# Comparison of Image Features Descriptions for Diagnosis of Leaf Diseases

Muhammad Waqas and Norishige Fukushima
Nagoya Institute of Technology, Japan

***Abstract*—The agricultural industries have always demanded technologies for the automatic discovery and diagnosis of plant diseases with high speed, accuracy, and low cost. Numerous studies have been conducted in response to this demand; however, significant issues remain in most cases where a large-scale dataset of field images is taken with different atmospheric conditions, lighting, scale, and in different directions. The large dataset often causes high computational and storage costs. To overcome this problem, we focus on methods based on efficient invariant image features. These methods are robust against such external factors added during image acquisitions with low computational cost and higher accuracy. We then use a well-known data clustering algorithm k-means to create visual features for lesions. We then create a group of robust visual features (BoVF) using the Term Frequency-Inverse Document Frequency (TF-IDF) weighting scheme that considers the most important visual features in the image for classification. Experimental results classify the BoVF using K-means clustering that categorizes a particular disease in the leaf image into their appropriate group.**

## I. Introduction

Quality control of agricultural products is one of the essential issues for supporting economic activities worldwide. Stably supplying these agricultural products to the market, farmers need to properly manage the fields such as to monitor the field conditions, ensure the appropriate use of water and pesticides based on climate and humidity change, and the sanitary management of farm equipment. Among these management measures, early detection of plant diseases, which is the direct cause of the decline in the commercial value of crops, is especially important. In particular, plant virus diseases are dangerous in the field because they can cause secondary damage due to the spread of other crops in the same field, and since it is not possible to treat the infected strains once infected. Due to such viral diseases, considerable crop losses have been attributed in Pakistan, India, United States, Australia, and Japan. Annual crop losses due to only plant viral diseases are estimated worldwide at 60 billion dollars [1].

The most common method for early disease discovery and diagnosis is a visual diagnosis by farmers and experts in plant pathology. However, this approach is very time-consuming because the diagnostician must judge the entire field one by one. Besides, visual diagnosis requires advanced knowledge to detect abnormalities in leaves, stems, and roots at an early stage. It is difficult to distinguish the presence and type of symptoms appearing in a plant without some experience, and especially the symptom in the early stage is almost indistinguishable from the healthy one. Also, when an ordinary person discovers such diseases, however, the symptoms may have already progressed to the end stage and are uncontrollable.

Using a microscope for genetic information is another solution for the visual diagnosis. LAMP method is a typical method for gene diagnosis [2]. LAMP is an effective method for plants of the family, e.g., Cucurbitaceae and Tomato, but this diagnosis requires a dedicated kit, and the types of viruses that can be diagnosed are limited. Furthermore, the kit is also hard for farmers who are inexperienced in farming or for ordinary people who do not have specialized knowledge.

The entire series of processes from the discovery to the diagnosis of plant diseases in the agricultural field is currently carried out based on the intuition and experience of farmers and specialists. This fact not only causes a major burden on farmers but in the case of a disease such as a viral disease that, if spreads, there is a possibility that the amount of damage caused by oversight will also increase. Therefore, automating techniques for the diagnosis of plant diseases characterized by speed, accuracy, and low computational cost have been demanded by the agricultural industry [3].

In response to such demand, this study aims to reduce this burden on farmers by image processing technology. In plants, the disease symptoms appear in different parts of the plant, such as leaves, stems, and roots. Among them, leaves are often used as clues for diagnosing disease symptoms, and leave observation with a camera is more accessible than stems and roots. The automatic diagnosis of plant leaf diseases is always affected by a large number of external factors such as light, shadows, shaking leaves by the wind, and the camera position. The conventional methods rely on a large dataset of field images taken under such circumstances to minimize the effect of such external factors for covering the issues. However, some approaches, such as CNN based methods, overcome the issues, but newly extracted issues, which are high computational cost and storage cost, are unrealistic for business.

In this study, we overcome the issues by assessing invariant image feature descriptors, which are robust against these external factors and can keep low computational and storage costs. We investigate various descriptors, e.g., SIFT, SURF, ORB, KAZE, and AKAZE. Further, we use k-means clustering [4], which creates the most important visual features for disease classification. The center of each cluster is called the visual word. Then we use TF-IDF [5] as the weighting scheme for evaluating the frequency of visual words [6]. In this study, we classified 680 plants leave images into seven different classes.

This work was supported by JSPS KAKENHI 17H01764, 18K19813.

Fig. 1: Example of sample images of PlantVillage Dataset.

## II. Related Work

In the past, the representative researches of the automatic diagnosis of plant leaf diseases are the gray-level co-occurrence matrix (GLCM) and the spatial gray-level dependence matrix (SGDM). For example, Huang et al. [7] trained a neural network with GLCM features. They identified a total of three types of lesions that appeared on orchid leaves and achieved an average identification performance of 89.6%. Kai et al. attempted to identify lesions on corn leaves. They trained a BP neural network with GLCM features and succeeded in diagnosing three types of lesions with a precision of 98% [8]. Bakshish et al. proposed a diagnostic framework that does not depend on the type of plant. They extracted the features using SGDM acquired on the HSV color space and achieved an average classification performance of 89.5% for a total of five lesions using a three-layer neural network [9]. Arivazhagan et al. have proposed an identification method using SVM. They trained the SVM classifier with SGDM features calculated from the HSI image of the leaves. This classifier achieved an average classification performance of 94.74% for many types of plant diseases such as banana and guava [10]. Plant disease diagnosis using these conventional methods are not invariant for external factors and often causes much computation and storage cost.

Recently, cutting edge methods, such as CNN, have been widely used in the agricultural field due to its high classification accuracy. From the viewpoint of the identification of plant leaves, Lee's method has high performance [11]. They extracted features from 44 types of plant leaves and succeeded in identifying plant leaves with 99.5% accuracy by placing a multi-layer perceptron in the identification part. Mohanty et al. have verified the utility of CNN for the diagnosis of plant diseases using multiple CNN architectures [12]. They conducted a verification experiment using a dataset containing 54,306 plant leaf images with 14 crop species and 26 types of diseases for 38 classes. They achieved a classification accuracy of 99.35%. Liu et al. [13] classify four types of apple leaf diseases using CNNs. They achieved an average classification accuracy of 97.62%. Jiang et al. [14] proposed a method that analyzes the wheat diseases using the WDD2017 dataset containing 9,230 images. The method diagnosed seven different types of wheat diseases by using a full convolutional VGG-based neural network. They achieved an average recognition accuracy of 97.5%. Fuentes et. al [15] proposed a method that detects and diagnoses lesions in tomato images based on VGG-net and Faster region-based CNN. They succeed in achieving a maximum of 0.83 mean average precision. The recent approaches demonstrated promising realization of plant disease diagnosis utilizing a large-scale dataset. However, extensive power and memory processing is still the issue remain unsolved.

The existing approaches for the diagnosis of leaf diseases use large-scale datasets taken under favorable conditions. However, one of the problems is that images to be taken in the agricultural fields are not always under favorable conditions. For example, the plants are often planted in a narrow susceptible to leaf overlap. In addition, the shooting environment is effected by many external factors, such as the effect of sunlight, shading from the sun, and shaking of the leaves by the wind. This study uses efficient image feature descriptors to implement an automatic plant disease diagnosis system that is independent of such external factors.

## III. Methodology

### A. Dataset

For creating an automatic diagnosis system, a dataset containing different plant leaf images is indispensable. For this research work, we used the PlantVillage dataset [16] available online on the internet, which consists of 54,393 images divided into 38 categories by species and disease. Figure 1 shows examples of some image samples of this dataset.

### B. Features Extraction and Description

In this paper, we compared five feature extraction methods for the diagnosis of leaf diseases, such as SIFT, SURF, ORB, KAZE, and AKAZE. We shortly describe the characteristics of these methods.

*1) SIFT:* Scale-invariant feature transform (SIFT) [17] finds interest points invariant to scale and rotation by searching the entire image locations and scales. SIFT uses linear diffusion to create a difference-of-Gaussian (DoG) and down-sampling the images. The algorithm detects local maxima and minima by choosing a pixel in the DoG image and compare it with the 26 neighborhood pixels around it in $3 \times 3$ regions of adjacent scales. Each pixel is investigated in several different scales, and the candidate keypoints are selected that gave the highest measurement in the frequency scale, causing the scale invariance. The algorithm then localizes these interest points. While localizing such interest points, there is a possibility of outliers due to either low contrast candidates or caused by noise or poorly localized candidates along the edges. Rejection is done by examining potential surrounding points of interest using a Harris corner detector that detects large gradients, i.e., derivatives, in all directions.

The algorithm then assigns the orientation to each interest point localized previously, which is useful for rotation invariance. An orientation histogram of the neighborhood, which contains gradients around the region of each keypoint, is essential for the rotation invariance. The histogram generates different peak values, but the peak with the dominant direction will be set as the orientation for the keypoint. If several peaks of the same magnitude are found, then multiple keypoint with different orientations, but the same location and scale will be created at that peak. Finally, the descriptions for each keypoint are computed using $16 \times 16$ neighborhood around the detected point. This neighborhood window is further divided into $4 \times 4$ blocks, and the 8 bin orientation histogram is created for each block. Then a 16 bin histogram of 128 dimensions in one long vector is concatenated. The final vector contains the descriptors of these keypoints.

*2) SURF:* DoG processing is the bottleneck of SIFT. Speeded up robust features (SURF) [18] accelerate the DoG processing by using box filtering with an integral image [19], [20], [21]. In this algorithm, both keypoint extraction and scale detection are approximated by combining Hessian-Laplace detection with box filtering.

*3) ORB:* Oriented FAST and rotated BRIEF (ORB) [22] is based on binary robust independent elementary features (BRIEF) descriptor [23] and features from accelerated segment test (FAST) detector [24]. Both methods are computationally efficient. ORB uses Fast and Harris corner detector to detect efficient keypoints. Since FAST is not rotation invariant, it uses the intensity centroid technique to make it rotation invariant. BRIEF descriptor in ORB has a weakness in orientation performance. It was improved by computing a rotation matrix using the orientation of patches. ORB then create a steer version of BRIEF according to the orientation.

*4) KAZE and Accelerated KAZE (AKAZE):* The idea behind the creation of KAZE and AKAZE is to detect and describe 2D features in a nonlinear scale-space extreme to obtain a better localization accuracy distinctiveness [25]. Gaussian blurring used in the other object recognition algorithms, such as SIFT, does not respect the natural boundaries of objects since image details and noise are smoothed to the same degree at all scale levels. Gaussian filtering is isotropic in the scale-space, the processing is performed by blurring the edge of the object, and thus local features cannot be easily detected.

Therefore, to make the blurring adaptive to image features, KAZE uses nonlinear diffusion filtering and additive operator splitting (AOS) to reduce the noise. Also, AKAZE defines a feature descriptor and uses a unique descriptor called modified-local difference binary (M-LDB).

## IV. Classification of Leaf Disease Images

After extracting features, we clustered the similar-looking features using k-means clustering. The goal of k-means clustering algorithm is the partition of $n$ local features $(X = [x_1, x_2, ..x_n])$ into $k$ clusters $(k \leq n)$ in order to minimize the residual sum of squares within each cluster.

$$\arg\min_C \sum_{i=1}^{k} \sum_{x_j \in C_i} \|x_j - \mu_i\|^2 \tag{1}$$

Equation (1) is the objective function for k-means clustering where $\mu_i$ is the mean of samples in $C_i$. The value selected for $k$ in our experiment is 250 to create a vocabulary of 250 visual words. Different values for $k$ were tried, but $k = 250$ is optimum enough that if we decrease or increase this value, the resultant accuracy decreases.

After creating the visual words, the next step is to calculate the most frequently occurring visual words. To do so, we used a technique called TF-IDF to evaluate the frequency of visual words to create a BoVW. The term TF (Term Frequency) means a visual word has a high frequency of occurrence in an image indicates that visual word can represent the content of the image well.

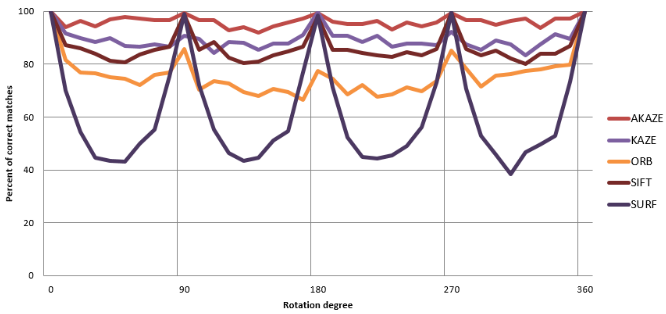
$$TF_{i,j} = \frac{n_{i,j}}{\sum_k n_{i,j}} \tag{2}$$

In (2), $n_{i,j}$ denotes the occurrence of visual words in the image and the denominator is the total number of visual words in the image. Inverse Document Frequency (IDF) is calculated that retrieve some common words that appear very frequently in each image, but these words do not represent the content of the image well, so this part of the word should be given a lower weight.

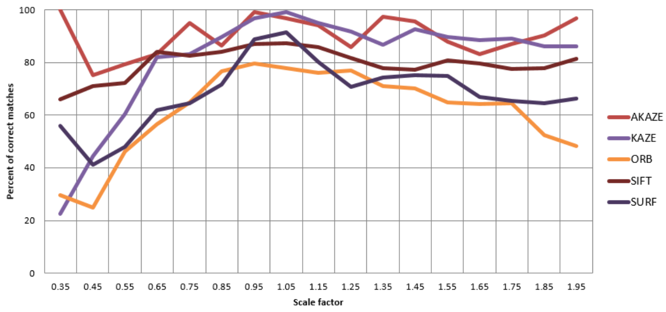
$$IDF_i = \log \frac{|D|}{|\{d_j : t_i \in d_j\}|} \tag{3}$$

In (3), $|D|$ is the total number of images $D = d_1, d_2, ...d_n$ whereas $|\{d_j : t_i \in d_j\}|$ is the number of total images where the word $t_i$ appears. Now the weight of TF and IDF is calculated by multiplying them with each other.

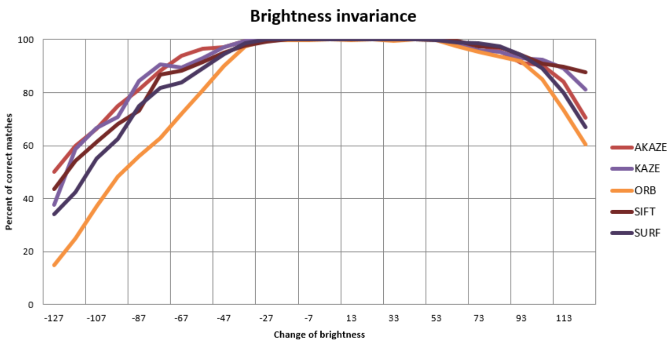
$$TFIDF_{i,j} = (TF_{i,j})\,(IDF_i) \tag{4}$$



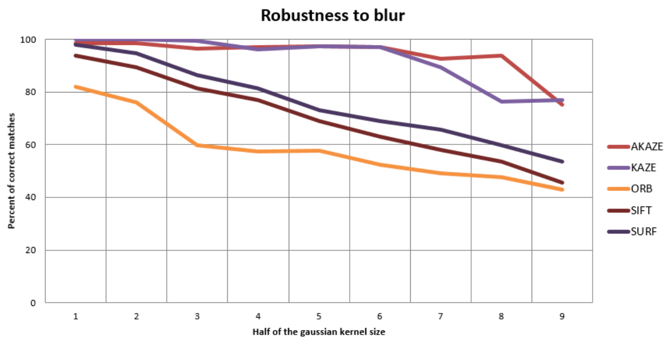

(a) Rotation invariance test




(b) Scale invariance test




(c) Brightness invariance test




(d) Blur invariance test

Fig. 2: Invariance test cases for each descriptor

TABLE I: Experimental results with the use of each image feature descriptor for 680 images.

| Descriptor | Time per frame [s] | Total keypoints | Accuracy [%] |
|---|---|---|---|
| SIFT | 2.062 | 899845 | 87.25 |
| SURF | 1.762 | 789640 | 85.73 |
| ORB | 1.320 | 392764 | 54.64 |
| KAZE | 1.132 | 538033 | 63.23 |
| AKAZE | 0.287 | 572143 | 81.56 |

The weight of TF-IDF is then assigned to BoVW vector that ultimately creates the BoVW. The system then uses k-means method again to classify these bag of features into their categories, which classified the leaf diseases as a result.

## V. Experimental Results and Analysis

In this section, we compare each method. The experimented code was implemented in Python with OpenCV. The CPU was Intel Core-i5 4690 3.50 GHz, and the code was compiled with Visual Studio 2017.

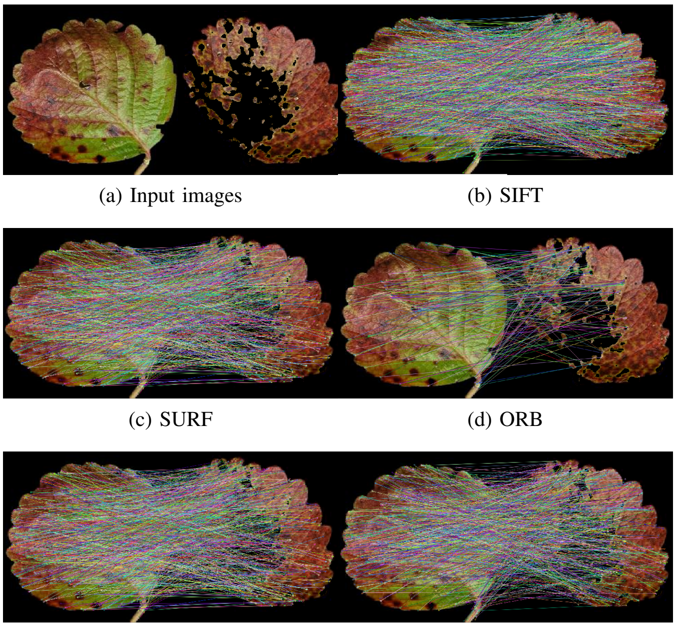
We compared the performance of each feature descriptor to evaluate their robustness against the brightness, blur, rotation, and scale-invariance, as shown in Fig. 2. The input image used for the test cases was Lena, as it has more variations than a leaf image to consider. In each test case, AKAZE, KAZE, and SIFT, respectively shows good accuracy.

Figure 3 shows the results of keypoint matching between two different images but the same condition type. In this test case, SIFT, SURF, and AKAZE show maximum numbers of keypoints matches.

We also compared the performance of each method based on the computation cost and accuracy in Table 1. The experimental results show that the computational speed of AKAZE is about seven times faster than that of SIFT and SURF and five times faster than that of ORB. The classification accuracy achieved with SIFT is 87.25 %, with SURF 85.73%, and using AKAZE is 81.56%, while the other two descriptors had the lower performance than SIFT, SURF, and AKAZE.

AKAZE showed better performance in terms of robustness to various variances, computation cost, and also the classification accuracy of AKAZE is slightly comparable to SIFT and SURF. Therefore, we justify AKAZE as the most efficient feature descriptor for the automatic diagnosis of leaf diseases.

Note that the famous approach of SIFT is fully accelerated by using hardware [26], [27]. Further, recent signal processing can accelerate the fundamental processing of Gaussian filtering by using sliding transform [28], [29]. The technique is utilized for SSIM computing [30], which has similar computational scheduling in DoG computation. Using adequate implementation can reduce the cost of SIFT, which has the highest accuracy.



(a) Input images (b) SIFT

(c) SURF (d) ORB

(e) KAZE (f) AKAZE

Fig. 3: Comparison based on Keypoint matching.

## VI. Conclusion

This paper proposes a method for the automatic diagnoses of leaf diseases using efficient image feature descriptors and a k-means clustering algorithm. We extracted the features from the lesion on the images using SIFT, SURF, ORB, KAZE, and AKAZE, respectively. We then used the K-means clustering algorithm to cluster the extracted features into visual features. The obtained visual features are then grouped by TF-IDF technique for creating a bag of visual features. We then used the K-means clustering algorithm again to classify these visual words, which in turn classify the leaf diseases. However, the major limitation of this study is the proposed method map the irrelevant interest points during the clustering process, which may reduce the accuracy of the system. Moreover, the proposed method manually learn the parameters of the encoders during the clustering.

## References


[1] D. V. R. Reddy, M. R. Sudarshana, M. Fuchs, N. C. Rao, and G. Thottappilly. *Chapter 6 - Genetically Engineered Virus-Resistant Plants in Developing Countries: Current Status and Future Prospects*. Academic Press, 2009.

[2] R. E. F. Matthews. *Diagnosis of Plant Virus Diseases*. CRC Press, 2019.

[3] E. Fujita, H. Uga, S. Kagiwada, and H. Iyatomi. A practical plant diagnosis system for field leaf images and feature visualization. *International Journal of Engineering and Technology*, 7:49–54, 2018.

[4] K. Alsabti, S. Ranka, and V. Singh. An efficient k-means clustering algorithm. In *Proc. First Workshop High Performance Data Mining*, 1998.

[5] G. Salton, A. Wong, and C.-S. Yang. A vector space model for automatic indexing. *Communications of the ACM*, 18(11):613–620, 1975.

[6] C. F. Tsai. Bag-of-words representation in image annotation: A review. *International Scholarly Research Notices*, 2012, 2012.

[7] K.-Y. Huang. Application of artificial neural network for detecting phalaenopsis seedling diseases using color and texture features. *Computers and Electronics in Agriculture*, 57(1):3–11, 2007.

[8] S. Kai, L. Zhikun, S. Hang, and G. Chunhong. A research of maize disease image recognition of corn based on bp networks. In *Proc. IEEE International Conference on Measuring Technology and Mechatronics Automation (ICMTMA)*, volume 1, pages 246–249, 2011.

[9] D. Al Bashish, M. Braik, and S. Bani-Ahmad. A framework for detection and classification of plant leaf and stem diseases. In *Proc. IEEE International Conference on Signal and Image Processing (ICSIP)*, pages 113–118, 2010.

[10] S. Arivazhagan, R. N. Shebiah, S. Ananthi, and S. V. Varthini. Detection of unhealthy region of plant leaves and classification of plant leaf diseases using texture features. *Agricultural Engineering International: CIGR Journal*, 15(1):211–217, 2013.

[11] S.-H. Lee, C.-S. Chan, P. Wilkin, and P. Remagnino. Deep-plant: Plant identification with convolutional neural networks. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 452–456, 2015.

[12] S. P. Mohanty, David P. H., and Marcel S. Using deep learning for image-based plant disease detection. *Frontiers in Plant Science*, 7, 2016.

[13] B. Liu, Y. Zhang, D. He, and Y. Li. Identification of apple leaf diseases based on deep convolutional neural networks. *Symmetry*, 10(1):11, 2018.

[14] J. Lu, J. Hu, G. Zhao, F. Mei, and C. Zhang. An in-field automatic wheat disease diagnosis system. *Computers and Electronics in Agriculture*, 142:369–379, 2017.

[15] A. Fuentes, S. Yoon, S. Kim, and D. Park. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors*, 17:2022, 2017.

[16] David. P. H. and Marcel S. An open access repository of images on plant health to enable the development of mobile disease diagnostics, 2015.

[17] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[18] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool. Speeded-up robust features (surf). *Computer Vision and Image Understanding*, 110(3):346–359, 2008.

[19] F. C. Crow. Summed-area tables for texture mapping. In *Proc. ACM SIGGRAPH*, pages 207–212, 1984.

[20] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 511–518, 2001.

[21] N. Fukushima, Y. Maeda, Y. Kawasaki, M. Nakamura, , T. Tsumura, K. Sugimoto, and S. Kamata. Efficient computational scheduling of box and gaussian fir filtering for cpu microarchitecture. In *Proc. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, 2018.

[22] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. Orb: An efficient alternative to sift or surf. In *Proc. IEEE International Conference on Computer Vision (ICCV)*, pages 2564–2571, 2011.

[23] M. Calonder, V. Lepetit, C. Strecha, and P. Fua. Brief: Binary robust independent elementary features. In *Proc. European Conference on Computer Vision (ECCV)*, pages 778–792, 2010.

[24] E. Rosten and T. Drummond. Machine learning for high-speed corner detection. In *Proc. European Conference on Computer Vision (ECCV)*, pages 430–443, 2006.

[25] P. F. Alcantarilla, A. Bartoli, and A. J. Davison. Kaze features. In *Proc. European Conference on Computer Vision (ECCV)*, pages 214–227, 2012.

[26] F. Alhwarin, D. Ristić-Durrant, and A. Gräser. Vf-sift: very fast sift feature matching. In *Proc. Joint Pattern Recognition Symposium*, pages 222–231, 2010.

[27] C. Griwodz, L. Calvet, and P. Halvorsen. Popsift: a faithful sift implementation for real-time applications. In *Proc. ACM Multimedia Systems Conference*, pages 415–420, 2018.

[28] K. Sugimoto and S. Kamata. Fast gaussian filter with second-order shift property of dct-5. In *Proc. IEEE International Conference on Image Process (ICIP)*, 2013.

[29] K. Sugimoto, S. Kyochi, and S. Kamata. Universal approach for dct-based constant-time gaussian filter with moment preservation. *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2018)*, pages 1498–1502, 2018.

[30] T. Sasaki, N. Fukushima, Y. Maeda, K. Sugimoto, and S. Kamata. Constant-time gaussian filtering for acceleration of structure similarity. In *Proc. International Conference on Image Processing and Robot (ICIPRoB)*, 2020.