutorial. The project was successfully built upon a multidisciplinary and international team with experience in image analysis, photogrammetry and computer vision in order to design and develop this feature matching tool.

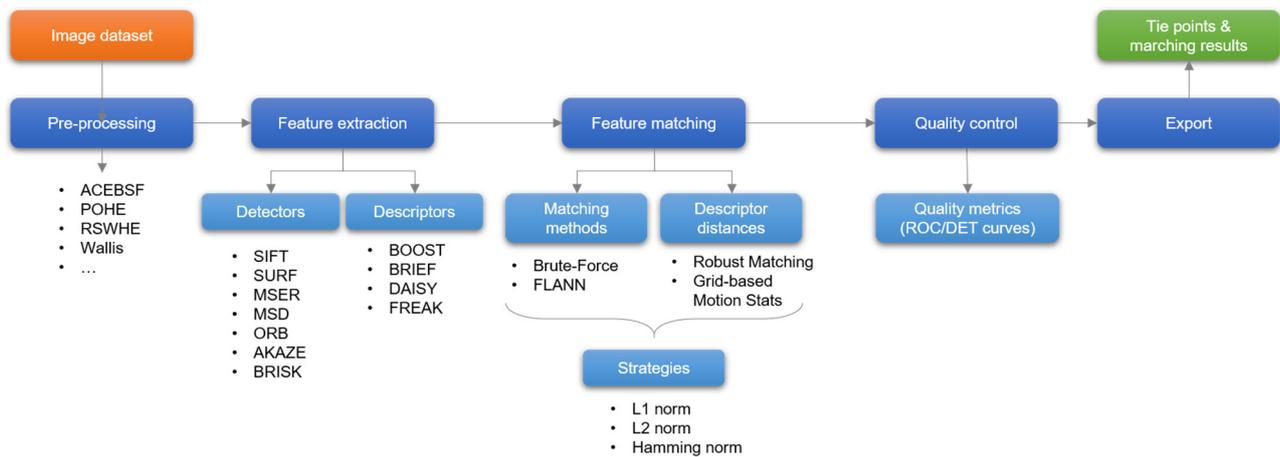


Figure 1. The overall workflow of the PhotoMatch tool.

## 2. PHOTOMATCH

### 2.2 Feature extraction

The developed open-source tool encloses a pipeline divided in 6 main steps applied sequentially to the set of loaded images (Figure 1):

1. Project/session definition: it allows to process the same dataset with different algorithms and/or parameters, and compare the achieved results.
2. Pre-processing: different enhancement and decolorization methods are available to improve the successive feature extraction steps.
3. Feature extraction and description: many detectors and descriptors algorithms are included (e.g. SIFT, SURF, etc.) and users can run tests modifying all necessary parameters and combinations to extract and describe keypoints in the images.
4. Feature matching: once keypoints are identified in two or more images, they are matched using different matching strategies (brute force, FLANN, etc.).
5. Quality control: feature matching results are evaluated using several options and metrics.
6. Export: PhotoMatch allows to export tie points and matching results in formats compatible with most of the common photogrammetric software in order to run a bundle adjustment and derive the orientation parameters.

In the same project, each of these steps can be repeated several times, allowing the assessment and comparison of different algorithms and parameters. More details on the different options are given in the following sections.

#### 2.1 Pre-processing

PhotoMatch allows to pre-process the input images in order to improve their radiometric content and support the successive feature extraction. The image pre-processing has been reported in many papers as a fundamental step, in particular in those cases where the texture quality is unfavourable (Aicardi et al., 2016; Gaiani et al., 2017; Jende et al., 2018).

Different pre-processing algorithms are available in PhotoMatch, including among others: ACEBSF (Lal and Chandra., 2014), POHE (Liu et al., 2013), RSWHE (Kim and Chung., 2008), Wallis (Wallis, 1974), etc. This step is optional but highly suggested in order to achieve better results in the subsequent feature extraction step. Note however that feature detectors are typically invariant to certain radiometric transformations, so not all the pre-processing algorithms that improve the visual perception have an impact on the extraction stage.

Many photogrammetric and computer vision tasks rely on feature extraction as primary input for further processing and analysis, including point matching, image registration, object detection, etc. Matching features have the following characteristics (Haralick and Shapiro, 1992): distinctness (clearly distinguished from the background), invariance (independent from radiometric and geometric distortions), stability (robustness against image noise), interpretability (the associated interest values should have a meaning and possibly usable for further operations) and uniqueness (distinguishable from other points).

Feature extraction consists on the identification of several meaningful features in the images, depicting a salient and distinctive part of the object scene seen in an image. Good features differ from other pixels as they have specific radiometric properties that make them distinctive and therefore re-detectable in different images with automated procedures. Image features can be categorized into corners, blobs and edges and their extraction consists in two consecutive steps: feature detection and description.

Detectors are operators which search for 2D locations in the images (i.e. a keypoint or a region) that are geometrically stable under different transformations and containing high information content.

On the other hand, descriptors analyse the surrounding of the detected feature (e.g. a keypoint) and produce a 2D vector of information. This information can be used to quickly classify the extracted points or in a matching process. Descriptors can be generally divided into floating and binary, according to the type of information stored in the vector.

Several extraction and detection algorithms have been proposed in the last decades in order to reliably detect features among images with geometric and radiometric transformations. However, many extreme operative conditions (e.g. multi-modal or multi-temporal images, wide baseline, etc.) still represent a challenge for most of the existing algorithms.

PhotoMatch implements diverse sets of detectors (e.g. SIFT (Lowe, 2004), SURF (Bay et al. 2006), MSER (Matas et al., 2004), MSD (Tombari and Di Stefano, 2014), ORB (Rublee et al., 2011), AKAZE (Alcantarilla et al., 2013), BRISK (Leutenegger et al., 2011), etc.) and descriptors (e.g. BOOST (Trzcinski et al., 2013), BRIEF (Calonder et al., 2011), DAISY (Tola et al., 2010), FREAK (Alahi et al., 2012), etc.) algorithms to let the user run and test different combinations and assess the results in different conditions. Any kind of combination is allowed in the software. For each algorithm, several advanced parameters can be defined by the user.

The implemented algorithms belong to the more traditional features category, as they are defined in advance and not learnt from the images. In the current implementation of PhotoMatch, modern feature descriptors embedding deep learning algorithms (Zbontar and Le Cun, 2015; Ono et al., 2018; Christiansen et al., 2019) are not implemented yet. As reported in the literature, these methods need different trainings according to the typology of processed data, therefore they are still difficult to be generalized to any type of scenarios, and to be handled in an educational software.

### 2.3 Feature matching

Once keypoints are identified in two or more images, they need to be matched among the images in order to find a set of correspondences or tie points. PhotoMatch contains different matching methods (Brute-Force and FLANN) and strategies (Robust Matching-RM and Grid-based Motion Statistics-GMS), and different typologies of descriptor distances (e.g. L1 or L2 norm and Hamming norm). For each strategy, a different set of parameters can be defined by the user to test the results. In order to validate the matches and remove outliers, homography (H) or Fundamental (F) matrix computation can be used as relative orientation backbone. The robust filtering can be performed using different statistical methods: threshold values can be also set to assess the sensitivity of the achieved results.

### 2.4 Quality control and export

PhotoMatch includes several options for validating and analysing the feature matching results. Quality assessment can be checked based on different approaches:

- manually defining a ground truth within the tool or importing an external one from an input file;
- computing H and F matrices transformation;
- analyzing different quality metrics such as repeatability, ROC/DET curves (Receiver Operating Characteristic/Detection Error Tradeoff) to measure precision and recall of the retrieved correspondences.

Thanks to a developed GUI, tie points can be directly checked and edited on the images to have a better understanding of the algorithm performance.

Last but not least, PhotoMatch allows to export the extracted tie points and matching results in format compatible with most of the common photogrammetric and SfM software in order to run a bundle adjustment and derive camera parameters.

### 2.5 Educational tutorial

PhotoMatch includes also short descriptions of the algorithms implemented in every step, i.e. pre-processing, feature extraction, feature matching and quality control. The tutorial has a dual purpose:

- to give users an overview of the algorithm functionalities and facilitate optimal combinations and parameter selection based on the specific needs of each project;
- to serve as an educational tool for non-expert users with respect to other black-box solutions.

Each algorithm has a short description in the Help page of PhotoMatch and relevant references are also provided to allow a deeper understanding of the used methods.

## 3. EXPERIMENTAL RESULTS

The implemented algorithms were applied to various multi-view (Section 3.1) and multi-modal (Section 3.2) datasets to demonstrate the potentials of our tool, with special focus on flexibility and rigorousness. Although various detector/descriptors and matching functions were combined and tested, only few combinations are afterwards reported. Note that PhotoMatch allows us to create different sessions in order to compare and analyse the different combinations between detector/descriptor and matching.

### 3.1 Multi-view multi-platform datasets

Figure 2 shows the main GUI of PhotoMatch, with the available combinations of detectors/descriptors and results on a UAV dataset.

Figure 3-4a report results on a set of 9 nadir and oblique images provided by the ISPRS/EuroSDR's Benchmark on High Density Image Matching for DSM Computation (Nex et al., 2015).

In order to improve the feature extraction, all input images were pre-processed based on the same algorithm: Recursively Separated and Weighted Histogram Equalization-RSWHE (Kim et al., 2008), since RSWHE preserves the image brightness more accurately and produces images with better contrast enhancement.

All employed detectors were limited in the maximum number of keypoints (5,000) and the used matching strategy was robust matching supported by RANSAC using the F matrix as geometric test (Gonzalez-Aguilera et al., 2018). All the computations were performed exploiting parallel and GPU capabilities of the hardware.

Table 3.1 shows the evaluation results for the different combinations (detector/descriptor) considered in the processing. The combinations were based on the following aspects:

- (i) prioritize those detectors with affine invariant performance such as SIFT, MSER, BRISK, MSD;
- (ii) use those detectors that incorporate their own descriptor and those which are invariant to rotation and scale such as BRIEF, BRISK and SIFT.

Figure 4b shows some quality analyses with the ROC curves for the extracted correspondences in the Dortmund dataset.

According to the achieved results on the aerial oblique Dortmund dataset, the following aspects can be highlighted:

- SIFT+SIFT (detector+descriptor) provides the best results in terms of number of matchings, as well as the true positive matching rate. However, its efficiency decreases considerably when it is combined with BRIEF and BRISK descriptors.
- The BRISK detector shows a good performance with its own descriptor, but when it is combined with SIFT descriptor its performance is even better.
- MSD detector is less efficient in extracting correspondences with respect to SIFT and BRISK detectors and the best results of MSD are those obtained when combining it with the BRISK descriptor.
- MSER detector is underperforming with respect to SIFT and BRISK detectors and its best results are achieved by combining it with SIFT and BRISK descriptors.
- BRIEF descriptor, in all our datasets and case studies, is delivering the worst results in terms of extracted correspondences.
- In some cases, BRISK+SIFT has improved considerably the results obtained by SIFT+SIFT.

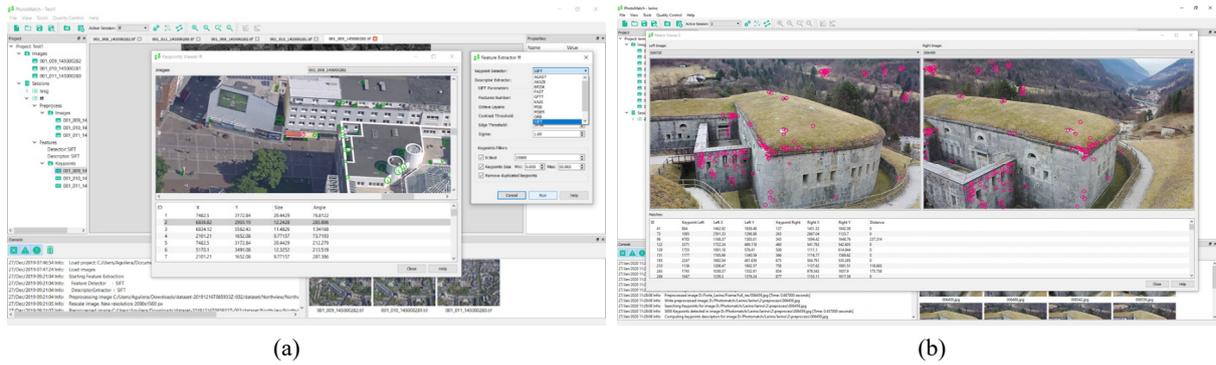


Figure 2: Different keypoint detectors and descriptors available in PhotoMatch (a) and results of UAV image pair matching (b).

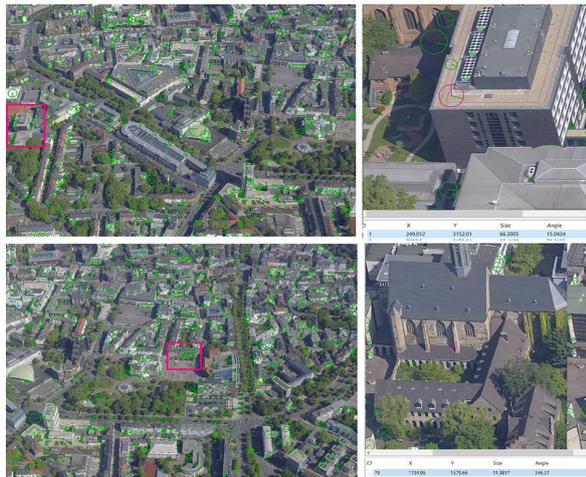
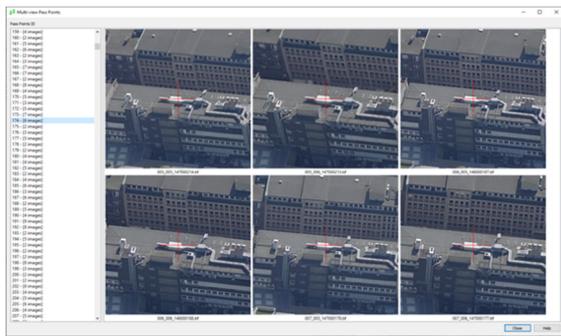


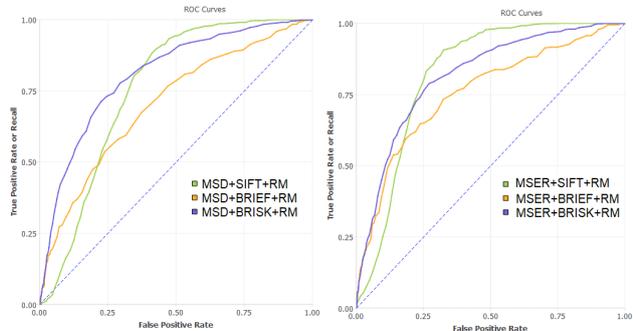
Figure 3: A view of the extracted correspondences between two images.

Detector	Descriptor	Avg. numb. of features per image	Avg. numb. of matches per image
SIFT	SIFT	4,732	1,344
SIFT	BRIEF	4,803	1,658
SIFT	BRISK	4,856	2,453
BRISK	BRISK	5,000	3,453
BRISK	SIFT	5,000	1,834
BRISK	BRIEF	4,756	3,052
MSD	SIFT	5,000	1,127
MSD	BRIEF	4,702	1,607
MSD	BRISK	4,827	632
MSER	SIFT	5,000	414
MSER	BRIEF	4,744	420
MSER	BRISK	4,509	728

Table 1: Various combinations of detectors/descriptors for the Dortmund dataset (with a limit of 5,000 features per image).



(a)



(b)

Figure 4: Multi-view visualization of the matching results (a) and ROC curves for the results presented in the Table 1 (b).

Figure 5 and Table 2 present some further tests performed on a multi-view dataset of 5 convergent and rotated terrestrial images of the main façade of the Modena Cathedral (Italy). Different combinations of feature detectors and descriptors were tested to assess the algorithm performances when rotations are present. Table 2 presents the total number of matches extracted with a robust matching strategy and adopting 9 detector and descriptor combinations.

In this case, the SURF-SURF combination returned the highest number of extracted correspondences, outperforming in identifying matches among differently tilted images (Figure 5).

Detector	Descriptor	Tot. numb. correct matches
SIFT	SIFT	450
SIFT	SURF	91
AKAZE	AKAZE	151
FAST	FREAK	148
SURF	SURF	539
SURF	SIFT	401
ORB	ORB	188
BRISK	BRISK	320
KAZE	KAZE	232

Table 2: Various combinations of detectors/descriptors for the terrestrial dataset of Figure 5.



(c) Figure 5: Examples of the features extraction based on SURF-SURF (a) and matching results (b and c) on the multi-view Modena Cathedral dataset.

### 3.2 Multi-modal datasets

A set of 10 thermographic and visible images (Figure 6) over an urban area were captured with a manned ultra-light motor (Lopez et al., 2015). In order to co-register the multi-modal dataset finding homologues points, a specific detector/descriptor combination together with a differential adaptation of the detector parameters was used. More specifically, the MSD detector combined with the SIFT descriptor was used, considering different salience thresholds ( $S$ ) and number of selected points ( $KNN$ ) which overpass the salience threshold for visible and thermographic images (Table 3). The different tests carried out demonstrate that both parameters yield remarkable performance on multi-modal images, turning out the best

parameters to be setup. The remaining parameters of MSD detector and SIFT descriptor were considered as suggested in the original implementation. On the other hand, it was performed a robust matching (RM) function supported by RANSAC estimator with different distance thresholds,  $D$ , and filtering coefficients,  $k$ , since both parameters were considered important in multimodal matching.

In order to check the results, the F matrix defined through the precise and reliable identification of a set of 12 well-distributed homologous points was used.

The results of the application of MSD+SIFT+RM are illustrated in Figure 7 and Table 3.

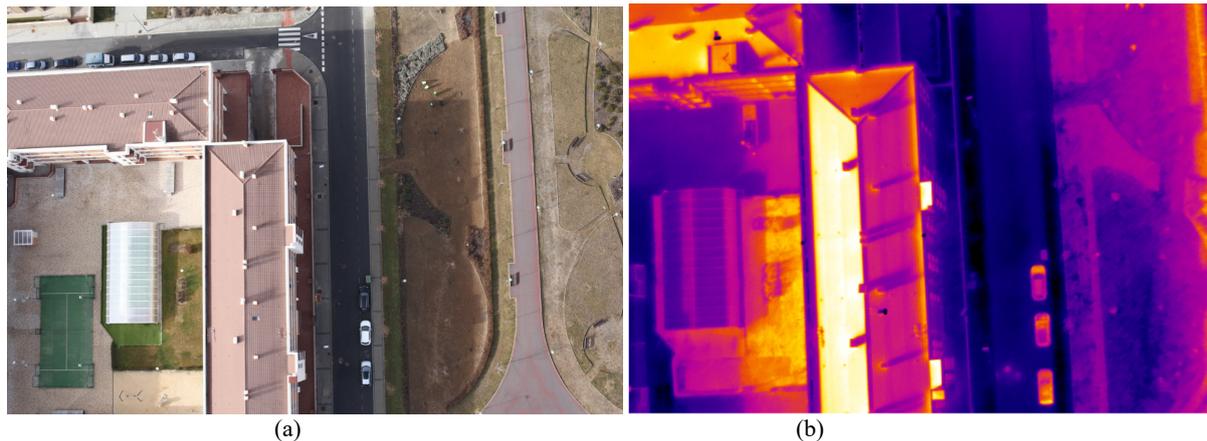


Figure 6: An example of visible (a) and thermal (b) images acquired with a manned ultra-light motor and that need to be automatically co-registered by finding homologues points.

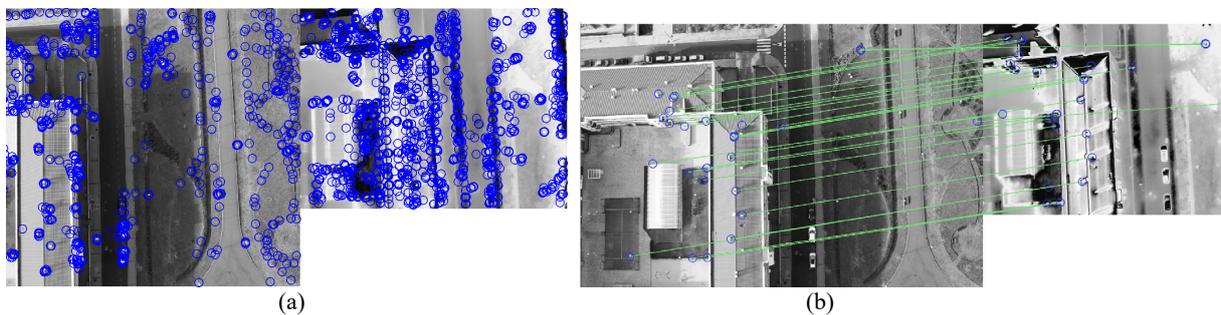


Figure 7: An example of the feature extraction and matching on the multi-modal dataset: feature extraction based on MSD+SIFT (a) and robust matching results using RANSAC (b).

		TRIPLET	$k=1$	$k=0.8$
		MSD+SIFT	$m_c/m_t$	$m_c/m_t$
$KNN=1$	RANSAC	$D=1$	80.0%	<b>84.2%</b>
		$D=3$	64.4%	61.5%
		$D=10$	61.6%	59.7%
		TRIPLET	$k=1$	$k=0.8$
		MSD+SIFT	$m_c/m_t$	$m_c/m_t$
$KNN=3$	RANSAC	$D=1$	80.0%	79.3%
		$D=3$	67.3%	77.6%
		$D=10$	66.3%	66.7%

Table 3. Different parameters analysed for the triplet detector/descriptor/matcher (MSD+SIFT+RM) in the multi-modal dataset.  $m_c$  and  $m_t$  refer to correct matchings and total matching, respectively. In bold the combination which provides the best efficiency.

According to the results presented in Table 3, it is worth to note that the saliency threshold  $S$  is related to the level of dissimilarity between neighboring pixels, i.e. the way a keypoint is different. So, it will be essential to make a different consideration of this threshold in visible ( $S=650$ ) and thermographic ( $S=65$ ) images, being higher in visible images (more demanding in dissimilarity) and lower in thermographic images (less demanding in dissimilarity).  $KNN$  indicates the minimum number of saliency points considered. If it was 3, it would keep the three points that have the highest saliency, i.e. more different. The parameter  $D$  indicates the orthogonal distance to the epipolar line in pixels.  $k$

is a weight factor based on the Norm-L2 distance: if it is 1, it will keep all the points and it is 0.8 will stay only with the 80% of the points prioritized by their lowest L2 distance.

#### 4. CONCLUSIONS

The paper documents the results of a ISPRS Scientific Initiative, led and managed by USAL in collaboration with UCLM, UNILEON, FBK, TWENTE and UDINE universities, aimed to develop an open-source tool for the image pre-processing, feature extraction and matching and system evaluation, including also an educational tutorial. The output is PhotoMatch tool, an open-source (<https://github.com/TIDOP-USAL/photomatch/releases>) educational tool that encloses different state-of-the-art algorithms for tie point extraction, including different detectors and descriptors as well as matching strategies. Extracted correspondences can be exported in various formats in order to launch a bundle adjustment with other tools. PhotoMatch features GPU and parallel computing, including CUDA programming capabilities. It offers various metrics to evaluate the matching results, including manually defined ground truth or ROC/DET curves. An educational tutorial and manual are also available in order to explain the implemented methods.

Some preliminary tests have been performed considering the available on-line benchmarks. Different tests on airborne, terrestrial and multi-modal (RGB-thermal) datasets have been performed, showing the performance of different combinations of algorithms and parameters. The performed tests have shown how the combination of different detector and descriptors can deliver higher accuracies in specific situations.

PhotoMatch could be further extended and improved, e.g. by adding other operators based on machine and deep learning approaches and especially focusing on multi-modal datasets.

PhotoMatch was developed with an educational approach in mind, nevertheless its GPU and parallel computing capabilities allow to quickly process the datasets. The secret of success has been to find a multidisciplinary and international team with experience in image analysis, photogrammetry and computer vision in order to design and develop this feature matching tool.

#### ACKNOWLEDGEMENTS

This work was supported by ISPRS through the 2019 Scientific Initiative entitled “Development of an open source multi-view and multimodal feature marching tool for photogrammetric applications”.

#### REFERENCES

- Aicardi, I., Nex, F.C., Gerke, M., Lingua, A.M., 2016. An image-based approach for the co-registration of multi-temporal UAV image datasets. *Remote Sensing*, 8(9), 779.
- Alahi, A., Ortiz, R., Vandergheynst, P., 2012. Freak: Fast retina keypoint. *IEEE Conference on Computer Vision and Pattern Recognition*, 510-517.
- Alcantarilla, P.F., Nuevo, J., Bartoli, A., 2013. Fast explicit diffusion for accelerated features in nonlinear scale spaces. *Proceedings British Machine Vision Conference*, 34(7), 1281–1298.
- Bay, H., Tuytelaars, T., Gool, L.V., 2006. SURF: Speeded-Up Robust Features. *European Conference on Computer Vision*, 404-417.
- Calonder, M., Lepetit, V., Ozuysal, M., Trzcinski, T., Strecha, C., Fua, P., 2011. BRIEF: Computing a local binary descriptor very fast. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(7), 1281-1298.
- Christiansen, P.H., Kragh, M.F., Brodskiy, Y., Karstoft, H., 2019. UnsuperPoint: End-to-end Unsupervised Interest Point Detector and Descriptor. *arXiv 1907.04011v1*.
- Gaiani, M., Apollonio, F.I., Ballabeni, A., Remondino, F., 2017. Securing color fidelity in 3D architectural heritage. *Sensors*, 17(11), 2437.
- González-Aguilera, D., López-Fernández, L., Rodríguez-González, P., Hernández-López, D., Guerrero, D., Remondino, F., Menna, F., Nocerino, E., Toschi, I., Ballabeni, A., Gaiani, M., 2018. GRAPHOS – open-source software for photogrammetric applications. *The Photogrammetric Record*, 33(161), 11–29.
- Haralick, R.M. Shapiro, L.G., 1992: *Computer and Robot Vision*, Addison-Wesley, 630 pp.
- Jende, P., Nex, F.C., Gerke, M., Vosselman, G., 2018. A fully automatic approach to register mobile mapping and airborne imagery to support the correction of platform trajectories in GNSS-denied urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 141, 86-99.
- Kim, M., Chung, M.G., 2008. Recursively separated and weighted histogram equalization for brightness preservation and contrast enhancement. *IEEE Transactions on Consumer Electronics*, 54(3), 1389-1397.
- Lal, S., Chandra, M., 2014. Efficient algorithm for contrast enhancement of natural images. *International Arab Journal of Information Technology*, 11(1), 95-102.
- Leutenegger, S., Chli, M., Siegwart, R., 2011. BRISK: Binary robust invariant scalable keypoints. *International Conference on Computer Vision*, 2548–2555.
- Liu, Y.F., Guo, J.M., Lai, B.S., Lee, J.D., 2013. High efficient contrast enhancement using parametric approximation. *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2444–2448.
- López, L., Lagüela, S., Picon, I., González-Aguilera, D., 2015. Automatic analysis and classification of the roof surfaces for the installation of solar panels using a multi-data source and multi-sensor aerial platform. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-5/W4, 171–178.
- Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60 (2), 91–110.
- Matas, J., Chum, O., Urban, M. Pajdla, T., 2004. Robust wide-baseline stereo from maximally stable extremal regions. *Image and vision computing*, 22(10), 761-767.
- Nex, F., Remondino, F., Gerke, M., Przybilla, H. J., Bäumker, M., Zurhorst, A., 2015. ISPRS Benchmark for Multi-Platform Photogrammetry. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences*, II-3/W4, 135-142.
- Ono, Y., Trulls, E., Fua, P., Yi, K.M., 2018. LF-Net: Learning Local Features from Images. *32nd Conference on Neural Information Processing Systems*, 6237-6247.
- Rublee, E., Rabaud, V., Konolige, K., Bradski, G.R., 2011. ORB: An efficient alternative to SIFT or SURF. *International Conference on Computer Vision*, 2564–2571.
- Tombari, F., Di Stefano, L., 2014. Interest Points via Maximal Self-Dissimilarities. *Asian Conference on Computer Vision*, 586-600.
- Tola, E., V. Lepetit, and P. Fua, 2010. Daisy: An efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(5), 815-830.
- Trzcinski, T. M. Christoudias, V. Lepetit. Learning Image Descriptors with Boosting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(3), 597-610.
- Wallis, K.F., 1974. Seasonal adjustment and relations between variables. *Journal of the American Statistical Association*, 69(345), 18-31.
- Žbontar, J., Le Cun, Y., 2015. Computing the stereo matching cost with a convolutional neural network. *Conference on Computer Vision and Pattern Recognition*, 1592-1599.