DE GRUYTER

Current Directions in Biomedical Engineering 2019;5(1):273-276

Konstantin Sieler*, Ady Naber, Werner Nahm

An Evaluation of Image Feature Detectors Based on Spatial Density and Temporal Robustness in Microsurgical Image Processing

Abstract: Optical image processing is part of many applications used for brain surgeries. Microscope camera, or patient movement, like brain-movement through the pulse or a change in the liquor, can cause the image processing to fail. One option to compensate movement is feature detection and spatial allocation. This allocation is based on image features. The frame wise matched features are used to calculate the transformation matrix. The goal of this project was to evaluate different feature detectors based on spatial density and temporal robustness to reveal the most appropriate feature. The feature detectors included corner-, and blob-detectors and were applied on nine videos. These videos were taken during brain surgery with surgical microscopes and include the RGB channels. The evaluation showed that each detector detected up to 10 features for nine frames. The feature detector KAZE resulted in being the best feature detector in both density and robustness.

Keywords: feature detection, KAZE, SURF, SIFT, Harris, MSER, MinEigen, BRISK, Neurovascular, Spatial, Temporal

https://doi.org/10.1515/cdbme-2019-0069

1. Introduction

Many brain surgery procedures involve and rely on technical equipment and diagnostic support. Surgical microscopes provide visual magnification and give access to video streams, such as in the Red (R), Green (G), Blue (B) or Infra-red channels. Diagnostic image processing algorithms have the ability to extract anatomic and physiological information of the patient's condition. Often no additional hardware is required, as in movement compensation, and most algorithms are based on features extracted directly from the images recorded by the surgical microscopes. This paper is focused on the evaluation of the spatial and temporal feature behavior. Unlike most state-of-the-art investigations [1], this paper is focused specifically in neurovascular videos, to reveal the best performing feature and which therefore can be recommended for usage in image feature-based applications on neurovascular data. The spatial evaluation aims to find the feature detector with the highest number of features detected in a region of interest (ROI). The temporal evaluation aims to assess the features persistence, so the feature that can be matched continuously over the most consecutive frames.

2. Method

The processing workflow is divided into three parts, beginning with the selection of the channel, the computation of the features and the evaluation. All implementations were done in MATLAB R2018b.

2.1. Data

The nine videos used in this work are hand chosen extracts of videos that were taken during cerebrovascular surgery and each one is 16 frames long. All videos contain the RGB channels recorded by a surgical microscope pointed and focused at the vessel of interest. The vessels are rather deeply located in the brain; therefore, a deep surgical access channel is partly visible. Clips of aneurysms and minimal surgical tool movement in the field of view are allowed.

2.2. Channel Selection

The image information from the surgical microscopes was obtained in 3 channels: R, G, B. In this work the R, G, and B

***Corresponding author: Konstantin Sieler:** Karlsruhe Institute of Technology (KIT), Institute of Biomedical Engineering (IBT), Kaiserstrasse 12 in 76135 Karlsruhe, Germany, e-mail: publications@ibt.kit.edu

Ady Naber, Werner Nahm: Karlsruhe Institute of Technology (KIT), Institute of Biomedical Engineering (IBT), Karlsruhe, Germany

channel and the grayscale image were evaluated separately. The grayscale image was computed using the recommended values of the International telecommunication Union (ITU): ITU-R BT.601-7 [2].

$$\boldsymbol{grayscale} = \mathbf{0.299} * \boldsymbol{R} + \mathbf{0.587} * \boldsymbol{G} + \mathbf{0.114} * \boldsymbol{B} \quad (\mathbf{1})$$

The coefficients in (1) are chosen so that the grayscale image produces the greatest luminance. This was done based on the non-linear perception of the eye. The channel with the highest number of features detected was selected for the following computation.

2.3. Feature Computation

Following the stages in figure 1, the features were tracked and mapped over time.

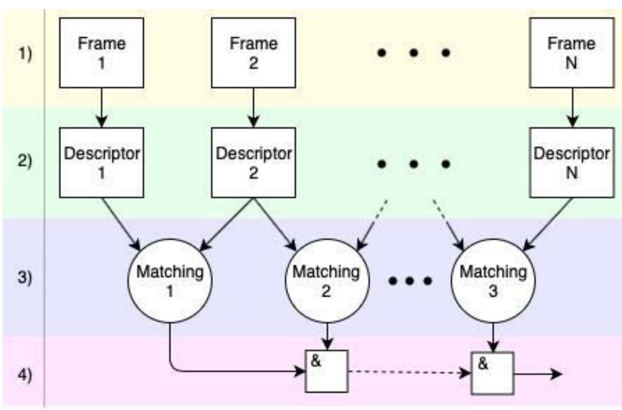

Figure 1: The four staged computational approach of the feature processing per sequence.

Stage 1) Video sequences were selected from the clinical videos for the analysis of the features. These sequences contain no changes in the microscopes settings as the magnification or focus.

Stage 2) The following detectors were applied to detect and extract features. The feature detectors (shown in Table 1) were selected according to the ones proposed in the literature and additional common features. After the detection of all features the descriptors were extracted separately of each frame in the sequence. Only features in a ROI were accepted. The ROI was segmented automatically according to the method proposed by Wirth et al. [9]. In this method the ROI is defined by color and focus segmentation. This is valid as long as the assumption that the information of interest is in focus and is not black or white holds.

Stage 3) The features were matched pairwise between 2 consecutive frames and for all consecutive frame-pairs in the sequence. To reduce the chance of mismatches the Euclidean distance between the initial and all other feature locations were computed. Features spatially shifted above a maximum accepted range of 40 pixels between consecutive frames were removed from the evaluation, as no such far pixel displacement was expected. The maximum range was set empirically.

Table 1: Feature detectors applied in this paper sorted by type. [2]

Corner Detectors	Blob Detectors
Harris [3]	MSER [4]
FAST [4]	SURF [6]
MinEigen [5]	BRISK [7]
	KAZE [8]

Stage 4) For temporal evaluation purposes the persistence of features that could be matched consecutively from frame to frame was computed. Any features that could not be matched were removed from the temporal evaluation.

2.1. Evaluation

The feature detectors were evaluated based on two key aspects: Spatial and temporal behavior.

- **Spatial behavior:** The average total number of features detected and feature density per image. The density is defined as the number of features per 100.000 pixels.
- **Temporal behavior:** The temporal robustness is evaluated using the number of persistent features. The parameter for comparison was set to be the number of features that could be matched consecutively up to each frame of the video.

1. Results

The results are presented in the order as in section two, showing the channel selection and the spatial and temporal behavior.

1.1. Channel Selection

Resulting in the highest number of features detected (Figure 2), the green channel was chosen for further investigation and the evaluation.

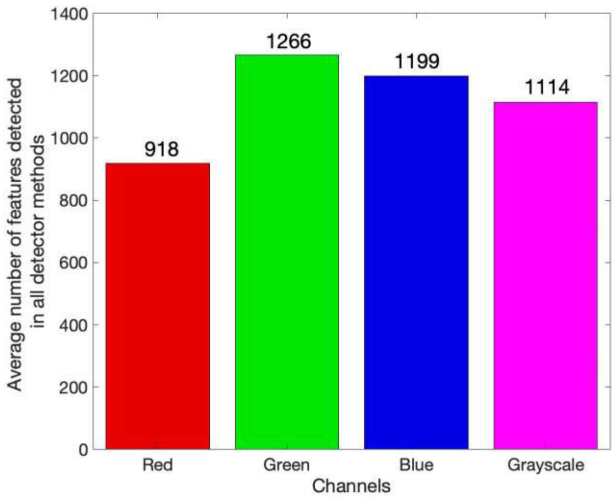

Figure 2: These results show the number of features detected for each investigated channel. The number of features presented per channel is the average of the number of features detected of all detector methods and all images in the video sequences.

1.2. Spatial Behavior

A sample frame with the detected SURF features is shown in Figure 3. The obtained feature count and computed density is shown in Table 2.

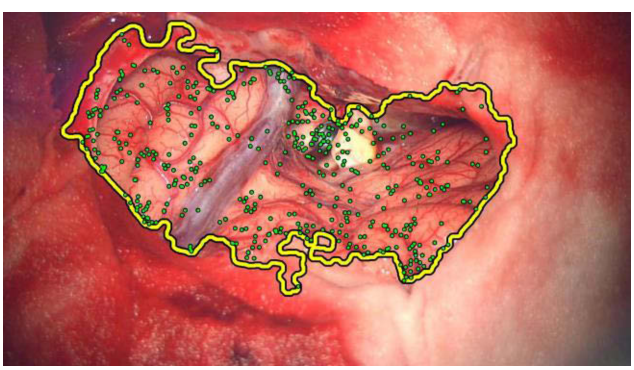

Figure 3: A single frame of the videos after analysis with the SURF detector. The green dots mark the location of the detected features. The yellow boundary marks the ROI only in which features were allowed for evaluation. [10]

Table 2: Average number of features detected in all frames of all videos of each detector and the feature density of 100 by 100 pixels. The results were rounded to two decimal digits. The half-life was calculated starting from the initial frame and the values were interpolated linearly between frames.

Detector	Number of Features	Density of Features	Half-Life in Frames
KAZE	1532	12,24	3,58
MinEigen	403	3,13	2,45
SURF	171	1,18	3,54
BRISK	85	0,63	2,58
Harris	69	0,51	2,82
MSER	59	0,38	1,32
FAST	26	0,20	2,54

1.3. Temporal Behavior

The half-life of the features is presented in Table 2. Figure 5 shows the number of persistent features detected in each frame.

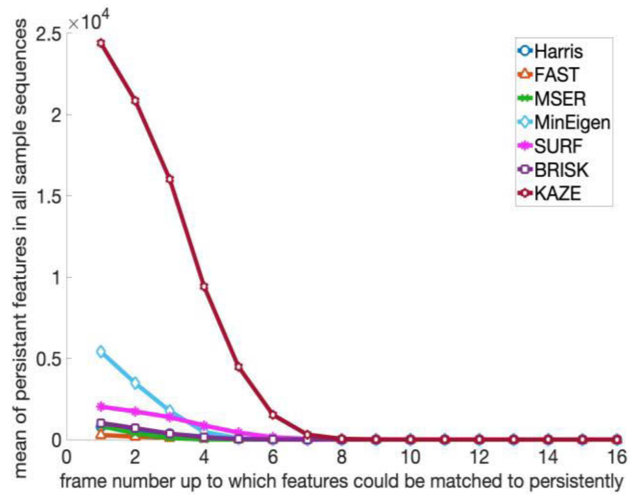

Figure 5: The number of persistent features that could be matched continuously starting from the initial frame in the sequence.

1. Discussion & Conclusion

As expected the result of the channel selection showed a clear dependency on the input channel (Figure 2). The green channel performed best due to the high contrast caused by the blood's optical properties. The linear combination of the RGB channels according to the ITU-R BT.601-7 standard did not increase the average total number of detected features. The green channel was therefore used for the further evaluation of the detector methods in this paper. The investigation of the spatial density of the features in the frames resulted in a clear best performing feature. The KAZE feature detector had at least a four times higher density (Table 2). Judging subjectively, the KAZE features are homogeneously distributed. Small regions in the image with little texture were, as expected, not suitable for the corner and edge detectors. The features evolution over time is shown in Figure 5 and shows clearly that the KAZE feature had the most total matches continuously in at least nine consecutive frames and the greatest half-life. Overall no feature detection, extraction and matching method was able to match more than ten features continuously from the initial frame to more than the ninth frames which does not limit the tracking since most methods use framewise matching only. The susceptibility to outliers could be reduced by using multiple consecutive frames and persistent features as the KAZE. The MinEigen and SURF feature detectors had the second and third highest feature density and therefore should be also considered to be used. The MinEigen feature further showed the best performance in the R-channel (Figure 6). The SURF feature had a higher half-life, which enables longer and more stable tracking. The BRISK, MSER, Harris and FAST features had overall a noticeable lower performance. These results do not totally comply with the results in literature where the KAZE feature did not perform best in the terms of the features density [11]. In the literature no surgical images are included and this reveals the need of problem related selection of features and a tailor-made processing chain.

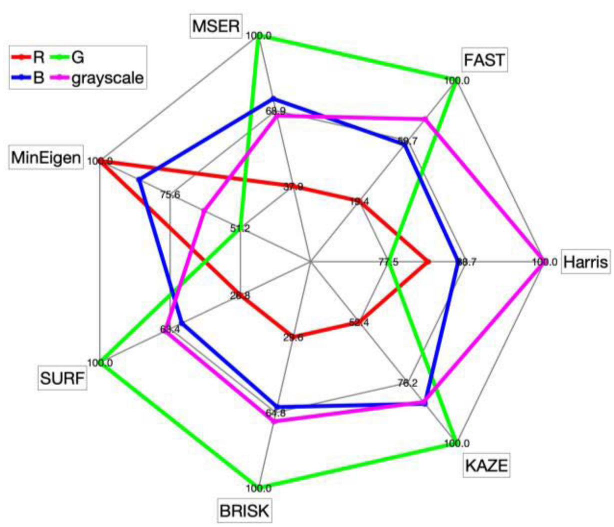

Figure 6: This Figure shows the average number of features detected in all frames. The results were scaled to 100 for each detector method.

2. Outlook

The evaluation of the input channel was done based on the average total number of features detected by the proposed detector methods in all videos. This can be further extended by also evaluating the temporal robustness of all features in the different channels. The outcome of the channel selection was to proceed with the green channel, which was expected due to the optical properties of the blood. But as shown in Figure 6, some features perform better in a different channel. Consequently, the combination of channels should enhance the total number of detected features. This approach would result in a novel and application tailored combination of the R, G and B channel (similar to the calculation of the grayscale image in Equation 1) to optimize the feature detection.

Author Statement

Research funding: Mr. Ady Naber received a scholarship of the Karlsruhe School of Optics & Photonics. Conflict of interest: Authors state no conflict of interest. Informed consent: Informed consent has been obtained from all individuals included in this study. Ethical approval: The research related to human use complies with all the relevant national regulations, institutional policies and was performed in accordance with the tenets of the Helsinki Declaration, and has been approved by the authors' institutional review board or equivalent committee.

Acknowledgement

We thank Priv. Doz. Dr. med. Uwe Spetzger of the Städtisches Klinikum Karlsruhe for the support by providing videos of neurovascular interventions.

References

1. Mouats T., Aouf N., Nam D. and Vias S. (2018). Performance Evaluation of Feature Detectors and Descriptors Beyond the Visible. Journal of Intelligent & Robotic Systems. 1-31. 10.1007/s10846-017-0762-8.
2. International Telecommunications Union. (2011, March). Recommendation ITU-R BT.601-7. Retrieved from https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.601-7-201103-I!!PDF-E.pdf (last visited 25.07.19).
3. Harris, C., and M. Stephens, "A Combined Corner and Edge Detector," Proceedings of the *4th Alvey Vision Conference*, August 1988, pp. 147-151.
4. Rosten, E., and T. Drummond. "Fusing Points and Lines for High Performance Tracking," *Proceedings of the IEEE International Conference on Computer Vision*, Vol. 2 (October 2005): pp. 1508–1511.
5. Shi, J., and C. Tomasi, "Good Features to Track," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 1994, pp. 593–600.
6. Bay, H., A. Ess, T. Tuytelaars, and L. Van Gool. "SURF:Speeded Up Robust Features." Computer Vision and Image Understanding (CVIU).Vol. 110, No. 3, pp. 346–359, 2008.
7. Leutenegger, S., M. Chli and R. Siegwart. "BRISK: Binary Robust Invariant Scalable Keypoints", Proceedings of the IEEE International Conference, ICCV, 2011.
8. Alcantarilla, P.F., A. Bartoli, and A.J. Davison. "KAZE Features." ECCV 2012, Part VI, LNCS 7577. 2012, p. 214
9. Wirth, T., Naber, A. & Nahm, W. (2018). Combination of Color and Focus Segmentation for Medical Images with Low Depth-of-Field. Current Directions in Biomedical Engineering, 4(1), pp. 345-349. Retrieved 12 May. 2019, from doi:10.1515/cdbme-2018-0083
10. Priv. Doz. Dr. Med. Uwe Spetzger. (Received 2019) Städtisches Klinikum Karlsruhe.
11. S. A. K. Tareen and Z. Saleem, "A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK," 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Sukkur, 2018, pp. 1-10. doi: 10.1109/ICOMET.2018.8346440