

# Extraction of Point-of-Interest in Multispectral Images for Face Recognition

Kossi Kuma KATAKPE

Institut de Mathématiques  
et de Sciences Physiques (IMSP)  
Université d'Abomey Calavi (UAC)  
Porto Novo, Benin

Lyes AKSAS

UFR Science & Technology  
IEM Department, ImVia  
Laboratory, Color, Sensor &  
Multispectral Imaging  
Burgundy University  
Dijon, France

Diarra MAMADOU

UFR Mathématiques et  
Informatique Université  
Félix Houphouët-Boigny,  
Abidjan, Côte d'Ivoire

Pierre GOUTON

UFR Science & Technology  
IEM Department, ImVia  
Laboratory, Color, Sensor &  
Multispectral Imaging  
Burgundy University  
Dijon, France

**Abstract**—Security systems in companies, airports, enterprises, etc. face numerous challenges. Among the major ones there is objects or face recognition. The problem with the robustness of recognition systems that usually affects color images nowadays can be addressed by multispectral image acquisition in the near infrared range with cameras equipped with new high performance sensors able to take images in dark or uncontrolled environments with much more accuracy. Multispectral CMOS (Complementary Metal Oxide Semi-conductor) sensors in a single shot record several wavelengths that are isolated and allow very specific analyses. They are equipped with new acquisition methods and provide observations that are more accurate. The current generation of these imaging sensors involve scientific and technical interest because they provide much more information than those that operate in visible range; precise nature and spatio-temporal evolution of the areas need to be analyzed. In this study, multispectral images acquired by camera equipped with a hybrid sensor operating in near infrared has been used. This camera is built in the ImViA laboratory of the University of Bourgogne as part of the European project EXIST (EXtended Image Sensing Technologies). The process involved in image acquisition, image mosaicing and image demosaicing by using mosaic filters. After acquisition process the interest points be extract in these bands of images in order to know how information is shared out all over the bands. The results were satisfactory because information is spread all over the images bands and the algorithms used also have detected many interest points. Based on the results, a large database can be set up for a face recognition system building.

**Keywords**—Multispectral image; hybrid sensor; image mosaicing; image demosaicing; mosaic filter

## I. INTRODUCTION

Nowadays, in almost every sector, security and attack problems have become a crucial challenge. Biometric imaging systems are appearing as a promising solution to increase levels of security. These biometric systems are mostly based on grayscale images, color images and spectral reflectance. But these systems still face tremendous difficulties when recognizing objects or faces.

In fact, conventional digital color cameras that generally operate in the visible spectrum seem to be limited in many situations where more information is needed and acquisition conditions are still difficult such as making acquisitions under a cloudy sky, while information beyond the visible range is required such as plants that emit in infra-red range, or when

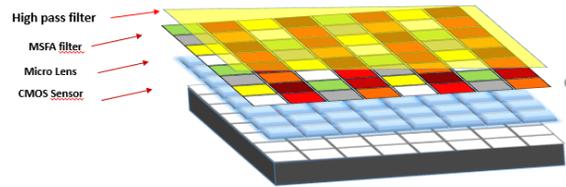
acquiring an image with more accuracy is needed, or when the calibration of the acquisition system is needed, or when making acquisitions in uncontrolled or dark environments is necessary, etc. Several studies have shown that images acquired in the visible spectrum present less information than those taken in Infra-red range [1], [2], [3], [4]. In addition Samuel ORTEGA et al. found that multispectral imaging technique able to obtain both spatial and spectral information within and beyond the human visual sensitivity, capture information regarding different wavelengths [5]. To overcome some of these problems, MAMADOU Diarra et al, in their studies on multispectral images, have merged information from visible range and thermal infra-red to increase information in the image [6], [7], [8], [9], [10]. They also presented multispectral imaging and especially merging of information from the visible and infrared as a very promising alternative for image recognition. Moreover, Xingbo Wang et al. also found that for having more accuracy, it is necessary to make good choice of spectral characteristics of the camera's filters [11], [12], [4]. Their results show that the filter bandwidth had an influence on the accuracy of the reflectance estimation. However, multispectral imaging with cameras equipped with hybrid sensors, operating in the field of Near infra-red are much more efficient and can capture more information [1]. For example to verify that a fingerprint comes from a living finger and not a copy of that finger, it is obvious that the near infra-red range is the best fit since veins are visible through the skin in this area, Laura Rey-Barroso et al. introduced Near infra-red (NIR) multispectral imaging system to evaluate deeper skin layers thanks to higher penetration of photon at this wavelengths [13]. This hybrid system, integrated into a camera with dedicated hardware and software computations, allows a performance in real-time application with 30 fps. It also provides finer detail analysis in recognition systems.

In the context of the problems listed above, a camera equipped with a hybrid sensor has been proposed, in which an optimization of the spectral bands from 680 nm to 950 nm (NIR) has been set up as optimal bands [14]. This camera that captures images on eight bands allowed good resolution for images. These images were used in order to extract the characteristics for the recognition so that performances of recognition systems could be improved. Based on the results, a large database of images taken in the NIR can be set up.

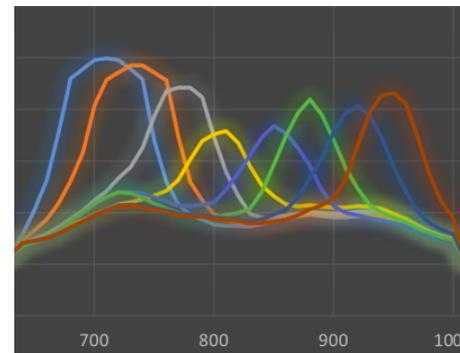
In the following, the process consists of acquiring an image that will be mosaicked before being transmitted by the camera. Once we have the image from the camera, we proceed to the separation of the different spectral bands using binary masks (Fig. 2). After separation each spectral band contains only one spectral component. In order to get a complete image, these image bands have to be interpolated. This process is called demosaicking and it allows us to have complete image bands Fig. 3, Fig. 4 shows the entire acquisition process. After this last step, the interest points will be extracted in these image bands for tests.



(a) Hybrid Sensor



(b) Hybrid Sensor Blocks



(c) Hybrid Sensor Spectral Response

## II. MULTISPECTRAL HYBRID SENSOR

### A. Hybrid Sensor

In this work a camera equipped with a hybrid sensor was used. This sensor was integrated into camera with a dedicated hardware unit, allowing the operation in real-time applications with 30 fps. In order to provide an optimal solution for the loss of spatial resolution inherent to MSFA, specific algorithms have been developed for multispectral demosaicking. The CMOS sensor is the physical element whose performance impacts on the quality of the final system. This sensor has been chosen respect to several criteria: ✓ Minimum pixel size is  $5\mu\text{m}$ ;

✓ The CMOS sensor resolution should be high enough to compensate the loss related to the MSFA system;

✓ The spectral sensitivity of the sensor must be extended to the near infra-red.

Taking into account the specifications above, our choice fell on the viimagic 9220H sensor. This sensor was provided by Grass Valley. Some modifications have been introduced in order to improve the final sensor.

The advantage of using CMOS sensors is that its manufacturing is much cheaper than CCD (Charged coupled Device) sensors. Furthermore, CMOS sensors consume less energy [15]. The ease of access to pixels available in CMOS sensors allows great flexibility for real-time data processing. All the above mentioned advantages bring about smaller systems, lower power consumption and lower manufacturing cost [15], [16]. As a result, we have chosen to use CMOS sensors rather than CCD sensors [17], [18]. The result of the mounted hybrid sensor and its spectral response are presented in Fig. 1.

### B. Multispectral Images

A multispectral (MS) image is an image acquired by a sensor that operates in several spectral bands; it can be defined as an image where each pixel contains essentially information on the reflectance of the scene. It is represented by the matrix of pixels as follows:

$$M = (M_1, M_2, \dots, \dots, M_j)$$

Where  $M_j$  is the associated matrix of  $j^{\text{th}}$  band of image.

$$M_j = \begin{pmatrix} x_{11}^j & x_{12}^j & \dots & x_{1n}^j \\ x_{21}^j & x_{22}^j & \dots & x_{2n}^j \\ \dots & \dots & \dots & \dots \\ x_{m1}^j & x_{m2}^j & \dots & x_{mn}^j \end{pmatrix}$$

Let  $I$  be a MS image, a pixel of the image is noted  $P(x, y)$ , where  $x$  and  $y$  are the coordinates of the pixel  $P$ . Each pixel  $P$  is associated to a point  $I(x, y, k)$  defined in a  $K$ -dimensional space ( $K$  being the number of component), and  $I_{(x,y)}^k$ ,  $k \in \{1, 2, \dots, K\}$  represents the value of each component. Therefore, for a multispectral image one needs  $k$  components plans  $I^k$ ,  $k \in \{1, 2, \dots, K\}$ . In this study  $K = 8$ , called an 8-band multispectral image.

III. MATERIAL AND METHODS

A. Mosaic Filters

Mosaic Filters are filters presented as a matrix where each filter is associated with a specific spectrum. These filters make it possible to divide finely the spectrum and thus to differentiate the bands. In this work, a set of 8 filters based on the principle of Fabry-Perot has been used. Table I illustrates the response of each of the eight filters. The resulting distribution of MSFA (Multispectral Filter Array) moxel is indicated in Table II.

TABLE I. FILTER BANDS RESPONSES

Bands	$\lambda$ (nm)	$\delta\lambda$ (nm)
P1	717	32
P2	751	32
P3	776	31
P4	810	31
P5	835	31
P6	870	30
P7	895	31
P8	930	31

TABLE II. SPATIAL DISTRIBUTION OF A MOXEL

P1	P5	P2	P6
P7	P3	P8	P4
P2	P6	P1	P5
P8	P4	P7	P3

B. Mosaic Images

Image Mosaicing is a technique that allows building an image by superimposing successive images by registration [16], [19]. It can therefore be defined as the process of assembling different images of the same scene to form a single image [20]. The aim of mosaic creation is to visualize a large area on a single image under perspective projection. One of its applications is the construction of large aerial and satellite images of small photographs collections [21].

C. Strips Extraction

There are many algorithms used for strips extraction [20], [22]. In this work we have chosen to multiply the mosaic image by different binary masks  $M^k_{(x,y)}$ ,  $k \in \{1, 2, \dots, K\}$  [20] which divides the mosaic image into  $K = 8$  components. These masks have the value 1 at the positions where the pixel is available, and 0 at the other positions. Each component plan is obtained by multiplying the mosaic image term by term by the corresponding  $M^k$  mask.

By multiplying the mosaic image with each mask, we obtain 8 uncorrelated image plans (Fig. 2) on which only one spectral component is available. Each mask corresponds to an image plan.

$$I^k = I \odot M^k \tag{1}$$

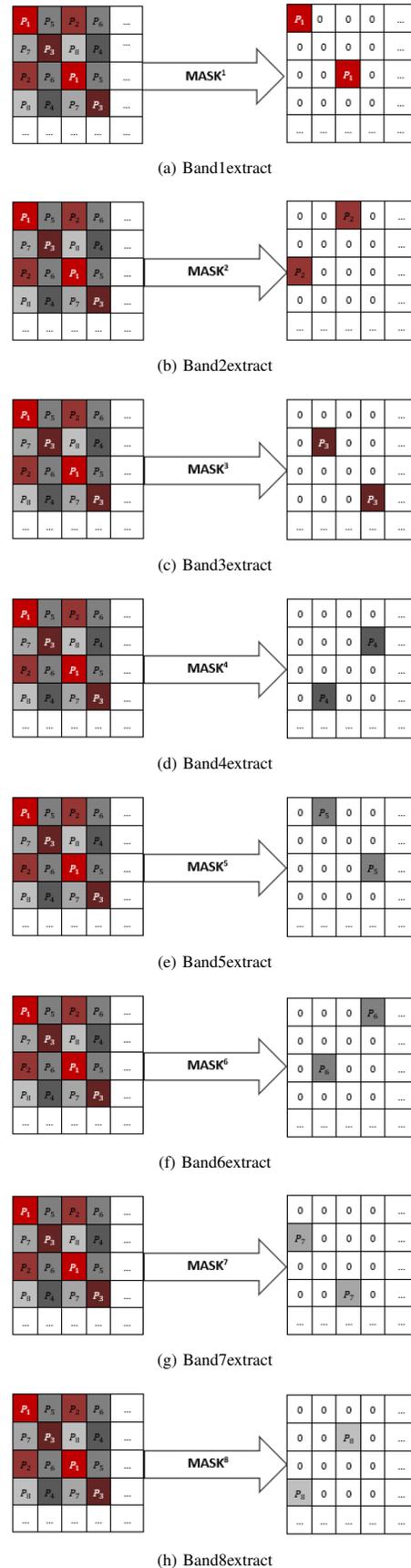


Fig. 2. Image Map after Applying Different Masks: After Applying Masks on Mosaic Image, Bands Obtained have only One Spectral Component

D. Multispectral Dematrixing by Bilinear Interpolation

After masks application, the resulting image plans contain only one spectral component. For complete image reconstruction, the missing pixels have to be interpolate. This process is called multispectral image demosaicing [23], [24], [18], [25], [3]. Bilinear interpolation [26], [16], [27], can be interpreted as a process of two linear interpolations, one in each direction. Linear interpolations can be made in several directions.  $P(i, j)$  being the missing pixel at the position  $(i, j)$ , we have:

- Diagonally:

$$P(i, j) = \frac{1}{4} \sum_{(m,n)=(-1,-1),(-1,1),(1,-1),(1,1)} p(i + m, j + n), \quad (2)$$

- Vertically:

$$P(i, j) = \frac{1}{2} \sum_{(m,n)=(-1,0),(1,0)} p(i + m, j + n), \quad (3)$$

- Horizontally:

$$P(i, j) = \frac{1}{2} \sum_{(m,n)=(0,-1),(0,1)} p(i + m, j + n), \quad (4)$$

The interpolation or demosaicing of a mosaic image is a method that estimates the missing pixel on different (chromatic) channel of the mosaic image. Several algorithms have been designed for image demosaicing [28], [29], [30]. The method used, consists of applying convolution filter  $H$  on each band of the image obtained [20]. This filter is fixed so that the contribution of the neighbors in the pixel estimation of missing level in this pixel depends on the spatial distance separating the neighbor from the central pixel. Given that the pixels have the same structure, the same filter as shown in Mihoubi's work [20] is used. Interpolated bands have been shown in Fig. 3. The acquisition process is depicted in Fig. 4.

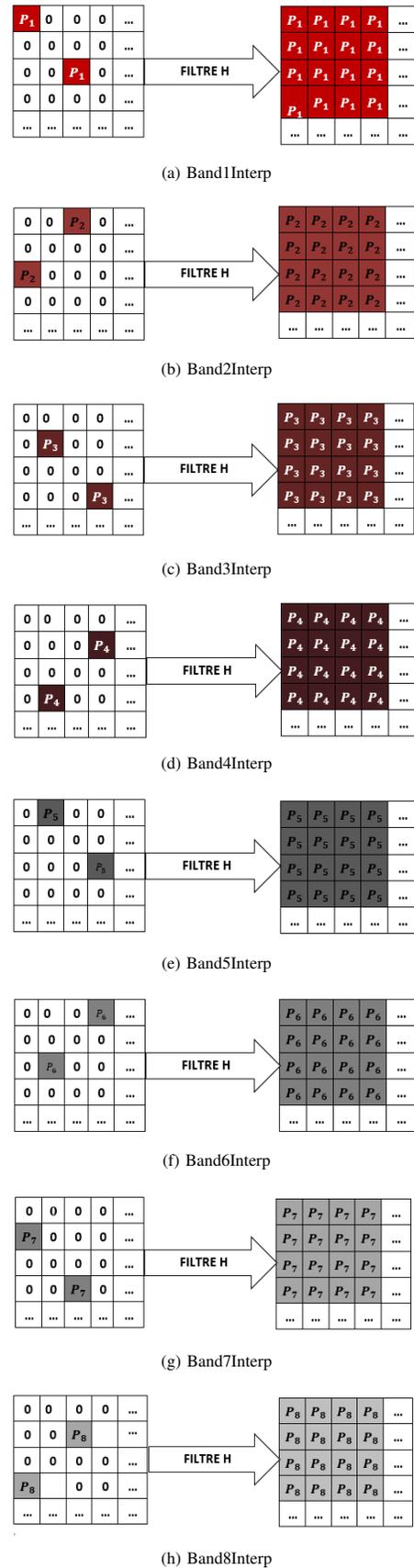


Fig. 3. Bands after Interpolation: The One Component Bands have been Interpolated using Filter H, to a Complete Image Bands

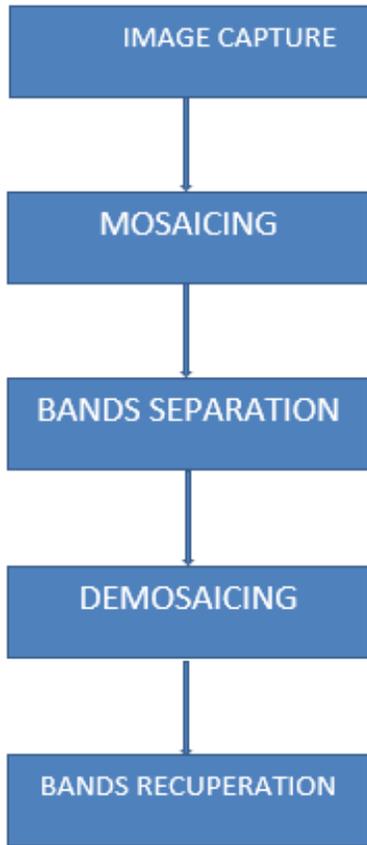


Fig. 4. Acquisition Process: The Process from Acquisition to the Last Band Recuperation

$$I^k = I'^k \odot H \quad (5)$$

$I^k$ , is the interpolated image band.

$$H = \frac{1}{9} \begin{pmatrix} 1 & 2 & 3 & 2 & 1 \\ 2 & 4 & 6 & 4 & 2 \\ 3 & 6 & 9 & 6 & 3 \\ 2 & 4 & 6 & 4 & 2 \\ 1 & 2 & 3 & 2 & 1 \end{pmatrix}$$

#### E. Point of Interest

A point of interest in an image is an area of pixel having remarkable properties often expressed by abrupt changes in intensity. They are regions of the image rich in terms of local information content and stable under affine transformations and illumination variations. In an image there must be few points whose local descriptors are similar [31].

#### IV. EXTRACTION OF INTEREST POINTS

The feature extraction methods are based on Scale Invariant Feature Transform (SIFT). The SIFT detector [32] is the best known of the detectors. This method combines a detector with a descriptor. SIFT's point of interest detection is based on a DoG (Difference of Gaussian), and has several versions. These algorithms are used in several contexts as multispectral imaging, face recognition under different criteria so that the

performance of such a feature extraction kernel be able to extract the parameters [33]. It should be remembered that the SIFT method is based on the determinant of the Hessian matrix.

$$H(x, \delta) = \begin{pmatrix} L_{xx}(x, \delta) & L_{xy}(x, \delta) \\ L_{xy}(x, \delta) & L_{yy}(x, \delta) \end{pmatrix} \quad (6)$$

where  $L_{xx}(x, \delta)$  is the convolution of the second order Differential of the Gaussian (DoG), which is the same for  $L_{xy}(x, \delta)$  and  $L_{yy}(x, \delta)$  to reduce the computation complexity of the determinant that uses the approximation of the wavelets of Haar.

$$H_{approx}(x, \delta) = \begin{pmatrix} D_{xx}(x, \delta) & D_{xy}(x, \delta) \\ D_{xy}(x, \delta) & D_{yy}(x, \delta) \end{pmatrix} \quad (7)$$

By using the expression of the integral image :

$$I_{\Sigma}(x) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i, j) \quad (8)$$

it can be deduced that:

$$\det(H_{approx}) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (9)$$

#### V. IMPACT OF THE WORK

The acquisition with hybrid sensors is made to measure the accuracy and the response of the resulting optical filters, which can ensure the accuracy and quality of the obtained multispectral images. These Multispectral images of the hybrid sensor can be less good, because of the demosaicking that compute the neighboring pixels which sometimes generate approximations. But the hybrid sensors are adequate for making snapshot acquisitions in real time application and it use in the case of this work for detecting faces in real time. The multispectral images from a filter wheel camera are very good quality [34], no approximation in the calculations, however, it is impossible to make the detection in real time. With this new camera, a multispectral images database will be set up. When the database contains enough images, a Deeplearning solution will be proposed in future work, as many research projects are moving towards this solution. In 2019 Shaukat Hayat et al. [35], proposed to use deep CNN-based features for Hand-Drawn sketch recognition via Transfer Learning Approach. Xi-ang Wang et al. introduced also method of privacy-preserving face recognition [36] where the convolutional neural network is used for face feature extraction. Moreover Bogdan BELEAN et al. [37] use CNN (Convolutional Neural Network) for images segmentation.

#### VI. PRESENTATION OF DETECTORS AND DESCRIPTORS

Face or shape recognition techniques require some tools such as detectors and descriptors that are complementary tools of object recognition.

### A. Detectors

Point-of-interest detection is a preliminary step in many computer vision processes. Detectors are used to isolate areas of interest in an image. For twenty years, several interest-point detectors have been developed. Schmid and Mohr compared the performance of several of these detectors. According to Schmid et al. [38], the most popular point-of-interest detector is the Harris detector [39]. The Harris corner detector was proposed by C. Harris and M. Stephens [40]. This easily detects the point of interest through a small window by moving this window in any direction. The Harris corner detection algorithm is performed by calculating the gradient of each pixel. Then, if the gradient values in the two directions are both large, the pixel is assumed to be a corner. Our experiences have been done using KAZE, Harris, ORB. KAZE, ORB, which are at the same time detectors and descriptors [41].

ORB (Oriented FAST and Rotated BRIEF) was introduced by Rublee et al. [42]. The Oriented Fast and Rotated Brief algorithm is based on the BRIEF keypoint descriptor and the FAST keypoint detector since both algorithms are computationally fast. It was presented in 2011 to provide a fast and efficient alternative to SIFT [43]. It is a variant of BRIEF to fill the lack of rotational invariance of it. The ORB method calculates a local orientation using an intensity centroid, which is defined as a weighted average of the pixel intensities in the local patch assumed not to coincide with the center of the entity.

The KAZE algorithm was developed in 2012 and it is in the public domain. The name comes from the Japanese word kaze which means wind and makes reference to the flow of air ruled by nonlinear processes on a large scale [44], [43]. For object recognition KAZE follows mainly the same steps as SIFT but with some differences in each step. KAZE algorithm [44], [45], instead of using DoG use AOS (Additive Operator Splitting) method and the Hessian matrix detector for blobs detection (DoH : Determinant of the Hessian) [43], [46].

$$L_{Hessian} = (\sigma)^2(L_{xx}L_{yy} - (L_{xy})^2) \quad (10)$$

where  $L_{xx}(x, \delta)$  is the convolution of the second order Differential of the Gaussian (DoG), which is the same for  $L_{xy}(x, \delta)$  and  $L_{yy}(x, \delta)$ .

### B. Descriptors

After detecting points of interest, descriptors are used to describe them. They analyze neighborhood of each point to produce a characteristic vector of the interest point area. This vector is called the descriptor vector and in our work this vector describes 64 features. The description vector associated with a point of interest is a set of values extracted from the image in the local neighborhood of the position of the detected point [47]. This work have utilized the detectors and binary feature descriptors in Table III that provide high performance and compact data representation [38], [39].

TABLE III. SET OF DETECTORS AND DESCRIPTORS USED

Detectors	Descriptors
ORB	ORB
KAZE	BRISK
KAZE	FREAK
KAZE	KAZE
Harris	BRISK
Harris	FREAK
Harris	KAZE

BRISK (Binary Robust Invariant Scalable Keypoints) descriptor algorithm has been proposed by Leutenegger et al. [48]. In its detection, it uses the AGAST (Adaptive and Generic Corner Detection Based on the Accelerated Segment Test) [49] which is an improved variant of FAST [50]. FREAK (Fast Retina Keypoint) is a binary descriptor proposed by Alahi et al. [51]. Like BRISK (Binary Robust Invariant Scalable Keypoints), this descriptor uses a sampling model and a compensation method orientation. This is a variant of BRISK improved using a selection of pairs of templates. FREAK organizes sampling points analogous to the structure of the biological retina. For the description of the point of interest, the tools used are weighted Gaussians, the motif functioning as the retina and an orientation assignment is made for the description.

## VII. RESULTS

In this part, the results of different stages of this work will literally be presented: mosaicing of the image, the decomposition into 8 bands and the points of interest tests.

### A. Mosaicing of the Image and Decompositions into 8 Bands

The image obtained after mosaicing is represented in Fig. 5.

$P_1$	$P_5$	$P_2$	$P_6$	...
$P_7$	$P_3$	$P_8$	$P_4$	...
$P_2$	$P_6$	$P_1$	$P_5$	...
$P_8$	$P_4$	$P_7$	$P_3$	...
...	...	...	...	...

Fig. 5. Motif of Filter: The 4x4 Moxel Used

By applying different masks, mosaic image will be separated into 8 bands of images whose pixels contained a single color component, Fig. 2.

After the separation of the strips, these strips are demosaiced in order to attribute the rest of the color components to each pixel (Fig. 3). Fig. 6 represents a sample of final results on 8 bands after acquisition.

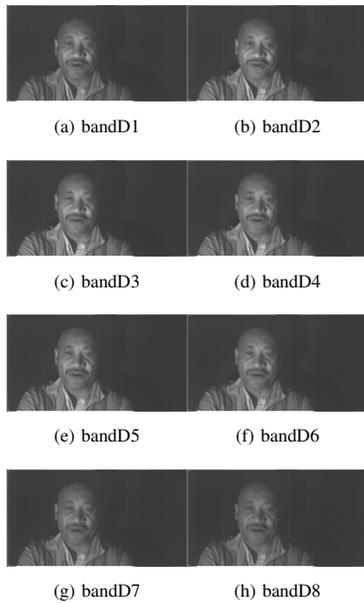


Fig. 6. Image on 8 Bands after Acquisition Process: Example of Bands Recuperated

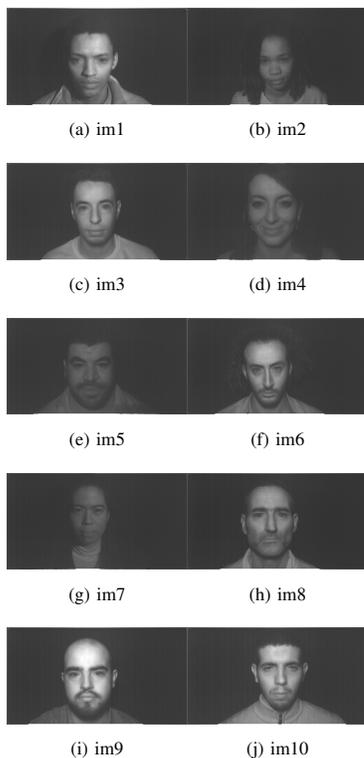


Fig. 7. Sample of Images of Database Used

### B. Test and Images Used

The images data base is set up with images taken by a camera equipped with a hybrid sensor that detect and acquire faces in real time. This camera takes images on 8 bands and can be used in real time applications. Most of the time, for

multispectral images, some bands contain less information than others. But the particularity of our camera is that the information is roughly spread over all the bands of image. The interest points are detected on all the 8 bands of images for all the algorithms mentioned above. Since the recognition is done on the face only, we used the algorithm of Viola Jones [52] to crop the face before the detection of those points of interest. This algorithm allows detecting only regions of interest. On the resulting image, different algorithms for the detection and the description of the points of interest has been applied. The tests have been done using Matlab v2020a with a sample of 30 images and the results are almost the same on each image. 10 images is used in this paper (Fig. 7). A sample of interest points by ORB/ORB has been shown in Fig. 8.

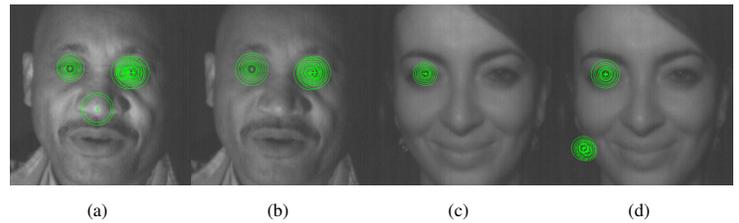


Fig. 8. Example of Key Points on Bands: Key Points with ORB / ORB Detector / Descriptor

Each detected point is described in 64 elements that are unique even if the position or the place of image acquisition change. The tests and results have been presented in Table IV.

Given the results in Table IV, the KAZE algorithm detects more points of interest than others and The results show that the points of interest are slightly more concentrated on the first bands for all the algorithms other than Harris algorithm which detects more interest points on the last bands. But in general, the information spread over all eight bands if the acquisition process has been successful.

### C. Entropy Test

Entropy in an image, makes it possible to measure the quantity of information contained in the image. In this work it allowed us to confirm that, information is spread over all the 8 bands. This entropy is computed by the formula below:

$$S = - \sum_{i=1}^{i \leq n} P_i \log(P_i) \quad (11)$$

Where  $P_i$  is the probability of each pixel occurrence.

Entropy tests have been done and the results are recorded in Table V and Fig. 9 represents the associated histogram.

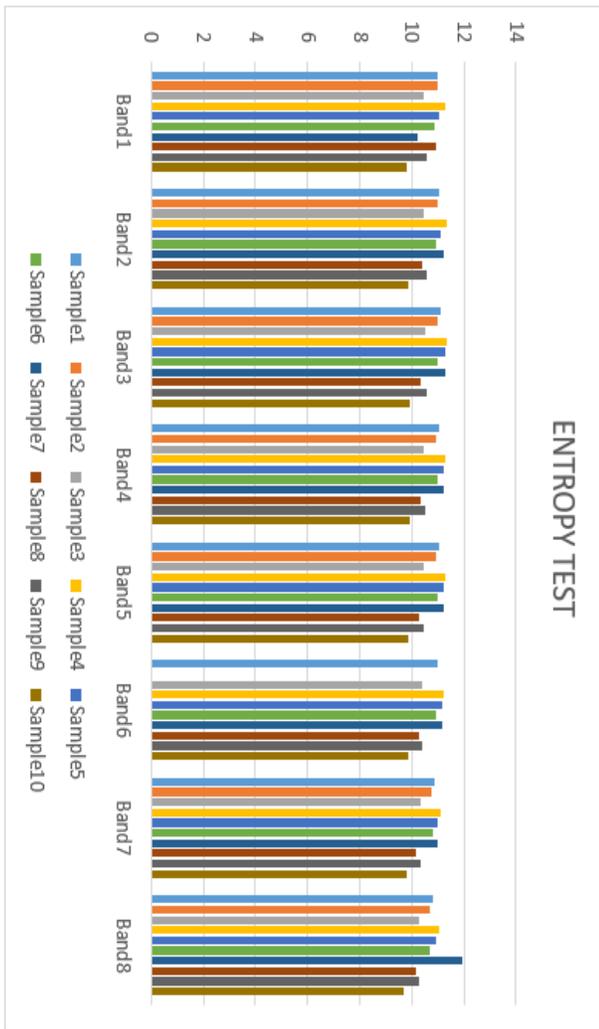


Fig. 9. Histogram of Entropy Test: Information Distributed over All the 8 Bands

### VIII. DISCUSSION

The results of Table IV allow to realize that when one used the pairs of detectors and descriptors Harris/FREAK, Harris/BRISK or Harris/KAZE, they did not detect enough interest points, especially on the first 6 bands. But on the seventh and eighth strips of some images, the number of interest points is quite enough. This phenomenon could be due to the lighting of the scene or the fact that this algorithm is not robust or suitable for these types of images. These algorithms have the worst performance in most cases in terms features detected. The algorithm ORB/ORB has correctly detected the points of interest on each band and each image. Therefore ORB/ORB demonstrates fairest precision with respect to the features, due to its performance one can say that it is better than the Harris/FREAK, Harris/BRISK and Harris/KAZE pairs which did not detect enough points and were also not stable for these types of images. For the pairs: KAZE/BRISK, KAZE/FREAK and KAZE/KAZE, the results were very satisfactory. The algorithms were performant at detecting high number of interest points. So, KAZE detector associate with BRISK, FREAK and KAZE descriptors are efficiency and robust enough for these

multispectral images. By making a comparison between these three pairs, KAZE/BRISK, KAZE/FREAK and KAZE/KAZE, in general, it is KAZE/KAZE detector/descriptor pair which is better represented than the others in terms of features detection. This result confirm that of Shaharya et al. [53], [54], where KAZE in term of interest points outperform ORB. Furthermore, Ratsimbazafy et al. have shown in [55] that in terms of detection: SURF and KAZE are in a high category compared to ORB. By talking about stability, KAZE is more stable. But on the other hand, in terms of execution time, KAZE is not as efficient as ORB. Shuvo Kumar Paul et al. studying detector pairs [56] find that the KAZE and AKAZE pairs perform better than other pairs. This algorithm has detected several points of interest on each image and on each strip. This proves that the KAZE/KAZE pair would be suitable for these multispectral images from the camera equipped with a hybrid sensor and operating in the near infrared range. One can notice that the interest points were almost spread on all the bands. For confirmation, we compute the entropy tests. This entropy tests in the Table V should show how the information in each image is distributed. These entropy results showed that for each image the information is almost roughly distributed over all the 8 bands except the 7th and eighth bands which detects less points of interest than the others. However, the results of the entropy tests confirm the interest points results.

Based on these results, a large database of images taken on 8 bands with this camera which operates in the NIR can be set up.

### IX. CONCLUSION

Security challenges of information systems keep increasing. Researchers have proposed different approaches and techniques. One of them is biometric imaging system. In recent years, studies have shown the limitation of this approach. This study focuses on multispectral (MS) imaging, primarily the use of the camera equipped with a hybrid sensor. This MS camera used in this work was built with a hybrid sensor, a Multispectral Filter Array (MSFA) mounted on a CMOS sensor that provided the best resolution for mosaic image, due to the small moxel used and due to the size of filter pitch ( $5 \times 5 \mu m^2$ ). This new camera system operates in the field of near infra-red in order to improve the process of object or image recognition. This study looked at the performance of this multispectral camera built in ImViA laboratory at the University of Bourgogne by extracting the points of interests on the bands of the multispectral images acquired by this camera. It also have been shown how to transform the obtained row images directly from the camera to a multispectral image through different steps namely: mosaicking, interpolation or demosaicing. Different descriptors have been used to extract interest points and the results were satisfactory. KAZE descriptor was the best and should be used to build recognition systems. However this project did not take place without difficulties: the filter based on the principle of Febray-Perrot is penalized by its secondary response; to compensate for the loss of sensitivity beyond 850 nm, a complex structure of moxel (6 x 6 pixels) simulation have been adopted, but this solution leads to none homogenous distribution. Ultimately, a regular distribution of the pixels (4x4) in the moxel is kept. The images are so contrasted that some lest robust algorithms are not able to extract interest points.

The Multispectral images of the hybrid sensor can be less good, because of the demosaicking that compute the neighboring pixels which sometimes generate approximations. But the hybrid sensors are adequate for making acquisitions and detecting faces in real time. Compared to the multispectral images from a filter wheel camera which are very good quality, no approximation in the calculations; however, it is impossible to make the detection in real time.

Future works will focus on building a large database of images acquired with this camera following by the usage of machine learning for recognition systems.

## REFERENCES

- [1] Wilder, Joseph and Phillips, P Jonathon and Jiang, Cunhong and Wiener, Stephen, Comparison of visible and infra-red imagery for face recognition, Proceedings of the Second International Conference on Automatic Face and Gesture Recognition, 182–187, 1996, IEEE
- [2] Liu, Jun and Kumar, Ajay, Detecting presentation attacks from 3d face masks under multispectral imaging, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pages=47–52, year=2018
- [3] Balbhimrao, Lamb Anupama and Khambete, Madhuri, No-Reference Perceived Image Quality Algorithm for Demosaiced Images, INTERNATIONAL JOURNAL OF ADVANCED COMPUTER SCIENCE AND APPLICATIONS, 7, 1, 285–289, 2016, Citeseer
- [4] He, Xin and Liu, Yajing and Ganesan, Kumar and Ahnood, Arman and Beckett, Paul and Eftekhari, Fatima and Smith, Dan and Uddin, Md Hemayet and Skafidas, Efstratios and Nirmalathas, Ampalavanapillai and others, A single sensor based multispectral imaging camera using a narrow spectral band color mosaic integrated on the monochrome CMOS image sensor, APL Photonics, 5, 4, 046104, 2020, AIP Publishing LLC
- [5] Ortega, Samuel and Halicek, Martin and Fabelo, Himar and Callico, Gustavo M and Fei, Baowei, Hyperspectral and multispectral imaging in digital and computational pathology: a systematic review, Biomedical Optics Express, 11, 6, 3195–3233, 2020, Optical Society of America
- [6] Diarra, Mamadou and Gouton, Pierre and Jérôme, Adou Kablan, Multi-spectral face recognition using hybrid feature, Electronic Imaging, 2017, 18, 200–203, 2017, Society for Imaging Science and Technology
- [7] Li, Wei and Dong, Mingli and Lu, Naiguang and Lou, Xiaoping and Zhou, Wanyong, Multi-sensor face registration based on global and local structures, Applied Sciences, 9, 21, 4623, 2019, Multidisciplinary Digital Publishing Institute
- [8] Zhang, Yongtao and Yin, Zhishuai and Nie, Linzhen and Huang, Song, Attention based multi-layer fusion of multispectral images for pedestrian detection, IEEE Access, 8, 165071–165084, 2020, IEEE
- [9] Grifoni, Emanuela and Campanella, Beatrice and Legnaioli, Stefano and Lorenzetti, Giulia and Marras, Luciano and Pagnotta, Stefano and Palleschi, Vincenzo and Poggialini, Francesco and Salerno, Emanuele and Tonazzini, Anna, A new Infrared True-Color approach for visible-infrared multispectral image analysis, Journal on Computing and Cultural Heritage (JOCCH), 12, 2, 1–11, 2019, ACM New York, NY, USA
- [10] Tariq Ahmad and Jinsong Wu and Imran Khan and Asif Rahim and Amjad Khan Human Action Recognition in Video Sequence using Logistic Regression by Features Fusion Approach based on CNN Features, International Journal of Advanced Computer Science and Applications, <http://dx.doi.org/10.14569/IJACSA.2021.0121103>, 2021, The Science and Information Organization, 12, 11,
- [11] Wang, Xingbo and Thomas, Jean-Baptiste and Hardeberg, Jon Y and Gouton, Pierre, Multispectral imaging: narrow or wide band filters?, JAIC-Journal of the International Colour Association, 12, 2014
- [12] Deng, Lei and Mao, Zhihui and Li, Xiaojuan and Hu, Zhuowei and Duan, Fuzhou and Yan, Yanan, UAV-based multispectral remote sensing for precision agriculture: A comparison between different cameras, ISPRS journal of photogrammetry and remote sensing, 146, 124–136, 2018, Elsevier
- [13] Rey-Barroso, Laura and Burgos-Fernández, Francisco J and Delpueyo, Xana and Ares, Miguel and Royo, Santiago and Malveyh, Josep and Puig, Susana and Vilaseca, Meritxell, Visible and extended near-infrared multispectral imaging for skin cancer diagnosis, Sensors, 18, 5, 1441, 2018, Multidisciplinary Digital Publishing Institute
- [14] Ansari, Keivan and Thomas, Jean-Baptiste and Gouton, Pierre, Spectral band Selection Using a Genetic Algorithm Based Wiener Filter Estimation Method for Reconstruction of Munsell Spectral Data, Electronic Imaging, 2017, 18, 190–193, 2017, Society for Imaging Science and Technology
- [15] Hizem, Walid, Capteur intelligent pour la reconnaissance de visage 2009, Evry, Institut national des télécommunications
- [16] Phelippeau, Harold, Méthodes et algorithmes de dématricage et de filtrage du bruit pour la photographie numérique, 2009 Université Paris-Est
- [17] Carlson, Bradley S, Comparison of modern CCD and CMOS image sensor technologies and systems for low resolution imaging, SENSORS, 2002 IEEE, 1, 171–176, 2002, IEEE
- [18] Frommen, Thorsten, Adaptive Homogeneity-Directed Demosaicing Algorithm, 2007
- [19] Capel, David, Image Mosaicing and super-resolution, Image Mosaicing and super-resolution, 47–79, 2004, Springer
- [20] Mihoubi, Sofiane and Losson, Olivier and Mathon, Benjamin and Macaire, Ludovic, Multispectral demosaicing using intensity-based spectral correlation, 2015 International Conference on Image Processing Theory, Tools and Applications (IPTA), 461–466, 2015, IEEE
- [21] Shum, Heung-Yeung and Szeliski, Richard, Systems and experiment paper: Construction of panoramic image mosaics with global and local alignment, International Journal of Computer Vision, 36, 2, 101–130, 2000, Springer
- [22] Miao, Lidan and Qi, Hairong and Snyder, Wesley E, A generic method for generating multispectral filter arrays, 2004 International Conference on Image Processing, 2004. ICIP'04., 5, 3343–3346, 2004, IEEE
- [23] Cao, Hong and Kot, Alex C, Accurate detection of demosaicing regularity for digital image forensics, IEEE Transactions on Information Forensics and Security, 4, 4, 899–910, 2009, IEEE
- [24] Hirakawa, Keigo and Parks, Thomas W, Adaptive homogeneity-directed demosaicing algorithm, IEEE Transactions on Image Processing, 14, 3, 360–369, 2005, IEEE
- [25] Baone, Gaurav A and Qi, Hairong, Demosaicking methods for multispectral cameras using mosaic focal plane array technology, Spectral Imaging: Eighth International Symposium on Multispectral Color Science, 6062, 60620A, 2006, International Society for Optics and Photonics
- [26] Alleysson, David and Susstrunk, Sabine and Héroult, Jeanny, Linear demosaicing inspired by the human visual system, IEEE Transactions on Image Processing, 14, 4, 439–449, 2005, IEEE
- [27] Longere, Philippe and Zhang, Xuemei and Delahunt, Peter B and Brainard, David H, Perceptual assessment of demosaicing algorithm performance, Proceedings of the IEEE, 90, 1, 123–132, 2002, IEEE
- [28] Brauers, Johannes and Aach, Til, A color filter array based multispectral camera, 12. Workshop Farbbildverarbeitung, 2006, Ilmenau
- [29] Chang, Edward and Cheung, Shiu-fun and Pan, Davis Y, Color filter array recovery using a threshold-based variable number of gradients, Sensors, Cameras, and Applications for Digital Photography, 3650, 36–43, 1999, International Society for Optics and Photonics
- [30] Li, Xin and Orchard, Michael T, New edge-directed interpolation, IEEE transactions on image processing, 10, 10, 1521–1527, 2001, IEEE
- [31] Calonder, Michael and Lepetit, Vincent and Strecha, Christoph and Fua, Pascal, Brief: Binary robust independent elementary features, European conference on computer vision, 778–792, 2010, Springer
- [32] Mikolajczyk, Krystian and Schmid, Cordelia, A performance evaluation of local descriptors, IEEE transactions on pattern analysis and machine intelligence, 27, 10, 1615–1630, 2005, IEEE
- [33] Albatat, Rami and Mulhem, Philippe and Chiamarella, Yves, Phrases Visuelles pour l'annotation automatique d'images., CORIA, 10, 3–18, 2010
- [34] Johannes Brauers and Nils Schulte and Til Aach, Multispectral Filter-Wheel Cameras: Geometric Distortion Model and Compensation Algo-

- rithms, IEEE Transactions on Image Processing, IEEE, 2008, Dec, 17, 2368–2380, 12,
- [35] Hayat, Shaukat and She, Kun and Yu, Yao and Mateen, Muhammad, Deep cnn-based features for hand-drawn sketch recognition via transfer learning approach, Editorial Preface From the Desk of Managing Editor..., 10, 9, 2019
- [36] Wang, Xiang and Xue, Heyu and Liu, Xuefeng and Pei, Qingqi, A privacy-preserving edge computation-based face verification system for user authentication, IEEE Access, 7, 14186–14197, 2019, IEEE
- [37] Belean, Bogdan, Active Contours Driven by Cellular Neural Networks for Image Segmentation in Biomedical Applications, STUDIES IN INFORMATICS AND CONTROL, 30, 3, 109–119, 2021, NATL INST R&D INFORMATICS-ICI PUBL DEPT, 8-10 AVERESCU BLVD, SECTOR 1 ...
- [38] Schmid, Cordelia and Mohr, Roger and Bauckhage, Christian, Comparing and evaluating interest points, Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271), 230–235, 1998, IEEE
- [39] Schmid, Cordelia and Mohr, Roger and Bauckhage, Christian, Evaluation of interest point detectors, International Journal of computer vision, 37, 2, 151–172, 2000, Springer
- [40] Mikolajczyk, Krystian and Schmid, Cordelia, Scale & affine invariant interest point detectors, International journal of computer vision, 60, 1, 63–86, 2004, Springer
- [41] Cowan, Bruce and Imanberdiyev, Nursultan and Fu, Changhong and Dong, Yiqun and Kayacan, Erdal, A performance evaluation of detectors and descriptors for UAV visual tracking, 2016 14th International Conference on Control, Automation, Robotics and Vision (ICARCV), 1–6, 2016, IEEE
- [42] Rublee, Ethan and Rabaud, Vincent and Konolige, Kurt and Bradski, Gary, ORB: An efficient alternative to SIFT or SURF, 2011 International conference on computer vision, 2564–2571, 2011, Ieee
- [43] Andersson, Oskar and Reyna Marquez, Steffany, A comparison of object detection algorithms using unmanipulated testing images: Comparing SIFT, KAZE, AKAZE and ORB, 2016
- [44] Alcantarilla, Pablo Fernández and Bartoli, Adrien and Davison, Andrew J, KAZE features, European Conference on Computer Vision, 214–227, 2012, Springer
- [45] Noble, Frazer K, Comparison of OpenCV's feature detectors and feature matchers, 2016 23rd International Conference on Mechatronics and Machine Vision in Practice (M2VIP), 1–6, 2016, IEEE
- [46] Ze-Ping, Cai and De-Gui, Xiao, Feature matching algorithm based on KAZE and fast approximate nearest neighbor search, 3rd International Conference on Computer Science and Service System, 2014, Atlantis Press
- [47] Harris, Christopher G and Stephens, Mike and others, A combined corner and edge detector., Alvey vision conference, 15, 50, 10–5244, 1988, Citeseer
- [48] Leutenegger, Stefan and Chli, Margarita and Siegwart, Roland Y, BRISK: Binary robust invariant scalable keypoints, 2011 International conference on computer vision, 2548–2555, 2011, Ieee
- [49] Mair, Elmar and Hager, Gregory D and Burschka, Darius and Suppa, Michael and Hirzinger, Gerhard, Adaptive and generic corner detection based on the accelerated segment test, European conference on Computer vision, 183–196, 2010, Springer
- [50] Rosten, Edward and Porter, Reid and Drummond, Tom, Faster and better: A machine learning approach to corner detection, IEEE transactions on pattern analysis and machine intelligence, 32, 1, 105–119, 2008, IEEE
- [51] Alahi, Alexandre and Ortiz, Raphael and Vanderghelynst, Pierre, Freak: Fast retina keypoint, 2012 IEEE Conference on Computer Vision and Pattern Recognition, 510–517, 2012, Ieee
- [52] Jones, Michael and Viola, Paul, Fast multi-view face detection, Mitsubishi Electric Research Lab TR-20003-96, 3, 14, 2, 2003
- [53] Tareen, Shaharyar Ahmed Khan and Saleem, Zahra, A comparative analysis of sift, surf, kaze, akaze, orb, and brisk, 2018 International conference on computing, mathematics and engineering technologies (iCoMET), 1–10, 2018, IEEE
- [54] Ramkumar, B and Laber, Rob and Bojinov, Hristo and Hegde, Ravi Sadananda, GPU acceleration of the KAZE image feature extraction algorithm, Journal of Real-Time Image Processing, 17, 5, 1169–1182, 2020, Springer
- [55] Ratsimbazafy, TH and Randriamitantoa, PA Comparaison de performances des détecteurs KAZE, ORB, SURF,
- [56] Paul, Shuvo Kumar and Hoseini, Pourya and Nicolescu, Mircea and Nicolescu, Monica, Performance Analysis of Keypoint Detectors and Binary Descriptors Under Varying Degrees of Photometric and Geometric Transformations, arXiv preprint arXiv:2012.04135, 2020

TABLE IV. KEY POINTS EXTRACTION WITH DIFFERENT METHODS

Images	Bands	ORB/ ORB	KAZE/ BRISK	KAZE/ FREAK	KAZE/ KAZE	Harris/ BRISK	Harris/ FREAK	Harris/ KAZE
Sample1	Band1	49	287	336	389	9	9	9
	Band2	44	344	408	480	4	4	4
	Band3	37	324	387	466	3	3	3
	Band4	37	291	355	425	5	5	5
	Band5	30	250	302	374	2	2	2
	Band6	27	227	282	343	3	3	3
	Band7	24	156	195	244	3	3	3
	Band8	18	124	156	204	2	2	2
Sample2	Band1	37	287	334	357	4	4	4
	Band2	34	251	300	335	3	3	3
	Band3	29	188	229	261	3	3	3
	Band4	25	141	172	202	7	7	7
	Band5	21	140	165	189	3	3	3
	Band6	24	102	133	151	5	5	5
	Band7	26	67	85	104	4	4	4
	Band8	17	46	62	75	3	3	3
Sample3	Band1	144	169	180	181	20	20	20
	Band2	139	178	191	195	24	24	24
	Band3	144	144	156	158	23	23	23
	Band4	145	128	134	136	25	25	25
	Band5	152	120	129	132	21	21	21
	Band6	126	99	106	107	19	19	19
	Band7	117	72	76	76	20	20	20
	Band8	95	55	58	58	15	15	15
Sample4	Band1	32	200	233	292	2	2	2
	Band2	30	201	248	310	7	7	7
	Band3	26	176	222	297	11	11	15
	Band4	20	151	185	262	2	2	2
	Band5	21	123	159	226	60	60	69
	Band6	13	113	143	197	2	2	2
	Band7	9	92	116	157	517	517	561
	Band8	8	60	78	116	106	106	121
Sample5	Band1	27	117	128	174	2	2	2
	Band2	24	132	172	230	2	2	2
	Band3	19	121	162	227	5	5	5
	Band4	20	111	142	209	8	8	8
	Band5	10	97	127	179	1	1	1
	Band6	11	77	98	152	341	341	361
	Band7	5	51	66	107	61	61	65
	Band8	5	43	50	80	172	172	192
Sample6	Band1	21	203	259	289	5	5	5
	Band2	21	217	285	330	3	3	3
	Band3	18	170	230	272	3	3	4
	Band4	23	151	211	246	18	18	18
	Band5	12	128	173	220	35	35	36
	Band6	12	114	160	189	7	7	7
	Band7	11	65	87	113	215	215	238
	Band8	11	42	68	85	253	253	270
Sample7	Band1	37	292	315	368	5	5	5
	Band2	40	308	345	402	5	5	5
	Band3	31	236	259	308	37	37	41
	Band4	31	190	203	239	4	4	4
	Band5	24	153	168	204	351	351	371
	Band6	23	133	144	180	9	9	9
	Band7	12	93	101	127	177	177	189
	Band8	11	63	69	86	278	278	293
Sample8	Band1	11	68	86	92	1	1	1
	Band2	11	69	86	94	2	2	2
	Band3	15	57	71	76	1	1	1
	Band4	9	45	56	59	1	1	1
	Band5	11	40	49	52	1	1	1
	Band6	12	33	41	44	2	2	2
	Band7	8	17	19	21	8	8	8
	Band8	6	16	18	18	1	1	1
Sample9	Band1	17	39	50	63	2	2	2
	Band2	6	25	40	55	4	4	4
	Band3	12	18	29	38	2	2	2
	Band4	13	16	20	26	3	3	3
	Band5	11	9	14	19	2	2	2
	Band6	5	5	11	17	3	3	3
	Band7	7	4	4	7	2	2	2
	Band8	7	4	4	6	11	11	11
Sample10	Band1	5	3	3	3	1	1	1
	Band2	7	3	3	3	3	3	3
	Band3	5	3	3	4	1	1	2
	Band4	3	2	4	2	15	15	17
	Band5	4	2	2	2	22	22	26
	Band6	3	2	2	2	45	45	49
	Band7	0	2	2	2	596	596	680
	Band8	0	2	2	2	273	273	295

TABLE V. ENTROPY TEST

Bands	Sample1	Sample2	Sample3	Sample4	Sample5	Sample6	Sample7	Sample8	Sample9	Sample10
Band1	10.9735	10.9679	10.4355	11.2693	11.0272	10.8871	10.2259	10.9361	10.5752	9.8139
Band2	11.0420	10.9775	10.4639	11.3286	11.1261	10.9460	11.2483	10.3736	10.5881	9.8807
Band3	11.1117	10.9752	10.4896	11.3205	11.2625	11.0012	11.2764	10.3494	10.5584	9.9089
Band4	11.0632	10.9234	10.4486	11.2925	11.2444	10.9879	11.2326	10.3120	10.4947	9.9028
Band5	11.0739	11.9339	10.4635	11.2739	11.2266	10.9748	11.2172	10.3028	10.4727	9.8868
Band6	11.0101	10.8779	10.4193	11.2393	11.1768	10.9426	11.1633	10.2719	10.4235	9.8687
Band7	10.8807	10.7572	10.3200	11.1083	11.0038	10.7959	11.0099	10.1796	10.3231	9.7819
Band8	10.8172	11.6894	10.2670	11.0272	10.9420	10.7253	11.9361	10.1363	10.2562	9.7181