Automatic Image-to-Geometry Registration in Varying Illumination Conditions using Local Descriptors

Christian Kehl and Simon J. Buckley

[1] UniResearch AS – CIPR, P.O.Box 7810, N-5020 Bergen
{Christian.Kehl,Simon.Buckley}@uni.no
[2] University of Bergen – Dept. Earth Science, P.O.Box 7803, N-5020 Bergen

Abstract. The registration of outdoor photograph to available 3D models is a common procedure in geoscientific workflows. Current algorithmic approaches are limited in their applicability by inaccurate initial external orientation and position estimates, and fail to produce satisfying registration results under varying radiometric conditions, such as illumination changes. This article presents new results in the application of digital elevation models for initial positioning. Furthermore, new algorithmic feature-based registration methods are introduced and assessed together with image processing methods to improve image-to-geometry registration accuracy under challenging outdoor illumination conditions. Our results indicate that alternative local descriptors in combination with image processing allow 2D-3D co-registrations that were previously impossible to retrieve.

1 Introduction

The mapping of photographic images to geoscientific surface model in 3D, referred to as image-to-geometry registration, received increasing attention in recent years. Major contributing factors for this interest is the large availability of surface models due to structure-from-motion (SfM) and lidar scanners, and the availability of photographs from ubiquitous mobile devices. The mapping can be used to re-texture existing surface models or introduce semantics in a model via annotated photographs. Several applications for the registration have been shown recently, ranging from outdoor navigation [1] over the detection of river flood events [2] to the mapping of geological interpretations [3].

The image-to-geometry registration on mobile devices is commonly realised using a 2D-3D feature correspondence matching. Although alternative registration approaches, such as mutual information [4], exist, feature-based registration is the prevalent approach across application domains because of the availability of algorithms (e.g. feature detection, pose-n-perspective (PnP) estimation), their simple parameterisation and their control via image processing. Furthermore, all necessary subroutines are available and accessible on novel computing platforms, such as Android mobile device. A major challenge within feature-based image registration is its sensitivity to radiometric variance (e.g. due to illumination changes) between the texture of the surface model and the to-be-registered photographs. This illumination sensitivity is illustrated in Fig. 1, where the same textured surface model is used to register photos obtained on different field campaigns for digital outcrop studies. The photo obtained simultaneously to the

texture (Fig. 1, top) is correctly registered, while the same algorithm fails in registering the photo of the other campaign (Fig. 1, middle). This is a problem also common to many other computer vision applications, such as outdoor navigation and structure-from-motion reconstruction. Initial experiments with alternative detectors have shown promising results (see Fig. 1, bottom). Furthermore, as discussed by Kehl et al. [5, 3], initial estimates for position and orientation need to be available for the registration. Mobile device sensors acquire infeasible position data due to GPS inaccuracies.

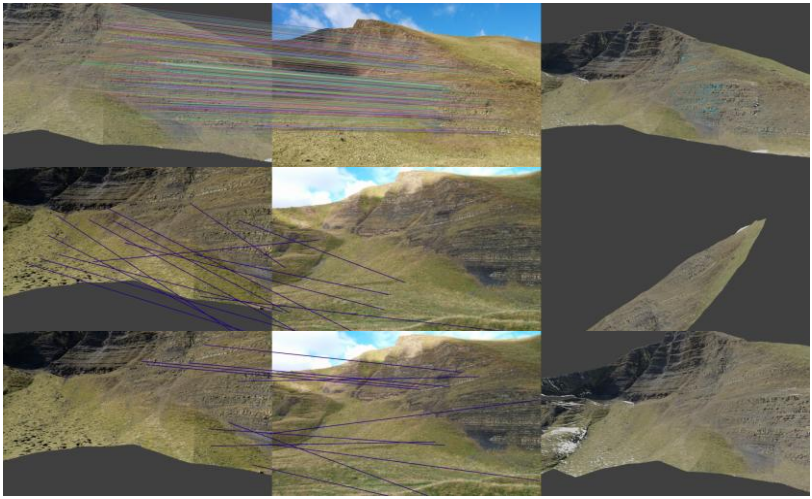


Fig. 1 Variability of registration accuracy on illumination conditions. A previously successful SIFT matching (top) may fail with changing illumination (mid). Alternative feature algorithms (e.g. MSCR, bottom) may resolve the illumination sensitivity.

This article introduces new approaches to overcome existing limitations in the coarse pose approximation and the sensitivity to illumination changes, targeted to mobile device implementations. In detail, the article shows the initial position improvement achievable by using compact, publicly available digital elevation models (DEMs). More importantly, the article assesses the impact of using the Wallis filter [6] and gamma adaptation [7] on the pose estimation accuracy of feature-based image-to-geometry registration methods. As a validation dataset for the approach photo series from two different field campaigns are presented which exhibit challenging, large radiometric variations. They are registered to a lidar-based digital outcrop model (DOM).

2 Related Work

The photogrammetric literature already contains studies on automatic, feature-based image-to-image registration, such as Jazayeri and Fraser[8]. This particular article builds on the work of Kehl et al. [5, 3], where openly available feature detection and pose estimation algorithms are used to register natural photographs

to lidar-derived DOMs. The approaches utilise mobile device sensor data to obtain a coarse, initial pose necessary for the subsequent pose estimation refinement. Recent developments along that research trajectory [9] have shown that feature detection-, description- and mapping combinations other than the established scale-invariant feature transform (SIFT) [10] and speeded-up robust features (SURF) [11] lead to improvements of image-to-geometry registration procedures under challenging radiometric variance. This article hence further compares the impact of image processing on SIFT-based, maximally-stable colour regions (MSCR)-based [12] and "Features from Accelerated Segment Test" (FAST)-based [13] feature correspondence- and registration methods.

The application of global greyscale image processing filters to adapt image collections to illumination changes is under-represented in the literature, although it represents a logical extension to consider when targeting image registration improvements. In many application scenarios that consider image series, such as outdoor navigation, a temporal dependency amongst images is observed, which is taken advantage of when developing new algorithms. This advantage vanishes when considering image series with little temporal or spatial overlap, which is why this article discussed image filters that are more generally applicable. The Wallis filter [6] has found many applications in the past, such as for image registration [8] and the reconstruction of featureless environments, showing promising improvements on feature-based algorithms. Furthermore, it is possible to smoothly adapt the contrast and brightness using gamma adaptation [7] to account for illumination variance.

Although the use of other image-to-geometry registration procedures, such as mutual information [4, 14], also potentially resolve the illumination sensitivity, these methods are not considered as their implementation on target mobile device platforms is currently not supported. In particular, mobile device platforms lack support for common linear algebra software (e.g. BLAS and LAPACK).

3 Method

The workflow in this study adapts the approach explained in [3, 9] and shown in Fig. 2 by applying the selected imaging filter before feature extraction and matching. Hence, for each obtained photo the related, textured surface model is rendered using a virtual camera, for which the external orientation and position is generated using mobile sensor data (i.e. GPS/DEM and magnetometer). The DEM-based altitude is obtained by interpolating height values on a given, local DEM using the GPS latitude and longitude. The rendered image and, in the case of Wallis filtering, the photo are subject to the image processing. The processed images are then used for salient point detection, feature description and feature mapping. The resulting 2D image features have a 3D correspondence in the surface model via the virtual camera of the rendered image. The 2D-3D pointsets are put into a PnP pose optimization to retrieve the final registration.

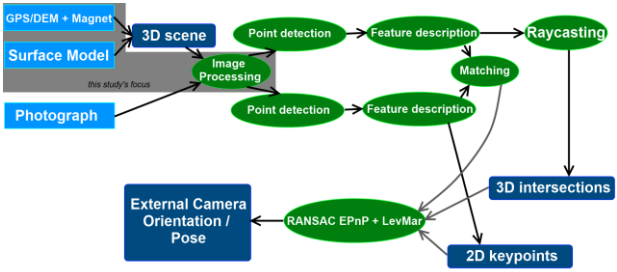
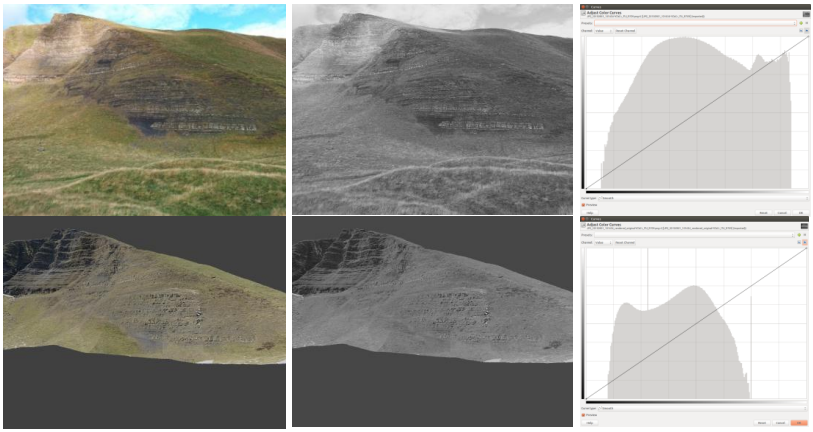


Fig. 2 Algorithmic pipeline of the adopted image-to-geometry approach. The focus of this article is on the initial positioning and the image pre-processing influence.

The application of the image processing and filtering is based on the fact that most salient point detectors and feature descriptors operate on the greyscale channel of an image. On the other hand, the greyscale channel (i.e. "luminance") is the one affected by radiometric variance, which is introduced by changing illumination. Therefore, adapting the greyscale channel of an image to account for illumination changes potentially results in an improved registration accuracy that resolves issues shown in Fig. 1. The Wallis filter is commonly known in photogrammetry and computer vision [6]. The gamma adaptation [7] controls brightness and contrast of the greyscale curves of images. Fig. 3 shows a photograph (top) and its corresponding rendering (bottom), their comparative greyscale channels (i.e. the Y channel using YCrCb decomposition) and intensity value curves.



(a) (b) (c)

Fig. 3 Luminance difference of a real (top) and synthetic (bottom, γ=1.6) image, their luminance channel (b) and histograms (c).

The registration evaluation procedure considers the following metrics, which are identical to other studies on illumination variance[9]: The feature detection results in a collection of salient 2D points in the photo I_p and the synthetic image I_s. The

number of correspondences states the number of mutually-unique point mappings from I_p to I_s. A point pair correspondence is considered matching if the points are within a bound ε of one of the epipolar lines connecting both images. The inlier ratio after feature extraction ($r_{inlier(feat)}$) is the ratio between the number of matches and the number of correspondences. The inlier ratio after pose optimisation ($r_{inlier(opt)}$) represents the ratio of matched points which adhere to the epipolar constraint before and after the pose optimisation. The reprojection error $\Delta(px_{feat}, px_{proj})$ is the common measure of average point distance between the initial 2D features and their image plane-projected 3D correspondences. Applying the reprojection error as the single evaluation metric is a potential risk as it is prone to erroneous correlations. The metric values presented in the result section are average values, obtained as the mean of each metric over the full image series. The success rate is determined by manual pose assessment of the final synthetic image and represents the ratio between correctly registered image and the size of the image series.

4 Dataset

In the experimental study, the image datasets and the DOM initially introduced by Kehl et al.[3] are used as they contain the challenging registration conditions this article attempts to resolve. The DOM is a TLS-based textured surface mesh with irregular spacing and possible mesh holes. The photos are split into two image series: the first series was obtained simultaneously to the DOM texture during an outcrop field campaign in March 2015, and the second series was obtained at the same locality during a separate field campaign in September 2015. On both campaigns, the photos were acquired with a Google Nexus 5 smarthphone using the integrated 8 megapixel camera. The initial pose (i.e. external orientation and position) is generated using the GPS or the DEM and the logged geo-magnetic orientation of the mobile device. Therefore, the image series are both subject to significant geometric variance (in respect to the accurate pose) and a radiometric variance (i.e. changing illumination condition) between each other.

As a reference for the accuracy improvement using image processing, the results presented in [9] under both variances are taken as a baseline. The experiments hereby further presented are reduced to the algorithmic combinations of FAST detector and SIFT descriptor, MSCR detector and SIFT descriptor, and a pure SIFT detection and description. Table 1 shows the obtained baseline results.

	SIFT	MSCR-SIFT	FAST-SIFT
# points	12310	**5727.7**	43020
# correspondences	8.31	**9.93**	3.45
# matches	**2.71**	0.84	1.12
$r_{inlier(feat)}$	**0.33**	0.09	**0.33**
$r_{inlier(opt)}$	0.02	0.13	**0.33**
$\Delta(px_{feat}, px_{proj})$	11.77	**10.05**	7392.7
success rate	0/52	**10/52**	9/52

Table 1 Original registration accuracy results, presented in Kehl et al. [8]

4 Results

a) DEM positioning

In order to assess the pose positioning improvement provided by the DEM, the differences of the altitudes and sensor data pose positions are compared. The differences (treated as mean error deviations) are measured for the GPS and the combined data of GPS latitude-longitude and DEM altitude, with respect to a manually-obtained, control points-based, PnP-optimized pose. The experiments used an external GPS receiver via Bluetooth to overcome consumer electronics limitations. The results are mean values of 9 selected data points within the studied dataset, shown in Fig. 4.

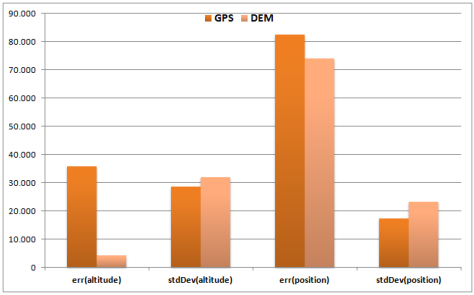


Fig. 4 Accuracy improvement of the initial positioning using DEMs instead of GPS.

As can be seen from the diagram, the application of the DEM for altitude determination improves upon the position accuracy by approximately 8m. On the other hand, this position is subject to a larger standard deviation than the GPS altitude signal in the field. Reasons for the large standard deviation are the low lateral resolution (in this case 25m x 25m), the linear interpolation scheme and the terrain variability close to outcrop sections (e.g. cliffs and hill sections). In such remote outdoor locations, the GPS satellite reception is good compared to urban environments, giving an accurate lateral position compared to the lateral DEM resolution. Despite the good reception, the GPS altitude measurement is still more erroneous after geoid processing than DEM altitude measurements that occur close to cell corners.

b) Wallis filtering

Table 2 shows the statistical results of the feature-based registration procedure on the Wallis-filtered photos and synthetic image data, compared to Table 1.

	SIFT	MSCR-SIFT	FAST-SIFT
# points	14129	**5516.6**	43104
# correspondences	8.69	**10.36**	6.62
# matches	**2.9**	0.83	1.41
$r_{inlier(feat)}$	**0.32**	0.08	0.15
$r_{inlier(opt)}$	-	0.14	**0.52**

$\Delta(px_{feat}, px_{proj})$	-	**13.09**	5019.7
success rate	0/52	11/52	**15/52**

Table 2 Improved registration accuracy results by applying the Wallis image filter.

Within the measured metrics, it can be seen that Wallis filtering improves matching and registration results across all the SIFT feature point alternatives. The registration success of a SIFT-exclusive algorithmic pipeline, on the other hand, is insignificantly affected by image processing. Aside from the statistical measures, it was observed that Wallis filtering resulted in a less stable registration procedure where some image pairs – easy to register without image modification – fail to produce satisfactory results.

In the FAST-SIFT combination, the increased contrast response of the salient points' neighbourhood on approximately doubles the retrieved correspondences. This leads to an increased success rate for registered image. Despite the observed improvement, a large amount of correspondences remain incorrect. This can be explained when considering the image content: the rendered image lacks background information and potentially contains holes in the model. The histograms of the synthetic image and the photo are in these cases not compatible. The negative impact of this histogram mismatching on the registration is illustrated in Fig. 5.

Fig. 5 Example of a failed pose estimation using FAST-SIFT after Wallis filtering (bottom), compared to the original (top), due to histogram mismatches.

Generally, the image processing via Wallis filtering improves the matching and, in some cases, is able to retrieve correct external orientation and position information for distinctly different images, as shown in Fig. 6.

Fig. 6 Prime example of the potential gain in registration accuracy using Wallis filtering. The hereby retrieved pose using FAST-SIFT was previously unretrievable.

c) Gamma adaptation

The gamma adaptation has been studied for multiple realisations of γ because a single adaptation value does not cover the range of illumination differences. The shown experiments cover γ=0.8 (Table 3) and γ=1.3 (Table 4), comparable to Table 1.

	SIFT	MSCR-SIFT	FAST-SIFT
# points	14120	**5242.5**	59116
# correspondences	9.27	**10.4**	7.98
# matches	**2.4**	0.79	1.62
$r_{inlier(feat)}$	**0.26**	0.08	0.12
$r_{inlier(opt)}$	0.02	0.11	**0.46**
$\Delta(px_{feat}, px_{proj})$	**2.81**	6.86	19.86
success rate	0/52	6/52	**21/52**

Table 3 Improved registration accuracy results after gamma adaptation (γ=0.8).

	SIFT	MSCR-SIFT	FAST-SIFT
# points	14068	**6833.3**	74447
# correspondences	9.23	**10.56**	5.53
# matches	**3.58**	0.63	2.14
$r_{inlier(feat)}$	**0.39**	0.06	0.23
$r_{inlier(opt)}$	0.03	0.09	**0.54**
$\Delta(px_{feat}, px_{proj})$	1223.9	**32.77**	4725.1
success rate	0/52	3/52	**13/52**

Table 4 Improved registration accuracy results after gamma adaptation (γ=1.3).

In the conducted experiments, SIFT- and MSCR detectors are negatively affected by the gamma adaptation while FAST-based registration has significantly improved. The registration accuracy achieved by FAST salient points on the γ=0.8 adaptation exceeds every formerly-achieved registration. On the other hand, the registration success rate and accuracy drops with increasing gamma values (see Fig. 7). Experiments with γ=1.6 only result in a success rate of 9 out of 52 images.

photo γ=0.8 γ=1.0

γ=1.3 γ=1.6 γ=1.9

Fig. 7 Registration results for gamma-adapted FAST-SIFT with different γ values.

5 Discussion

In this article, the accuracy impact of DEM-based altitude estimation and different greyscale channel image filters on feature-based image-to-geometry registrations is assessed. A challenging outdoor dataset has been chosen for the assessment that is affected by large illumination changes between two field campaigns in March 2015 and September 2015. The photos are acquired with a commonly-available Google Nexus smartphone.

In the DEM experiments, a significant positioning improvement is observed in comparison to the use of pure GPS positioning. The obtained positioning accuracy is limited the coarse available resolution of the DEM. This means that the results can be improved when higher-resolution DEMs for field locations become publically available. On the other hand, the low-resolution, small grids can be parsed in real-time on mobile devices. DEM-supported positioning also has its limits, as DEM queries close to vertically-straight cliff sections commonly result in large errors.

The experiments of image processing for the reduction of radiometric variance effects provides multiple conclusions: While the Wallis filter resulted in a minor or intermediate accuracy improvement for MSCR- and FAST detector alternatives, the gamma adaptation has presented the potential for significant accuracy improvements for FAST salient point-based registration. Achieving this accuracy improvement across different scenarios by selecting γ parameters remains a challenge.

In case studies with less drastic radiometric differences or more accurate initial pose information, the application of the Wallis filter yields sufficient improvements to co-register natural photographs on the related, textured surface model.

Acknowledgement

This research is part of the “VOM2MPS” project (Petromax 2, grant number 234111/E30), funded by the Research Council of Norway, the FORCE consortium and SAFARI.

References

[1] C. Sweeny, J. Flynn, B. Nuernberger, M. Turk und T. Höllerer, „Efficient

Computation of Absolute Pose for Gravity-Aware Augmented Reality," in s *International Symposium on Mixed and Augmented Reality (ISMAR)*, Adelaide, Australia, 2015.

[2] M. Kröhnert, "Automatic Waterline Extraction from Smartphone Images," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* vol. XLI, no. B5, pp. 857-863, 2016.

[3] C. Kehl, S. J. Buckley, R. L. Gawthorpe, I. Viola and J. A. Howell, "Direct Image-to-Geometry Registration Using Mobile Sensor Data," *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information,* vol. 3, no. 2, pp. 121-128, 2016.

[4] M. Corsini, M. Dellephiane, F. Ganovelli, R. Gherardi, A. Fusiello and R. Scopigno, "Fully Automatic Registration of Image Sets on Approximate Geometry," *International Journal of Computer Vision,* vol. 102, no. 1-3, pp. 91-111, 2013.

[5] C. Kehl, S. J. Buckley and J. A. Howell, "Image-to-Geometry Registration on Mobile Devices - An Algorithmic Assessment," in *Proceedings of 3D NordOst*, Berlin, 2015.

[6] K. F. Wallis, "Seasonal adjustment and relations between variables," *Journal of the American Statistical Association,* vol. 69, no. 345, pp. 18-31, 1974.

[7] P. A. Charles, Digital Video and HDTV: Algorithms and Interfaces, Morgan Kaufmann, 2003, p. 260.

[8] I. Jazayeri and C. S. Fraser, "Interest operators for feature-based matching in close range photogrammetry," *The Photogrammetric Records,* vol. 25, no. 129, pp. 24-41, 2010.

[9] C. Kehl, S. J. Buckley, I. Viola, S. Viseur, R. L. Gawthorpe und J. A. Howell, „Automatic Illumination-Invariant Image-to-Geometry Registration in Outdoor Environments," *The Photogrammetric Record,* 2017 (to be published).

[10] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision,* vol. 60, no. 2, pp. 91-110, 2004.

[11] H. Bay, T. Tuytelaars and L. Van Gool, "SURF: Speeded Up Robust Features," in *European Conference on Computer Vision (ECCV)*, 2006.

[12] P.-E. Forssén, "Maximally Stable Colour Regions for Recognition and Matching," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007.

[13] E. Rosten, R. Porter and T. Drummond, "Faster and Better: A Machine Learing Approach to Corner Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 32, no. 1, pp. 105-119, 2010.

[14] G. Caron, A. Dame and E. Marchand, "Direct model based visual tracking and pose estimation using mutual information," *Image and Vision Computing,* vol.

32, no. 1, pp. 54-63, 2014.