

# Comparison of Four Demosaicing Methods for Facial Recognition Algorithms

M. Eléonore Elvire HOUSSOU<sup>1</sup>, A. Tidjani SANDA MAHAMA<sup>2</sup>, Pierre GOUTON<sup>3</sup>, Guy DEGLA<sup>4</sup>

ImVIA, University of Bourgogne, Dijon, France<sup>1, 2, 3</sup>  
IMSP, University of Abomey-Calavi, Dangbo, Benin<sup>1, 2, 4</sup>

**Abstract**—Multispectral imaging has become more important in several areas during this decade to overcome the limitations of color imaging. There are several types of multispectral acquisition systems, including single-shot cameras that incorporate Multispectral Filter Arrays (MSFA). MSFA is an extension of the color filter array. Acquisition systems that incorporate spectral filter arrays are very fast, lightweight, and able to acquire moving scenes. But these cameras are manufactured with at best software for filter positioning correction without demosaicing software. Hence there is a need to identify a suitable demosaicing algorithm in terms of image quality, computation time, and decorrelation factor. This paper presents a comparative study of four relevant demosaicing methods in the facial recognition process using images acquired with a single-shot MSFA camera designed in our laboratory. To achieve this goal, the four demosaicing methods named bilinear interpolation, discrete wavelet transform, binary tree, and median vector were adapted to multispectral images acquired using a MSFA camera. Evaluations were first performed using the NIQE performance metric and the correlation coefficient. Then Demosaiced images were used to train VGG19 neural network to know which demosaicing method better contains relevant features for recognition and better computation time. Results reveal that bilinear interpolation provides the less correlated images and the binary tree gives the best quality images with a NIQE of 8.99 and an accuracy of 100% for face recognition.

**Keywords**—Multispectral image database; multispectral imaging; multispectral filter array (MSFA); one-shot camera; facial recognition system

## I. INTRODUCTION

Most of today's color cameras incorporate color Filter Array (CFA) or Bayer filters [1], [2]. The color filter array contains three filters: green, red, and blue, each one responsible for acquiring the image in each spectral band. These cameras are very fast, displaying the acquired image in a matter of seconds. Despite this performance, facial recognition [3] with color cameras is affected by problems of light variations, occlusions, and pose variations [4], [5]. Multispectral imaging corrects these problems, with more information available in the image bands. There are three types of multispectral acquisition systems: multi-camera systems, multi-sensors systems, and one-shot MSFA systems. The first two types are very slow, heavy, and consume more energy. One-shot MSFA cameras help to overcome the

problems associated with the other two types. The one-shot MSFA cameras are equipped with MSFA which is an extension of CFA [6]–[8]. The MSFA includes more than three filters. Each filter is responsible for the acquisition of the image on a given wavelength. The acquisition systems that incorporate the multispectral filters array allow the acquisition of a single image on several spectra simultaneously. These are compact, one-shot cameras, very fast, and capable of capturing moving scenes. MSFA one-shot cameras solve the problems associated with conventional multispectral cameras, which are the heaviness and slowness during the acquisition of multispectral images. MSFA one-shot cameras are used in several fields such as agriculture, medical imaging, and pattern recognition [9]–[12].

A recognition system is composed of four main modules namely acquisition, feature extraction, matching, and decision. The performance of the system depends on each of the modules. There are several facial recognition systems in the multispectral but most of them use multispectral cameras consisting of multiple single-shot cameras or a single-shot camera in scanning mode. For the acquisition module, a database of face images was collected with a single-shot MSFA camera for facial recognition. This camera is mainly equipped with a viamagic CMOS sensor, a MSFA with eight filters, micro-lens, an electronic board to drive the sensor, and a camera board for image acquisition. The MSFA bands were selected theoretically with a genetic algorithm combined with a facial recognition application [13]. This acquisition system covers the spectral range from 650 to 950 nm and produces raw or mosaic images which require demosaicing before use. Demosaicing is a method of estimating the value of missing pixels in a given band. At the end of demosaicing the number of multispectral images obtained is equal to the number of filters that compose the MSFA. Some demosaicing methods [14]–[18] are developed during the theoretical design of the MSFA, but the industrial constraints of MSFA manufacturing mean that effective demosaicing methods are developed after the MSFA is manufactured. In this case, we already have a MSFA camera, and we want to determine suitable demosaicing methods using its images. Demosaicing methods impact the image quality and thus the performance of the facial recognition system with a single-shot MSFA camera. Single-shot MSFA cameras are very efficient and fast, but the quality of the demosaiced images depends on the demosaicing method used.

This paper presents the comparative study of four demosaicing methods to identify the one that gives:

- The best image quality result.
- The best decorrelation factor.
- The best facial recognition score.
- And the best computation time.

For these purposes, we adapted some currently used algorithms namely demosaicing algorithms Bilinear Interpolation, Discrete Wavelet Transform (DWT), Binary Tree, and Median Vector to demosaic the images acquired with our camera. A comparison of the demosaiced images is made with the NIQE metric and the intercorrelation between the demosaiced image bands is analyzed with the coefficient correlation. The convolutional neural network VGG19 is used to evaluate the impact of demosaicing on facial recognition in terms of accuracy and computation time. To the best of our knowledge, this is the first study that focuses on demosaicing after the MSFA camera is manufactured.

This paper is structured as follows: Section II briefly presents the MSFA one shot camera, material and methods in Section III, followed by experimental results and discussion in Section IV. The paper ends with the conclusion and future perspectives in Section V.

## II. THE MSFA ONE SHOT CAMERA

Multispectral imaging with a single camera equipped with a spectral or multispectral filter array allows acquiring multispectral images simultaneously on several spectral bands. The concept of a multispectral filter array is an extension to  $n$  bands of the color filter array that revolutionized digital cameras. A spectral filter array is composed of  $n$  filters and each of them is responsible for the acquisition of the image on a given band. The MSFA camera used in this work is composed of a single Viimagic 9220H sensor, an MSFA for the single-tap imaging system, optical lenses, an electronic board to drive the sensor, and a camera board for image acquisition. This acquisition system was designed in the Imaging and Artificial Vision (ImViA) laboratory formerly known as the Laboratory of Electronics, Informatics, and Image (LE2I) as part of the EU H2020 project called EXIST (Extended Image Sensing Technologies) [13]. It is a light and compact camera that covers wavelengths from 650 to 950 nm. The spectral filter array integrated into this camera is made up of eight optimal filters selected in the wavelengths {685, 720, 770, 810, 835, 870, 895, 930}. The design of this custom filter array uses the Fabry-Perot interferometer. The MSFA system integrated into this camera with dedicated hardware and software calculations allows working in real time with 30 frames per second. The filters used to overcome the problems caused by illumination variation, motion blur noise, and SNR noise that severely affect the performance of facial recognition systems using CMOS. The multispectral filter array is characterized by its moxel which is defined by a mosaic of elementary filters repeated across an MSFA. Fig. 1 illustrates the moxel used.

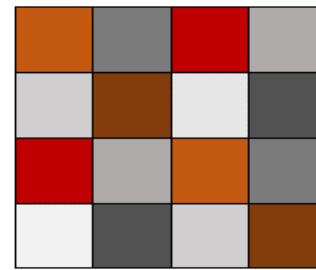


Fig. 1. Final moxel of the MSFA.

## III. METHODS

### A. Motivation

The literature distinguishes three categories of demosaicing methods: pixel interpolation, frequency transformation, and probability of appearance (POA). Pixel interpolation consists in using the value or weight of neighboring pixels to estimate the value of a missing pixel. Methods such as bilinear interpolation, weighting bilinear interpolation, and binary tree use pixel interpolation. Frequency transformation involves using wavelets to extract essential information from neighboring pixels to estimate the value of the missing pixel. Frequency transformation includes methods such as Discrete Wavelet Transform. Probability of appearance refers to methods that determine the probability of the occurrence of a band and a pixel in a selected band to estimate the value of the missing pixel. These methods include approaches based on binary trees. In general, demosaicing methods depend on the MSFA moxel but nowadays there are generic demosaicing methods that can be adapted to any type of MSFA. Discrete Wavelet Transform, Bilinear Interpolation, Binary tree, and vector median filtering were selected for the comparative study. These four demosaicing methods were selected because each one presents some interesting characteristics for the study. These were chosen for the following reasons:

- Bilinear interpolation uses the value or weights of neighboring pixels to estimate a missing pixel in a given band. Most demosaicing algorithms combine their demosaicing technique with bilinear interpolation [17], [19]–[21].
- Binary Tree-based Edge-Sensing method is a generic approach that uses the notion of POA to select the band and the pixel in the selected band. This method combines POA, and bilinear interpolation based on the edge-sensing information [22] to determine the value of the missing pixel.
- Discrete Wavelet Transform based MSFA demosaicing [17] is a technique that uses frequency information and Weighted Bilinear interpolation to approximate the value of the missing pixel.
- Vector median filtering [23] is a demosaicing method whose specificity is to use vector based operations and the concept of pseudo pixel. This method groups neighboring pixels according to the size of the moxel to evaluate the value of the pixel missing.

### B. Preprocess

The images acquired with the MSFA one-shot cameras are raw images or mosaic images and must be process or demosaic before using. Demosaicing is a technique that consists in reconstructing each band of the multispectral according to the number N (N=8 in this study) of filters contained in the MSFA. Before demosaicing, a preprocess of band extraction is first performed. This preprocess consists in multiplying each mosaic image by different binary masks  $M^k$  as defined by the Eq. (1).

$$M_{(x,y)}^k = \begin{cases} 1 & \text{si } (x \bmod \sqrt{K}) + (y \bmod \sqrt{K}) \times \sqrt{K} = k \\ 0 & \text{else} \end{cases} \quad (1)$$

This transformation gives eight planes of shifted images in which only one component is available at each pixel. Fig. 2 illustrates this transformation.

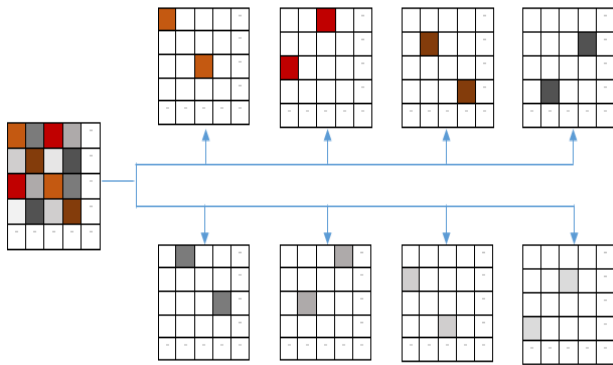


Fig. 2. Images bands obtained after applying the mask.

After pre-processing, the missing pixels in each image band are estimated using the four selected demosaicing methods.

### C. Bilinear Interpolation based MSFA Demosaicing

The bilinear interpolation based demosaicing is the simplest method for calculating the value of pixels missing during the demosaicing process. Other demosaicing methods combine their specificities with bilinear interpolation [24]. Bilinear interpolation approximates each missing pixel value by means of a distance weighted average of its neighboring pixels. As its name indicates, bilinear interpolation is a succession of two linear interpolations. The linear interpolation can be performed in multiple directions. For a missing pixel  $P(i, j)$  at position  $(i, j)$ , the linear interpolation is defined in Eq. (2), Eq. (3) and Eq. (4) as follows:

- 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

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

For this study, a convolution filter  $H$  is used for demosaicing. This filter is defined according to the spatial distance between the neighbors of the central pixel. The interpolated image band  $I^k$  is defined by Eq. (5) as:

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

with

$$H = \begin{pmatrix} 5 & 5 & 1 & 1 & 6 & 6 & 8 & 8 & 1 & 1 & 6 & 6 \\ 5 & 5 & 1 & 1 & 6 & 6 & 8 & 8 & 1 & 1 & 6 & 6 \\ 4 & 4 & 8 & 8 & 3 & 3 & 4 & 4 & 5 & 5 & 3 & 3 \\ 4 & 4 & 8 & 8 & 3 & 3 & 4 & 4 & 5 & 5 & 3 & 3 \\ 2 & 2 & 7 & 7 & 5 & 5 & 2 & 2 & 7 & 7 & 8 & 8 \\ 2 & 2 & 7 & 7 & 5 & 5 & 2 & 2 & 7 & 7 & 8 & 8 \\ 5 & 5 & 1 & 1 & 6 & 6 & 8 & 8 & 1 & 1 & 6 & 6 \\ 5 & 5 & 1 & 1 & 6 & 6 & 8 & 8 & 1 & 1 & 6 & 6 \\ 4 & 4 & 8 & 8 & 3 & 3 & 4 & 4 & 5 & 5 & 3 & 3 \\ 4 & 4 & 8 & 8 & 3 & 3 & 4 & 4 & 5 & 5 & 3 & 3 \\ 2 & 2 & 7 & 7 & 5 & 5 & 2 & 2 & 7 & 7 & 8 & 8 \\ 2 & 2 & 7 & 7 & 5 & 5 & 2 & 2 & 7 & 7 & 8 & 8 \end{pmatrix}$$

And  $I'^k$  the image band.

### D. Discrete Wavelet Transform based MSFA Demosaicing

A Discrete Wavelet Transform (DWT) is a transform that decomposes a given signal into a number of sets, where each set is a time series of coefficients describing the time evolution of the signal in the corresponding frequency band. When applied to an image, DWT transforms the image into different frequency bands. DWT is used to decompose images into a series of sub-bands with different frequency components. Over the decades, several DWT-based methods have been proposed in [25]–[29] for Color Filters Array (CFA) demosaicing. Xingbo et al. [16] extended the application of DWT into MSFA demosaicing. MSFA demosaicking based on DWT encompasses the concept of Down Sampling images, the Haar wavelet (D2), the “replace” rule for the estimation of high-frequency sub-bands and bilinear interpolation for the estimation of low-frequency sub-bands. This approach is applicable to any MSFA with a regular mosaic pattern, regardless of the number of channels. The algorithm is performed in three successive steps as follows:

- High-frequency estimation: First, the image is divided into K down sampled images and each of them is decomposed into spatial frequency sub-bands by DWT using Haar wavelet. Then, estimate the coefficients of the missing DS images in the high frequency sub-bands according to the “replace” rule.
- Low-frequency estimation: Apply bilinear interpolation to the mosaicked image plane by plane and extract the low frequency by image decomposition using Haar wavelet. Then replace the coefficients of the missing DS images at low-frequency sub-bands with those of the interpolated DS images.
- Recompose the low-frequency and high-frequency components and compute inverse discrete wavelet transform to reconstruct the demosaiced image.

This method has been tested by the authors with CFA and two MSFA of four-bands and eight-bands visible. Given the compatibility of this method with all MSFA with regular patterns, regardless of the number of channels, we have modified it to our eight-bands MSFA described above with two pixels per band. For this algorithm, the wavelengths of the bands have been modified by considering the following wavelengths {685, 720, 770, 810, 835, 870, 895, 930 nm} for the eight bands.

#### E. Binary Tree based MSFA Demosaicing

Lidan Miao et al. [20]–[22] had proposed a generic demosaicing method based on a binary tree. Specifically, this approach uses the Binary Tree Edge Sensing Method (BTES). The missing pixels are estimated progressively using all the edge correlation information of all spectral bands. Binary tree based MSFA demosaicing can be adapted to any  $4 \times 4$  MSFA. The generic demosaicing algorithm consists of three interconnected modules: band selection, pixel selection and interpolation [20].

- Band selection: This module defines the interpolation order of the different spectral bands by using the POA. The spectral bands have different POA, and the spectral band with the highest POA contains the most detailed information. Band selection is equivalent to selecting leaf nodes (spectral bands) at different levels of the tree. Nodes located at the same level have the same POA, and the deepest nodes have the smallest POA. The band selection process is described as follows: first, the band with the highest POA is selected, the leaf node at the first level of the binary tree. Then, for levels with more than one leaf, the bands at the next level are randomly selected. Finally, this process is repeated until the last level of the tree.

The band with the most edge information is interpolated first and the edge information from the first interpolated band is used to estimate the other bands.

- Pixel selection: This module determines the order of interpolation of pixel locations in each spectral band. The estimation of the pixel values is done gradually. First, some of the missing pixel values are estimated. The other unknown pixel values are then estimated using these estimated pixel values and the MSFA samples. The algorithm uses the binary tree for the pixel's selection. It takes as input the leaf patterns selected during band selection and interpolates for each of them first the missing band information at the pixel locations where its sibling pattern is located, and then the algorithm goes up one level in the binary tree to find the sibling of its parent pattern. If the latter is an internal node, then the leaf patterns in the subtree below that sibling pattern are examined. This part is repeated until the root of the tree.
- Interpolation is used to estimate the value of the missing pixels with selected pixel for a selected band. The estimate of the value of a pixel  $P$  at position  $(i, j)$  is calculated by the weighted sum of these four neighboring pixels and their contributions. The weights

of these four pixels are estimated based on their edge magnitudes. The weights of two neighboring pixels along the vertical and horizontally direction are calculated by Eq. (6) and Eq. (7).

Vertically:

$$W_{m,n} = \left( 1 + |P_{m+2,n} - P_{m,n}| + |P_{m-2,n} - P_{m,n}| + \frac{1}{2} |P_{m-1,n-1} - P_{m+1,n-1}| + \frac{1}{2} |P_{m-1,n+1} - P_{m+1,n+1}| \right)^{-1} \quad (6)$$

With  $m \in \{i - 1, i + 1\}, n = j$

Horizontally:

$$W_{m,n} = \left( 1 + |P_{m,n+2} - P_{m,n}| + |P_{m,n-2} - P_{m,n}| + \frac{1}{2} |P_{m+1,n-1} - P_{m+1,n+1}| + \frac{1}{2} |P_{m-1,n-1} - P_{m-1,n+1}| \right)^{-1} \quad (7)$$

With  $m = i, n \in \{j - 1, j + 1\}$

The estimated value  $\hat{P}_{i,j}$  of pixel at position  $(i, j)$  is defined in Eq. (8).

$$\hat{P}_{i,j} = \frac{\sum_{s,t(|s+t|=1)} W_{i+s,j+t} P_{i+s,j+t}}{\sum_{s,t(|s+t|=1)} W_{i+s,j+t}} \quad (8)$$

The algorithm has been developed for 7 bands. Since the MSFA used has 8 bands, this algorithm was modified for 8 bands, considering the probability of bands appearing.

#### F. Vector Median based MSFA Demosaicing

Gupta et al. [30] had proposed a CFA vector demosaicing algorithm. This approach selects the color vector that minimizes the sum of the distances to the neighboring pixels to estimate the missing colors. This demosaicing approach is based on the notion of pseudo-pixel, which is defined by a group of neighboring values of red, green, and blue pixels (horizontally and vertically). Xingbo et al. [22] had extended this technique for MSFA demosaicing. According to the authors, this method is based on two specificities: first, the pseudo-pixels are formed according to the dimension of a moxel. The Moxel is a mosaic element corresponding to a mosaic of elementary filters repeated via an MSFA. Second, the pseudo-pixels are those that are connected horizontally, vertically, and diagonally.

The median vector  $\underline{x}_{VM}$  of  $\underline{x}_1, \dots, \underline{x}_N$  is defined as follows:

$$\underline{x}_{VM} \in \{\underline{x}_i | i = 1, \dots, N\}$$

For all  $j = 1, \dots, N$

$$\sum_{i=1}^N \|\underline{x}_{VM} - x_i\|_2 \leq \sum_{i=1}^N \|\underline{x}_j - x_i\|_2 \quad (9)$$

The algorithm of calculation of the median vector for the demosaicing of the multispectral images proceeds as follows:

- For each vector  $x_i$ , compute the sum of the distances to all other vectors using the  $L^1$ -norm or  $L^2$ -norm as presented in Eq. (10)

$$S_i = \sum_{j=1}^N \|\underline{x}_j - x_i\|, i = 1, \dots, N \quad (10)$$

- The median vector is  $x_{min}$  corresponding to  $S_{min}$ , which is the minimum of  $S_i$

This method was tested by the authors with CFA, 4-bands and 8-bands MSFA. For this algorithm, the wavelengths of the bands have been modified by considering the following wavelengths {685, 720, 770, 810, 835, 870, 895, and 930} nm for the eight bands.

#### IV. EXPERIMENTAL RESULTS AND DISCUSSION

##### A. EXIST Database and Experimental Setup

EXIST is a multispectral image database that was collected with the MSFA one-shot camera described. The EXIST dataset is composed of 2100 raw images of faces of 105 subjects. Each multispectral image is 2072 x 1104 size. After demosaicing, the images obtained are of size 2072 x 1104x8 each corresponding respectively to the wavelengths {685, 720, 770, 810, 835, 870, 895, 930} in nm. Fig. 3 shows some images of the EXIST database.



Fig. 3. Samples of raw images of EXIST database.

The experiments were carried out on Microsoft System Windows, version 2010, with two computers. The first was equipped with an Intel(R) Core (TM) i7-8565U CPU, 8 GB of RAM memory. The second has a NVIDIA Quadro P400 graphics processing unit (GPU) with 32 GB of Random Access Memory (RAM). All code is written in the MATLAB 2020 and Python 3.7 programming languages.

##### B. Evaluation Metrics

To identify a suitable demosaicing algorithm, evaluations were performed on three criteria: image quality, decorrelation factor, and recognition rate, using NIQE, correlation coefficient, and recognition accuracy as metrics, respectively. Based on the literature [31]–[33], there are two types of performance metric for image quality: full reference quality, such as Peak Signal Noise Ratio (PSNR), Structural Similarity Index (SSI), Mean Square Error (MSE) and no-reference quality measurements, such as NIQE, Brisqe and piqe. Since the EXIST image database only contains raw images without references, the NIQE performance measure is used to assess the quality of images demosaiced using different methods.

NIQE is developed by researchers of University of Texas [34]. C.Kawan et al. [35] used NIQE to compare demosaicing method for Mastcam Images. NIQE is no-reference image quality metrics that use statistical features of the input image to evaluate the image quality. To calculate NIQE value, the image is divided into smaller patches, and the features are modeled as MultiVariate Gaussian (MVG) distributions.

After image quality evaluation with NIQE, intercorrelation of demosaiced images is evaluate with correlation coefficient computation. This coefficient allows to identify the degree of similarity between a pair of images. Images are identical when the correlation coefficient is equal to 1. To calculate this correlation coefficient, the method described in [36] was used. Also, to identify the best demosaicing method with the best image quality, decorrelation factor, the demosaiced images were used to train four models based on the VGG19 architecture. VGG19 is an architecture of VGGNet proposed by K. Simonyan et al. [37] in 2015. VGG19 contains 19 weight layers consisting of 16 convolutional layers with three fully connected layers and the same five pooling layers. VGG-19 CNN is used as a pre-training model. The accuracy is used to evaluate the images after recognition. Accuracy indicates the percentage of correct prediction.

##### C. Result and Discussion

To carry out the test, the following steps were followed for each method:

- Demosaicing raw images.
- Computation of NIQE values and time of demosaiced images.
- Calculation of correlation coefficients between demosaic images.
- Calculation of the average demosaicing.
- Recognition of demosaiced images.
- Computation of recognition time.

This algorithm led to the results presented in Fig. 4.



Fig. 4. Demosaic images by methods.

Visually, we can see that the binary tree presents better images in terms of sharpness of the different parts of the face than the other three demosaicing methods. But visual evaluation alone is not enough. The NIQE metric was therefore used to confirm or invalidate this visual assessment. The quality of the image varies according to the value of NIQE. The smaller the NIQE value, the better the quality of the demosaic image.

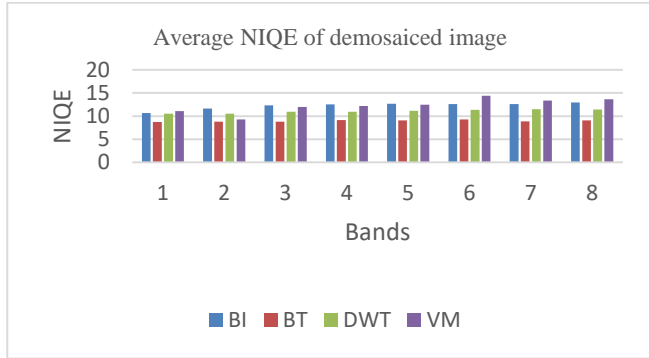


Fig. 5. Average NIQE of demosaiced image.

The Fig. 5 shows the average NIQE calculated on all images obtained for the four demosaicing methods.

For all demosaicing methods, the NIQE value of the corresponding images varies between 8.75 and 14.38 according to the band and demosaicing method. While the NIQE values of the images demosaiced by bilinear interpolation are between 10.68 and 12.97, those of BT are between 8.75 and 9.29, those of DWT between 10.51 and 11.42, and those of VM between 11.68 and 13.66.

The analysis of the NIQE value per band and per demosaicing method shows that the binary tree has the lowest NIQE value per band and for all methods with average NIQE value of 8.99. Based on this NIQE value, the binary tree provides better quality images than others. This analysis confirms the visual observation made after the experimental results.

Time simulations were then made to compare the four methods. The Figure 6 shows the execution time for each.

Interpretation of Fig. 6 shows that the minimum running time of a demosaicing algorithm is 12.78s and is obtained with bilinear interpolation. This time is slightly less than that of the binary tree. The median vector is the method with the longest execution time. The crossing of the execution time and the NIQE value of the demosaicing algorithms shows that the binary tree is the demosaicing method which allows having best quality images in a reasonable time.

To study the decorrelation factor of demosaiced images, we compute the correlation coefficients between them.

The Fig. 7 shows the average correlation coefficient between each band by demosaicing methods.

The demosaiced images are correlated with an average correlation coefficient of 0.9 for all methods. The analysis of

these different figures allows us to conclude that the intercorrelation factor between the different bands is lower for the Bilinear interpolation than for the other demosaicing methods. Bilinear interpolation allows us to have less correlated images than the other methods.

After the demosaicing process and comparison, the VGG19 convolutional neural network was used for feature extraction and classification of the demosaiced images obtained with each method. The dataset was identical for all the methods. The training and test datasets were separated with random selection with 80% for training, 10% for validation and 20% for test. In all, 2,200 demosaic images of size 300 x 300 pixels, organized into 110 classes and were used.

Tables I to III describe respectively the training parameter of VGG19, the accuracy, and the recognition time for each image by method.

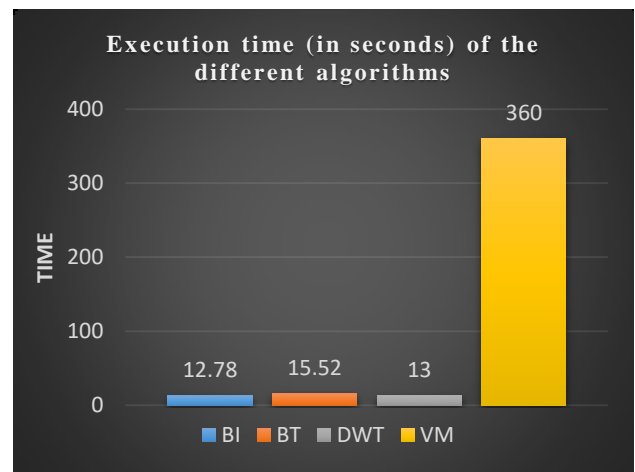


Fig. 6. Execution time of demosaicing algorithms.

TABLE I. PARAMETERS USED IN THE TRAINING PROCEDURE

Parameters	VGG19
Batch size	32
Optimization algorithm	SGD
Learning rate	0.0001
Epoch number	20

TABLE II. VGG19 RESULT

Accuracy	BI	DWT	BT	VM
VGG19	100	89	100	80

TABLE III. EXECUTION TIME

Accuracy	BI	DWT	BT	VM
Time(s)	3	4	2	6



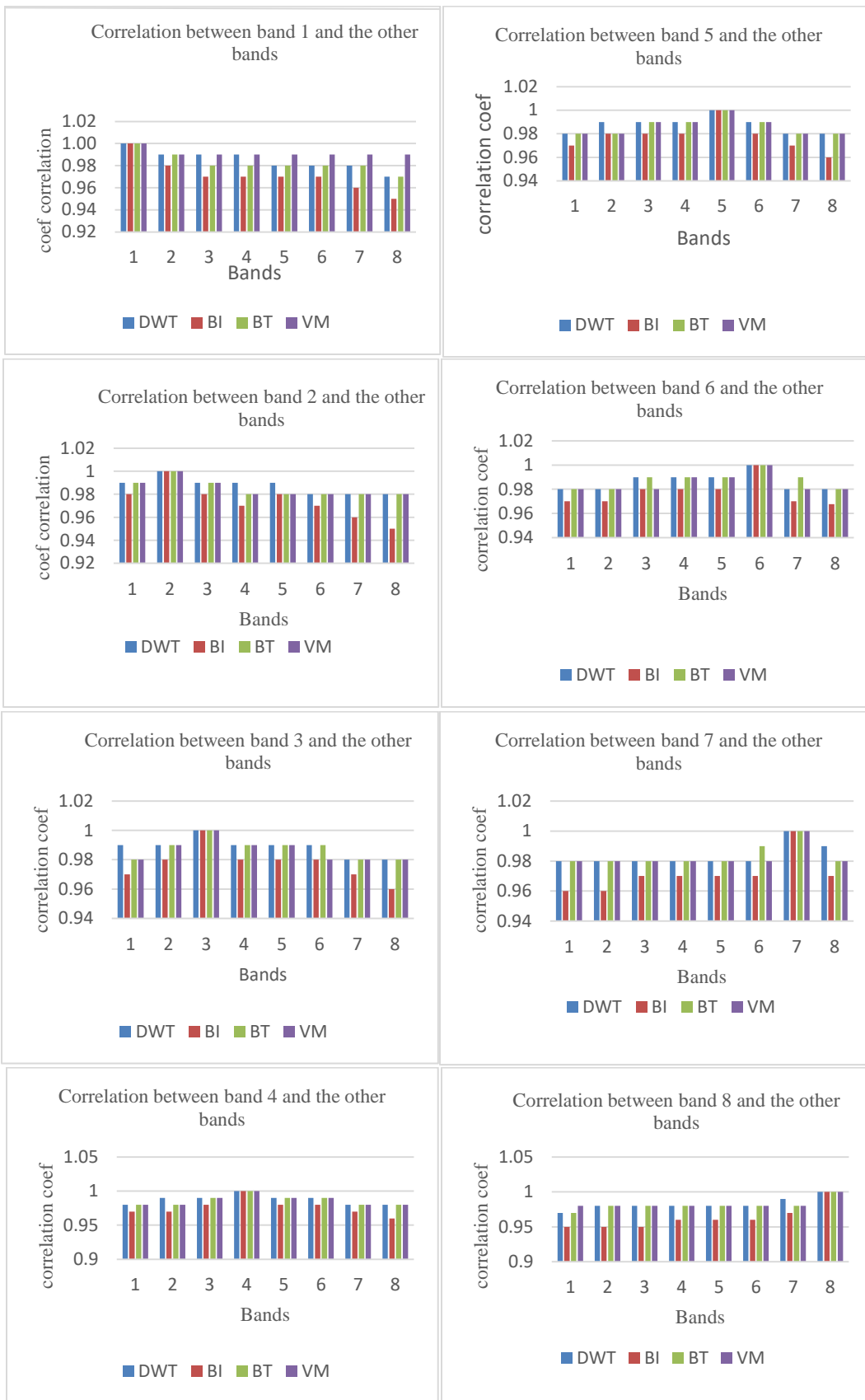


Fig. 7. Average correlation coefficient between bands by demosaicing methods.

Analysis of the results in Table II shows that the accuracies obtained vary between 80% and 100%. Images demosaiced with bilinear interpolation and binary tree have the best accuracy of 100%. A comparison of the results in Table III shows that the execution time for recognizing demosaiced images varies between two and six sec. Images demosaiced with the binary tree are recognized faster than those of the other methods.

In conclusion, the different results show that the binary tree-based demosaicing method is the best of these four methods in terms of image quality, computational time, and accuracy for facial recognition.

## V. CONCLUSION

We proposed an evaluation of demosaicing methods for facial recognition using MSFA one-shot cameras to identify a suitable demosaicing algorithm in terms of image quality, computation time and decorrelation factor. This study use and compare four demosaicing algorithms on a base of raw images acquired with a MSFA one-shot camera. The binary tree demosaicing method was used to obtain the best quality images with the best computation time and, accuracy for facial recognition. Demosaicing affects the facial recognition system in terms of image quality and time. The better the demosaic images, the better the accuracy of the facial recognition system. The minimum demosaicing time is 12.78 s with bilinear interpolation. This is huge compared to the demosaicing time of CFA which is a few milliseconds.

An optimization of this system should be done in our next work to make it real time.

## REFERENCES

- [1] R. Lukac and K. N. Plataniotis, "Color filter arrays: design and performance analysis," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 4, pp. 1260–1267, Nov. 2005, doi: 10.1109/TCE.2005.1561853.
- [2] Y.-P. Fan, L. Wei, L. Li, L. Yang, Z.-Q. HU, Y.-H. Zheng, and Y.-H. Wang, "Research on the Modulation Transfer Function Detection Method of a Bayer Filter Color Camera," *Sensors*, vol. 23, no. 9, Art. no. 9, Jan. 2023, doi: 10.3390/s23094446.
- [3] M. Lal, K. Kumar, R. H. Arain, A. Maitlo, S. A. Ruk, and H. Shaikh, "Study of Face Recognition Techniques: A Survey," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 9, no. 6, Art. no. 6, 29 2018, doi: 10.14569/IJACSA.2018.090606.
- [4] P. Martins, J. S. Silva, and A. Bernardino, "Multispectral Facial Recognition in the Wild," *Sensors*, vol. 22, no. 11, Art. no. 11, Jan. 2022, doi: 10.3390/s22114219.
- [5] S. Arya, N. Pratap, and K. Bhatia, "Future of Face Recognition: A Review," *Procedia Computer Science*, vol. 58, pp. 578–585, Jan. 2015, doi: 10.1016/j.procs.2015.08.076.
- [6] Y. Monno, S. Kikuchi, M. Tanaka, and M. Okutomi, "A Practical One-Shot Multispectral Imaging System Using a Single Image Sensor," *IEEE Transactions on Image Processing*, vol. 24, no. 10, pp. 3048–3059, Oct. 2015, doi: 10.1109/TIP.2015.2436342.
- [7] P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton, "Multispectral Filter Arrays: Recent Advances and Practical Implementation," *Sensors*, vol. 14, no. 11, pp. 21626–21659, Nov. 2014, doi: 10.3390/s141121626.
- [8] S. Mihoubi, O. Losson, B. Mathon, and L. Macaire, "Multispectral demosaicing using intensity-based spectral correlation," in *2015 International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Nov. 2015, pp. 461–466. doi: 10.1109/IPTA.2015.7367188.
- [9] S. Gutiérrez, A. Wendel, and J. Underwood, "Spectral filter design based on in-field hyperspectral imaging and machine learning for mango ripeness estimation," *Computers and Electronics in Agriculture*, vol. 164, p. 104890, Sep. 2019, doi: 10.1016/j.compag.2019.104890.
- [10] M. Ewerlöf, M. Larsson, and E. G. Salerud, "Spatial and temporal skin blood volume and saturation estimation using a multispectral snapshot imaging camera," in *Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues XV*, SPIE, Feb. 2017, pp. 105–116. doi: 10.1117/12.2251928.
- [11] K. Shinoda, S. Ogawa, Y. Yanagi, M. Hasegawa, S. Kato, M. Ishikawa, Y. Komagata, and N. Kobayashi in *2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, Dec. 2015, pp. 697–703. doi: 10.1109/APSIPA.2015.7415362.
- [12] M. E. E. Houssou, A. T. S. Mahama, P. Gouton, and G. Degla, "Robust Facial Recognition System using One Shot Multispectral Filter Array Acquisition System," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 13, no. 1, Art. no. 1, 55/31 2022, doi: 10.14569/IJACSA.2022.0130104.
- [13] Porteur UBFC: Gouton, "EXIST (EXTended Image Sensing Technologies), H2020 (ECSEL)," au 12/2018 2015.
- [14] S. Mihoubi, O. Losson, B. Mathon, and L. Macaire, "Multispectral demosaicing using intensity-based spectral correlation," in *2015 International Conference on Image Processing Theory, Tools and Applications (IPTA)*, Nov. 2015, pp. 461–466. doi: 10.1109/IPTA.2015.7367188.
- [15] M. Gupta, V. Rathi, and P. Goyal, "Adaptive and Progressive Multispectral Image Demosaicking," *IEEE Transactions on Computational Imaging*, vol. 8, pp. 69–80, 2022, doi: 10.1109/TCI.2022.3140554.
- [16] M. Gupta and P. Goyal, "Demosaicing Method for Multispectral Images Using Derivative Operations," *American Journal of Mathematical and Management Sciences*, vol. 40, no. 2, pp. 163–176, Apr. 2021, doi: 10.1080/01966324.2021.1939206.
- [17] X. Wang, J.-B. Thomas, J. Y. Hardeberg, and P. Gouton, "Discrete wavelet transform based multispectral filter array demosaicking," in *2013 Colour and Visual Computing Symposium (CVCS)*, Sep. 2013, pp. 1–6. doi: 10.1109/CVCS.2013.6626274.
- [18] S. Mihoubi, O. Losson, B. Mathon, and L. Macaire, "Multispectral Demosaicing using Intensity-based Spectral Correlation," presented at the *5th IEEE International Conference on Image Processing Theory, Tools and Applications (IPTA 2015)*, Nov. 2015, p. 461. doi: 10.1109/IPTA.2015.7367188.
- [19] M. Gupta, "Generalizing Spectral Difference Method for Multispectral Image Demosaicking and Analyzing the Role of MSFA Patterns," in *2020 9th International Conference on System Modeling and Advancement in Research Trends (SMART)*, Dec. 2020, pp. 416–420. doi: 10.1109/SMART50582.2020.9337090.
- [20] L. Miao, H. Qi, and R. Ramanath, "A Generic Binary Tree-Based Progressive Demosaicking Method for Multispectral Filter Array," in *2006 International Conference on Image Processing*, Oct. 2006, pp. 3221–3224. doi: 10.1109/ICIP.2006.312909.
- [21] L. Miao, H. Qi, R. Ramanath, and W. E. Snyder, "Binary Tree-based Generic Demosaicking Algorithm for Multispectral Filter Arrays," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3550–3558, Nov. 2006, doi: 10.1109/TIP.2006.877476.
- [22] L. Miao, H. Qi, and R. Ramanath, "Generic MSFA demosaicking and demosaicking for multispectral cameras," in *Digital Photography II*, SPIE, Feb. 2006, pp. 88–97. doi: 10.1117/12.642366.
- [23] X. Wang, J.-B. Thomas, J. Y. Hardeberg, and P. Gouton, "Median filtering in multispectral filter array demosaicking," in *Digital Photography IX*, SPIE, Feb. 2013, pp. 103–112. doi: 10.1117/12.2005256.
- [24] N. Hounsou, A. T. Sanda Mahama, and P. Gouton, "Extension of luminance component based demosaicking algorithm to 4- and 5-band multispectral images," *Array*, vol. 12, p. 100088, Dec. 2021, doi: 10.1016/j.array.2021.100088.
- [25] J. Driesen and P. Scheunders, "Wavelet-based color filter array demosaicking," in *2004 International Conference on Image Processing*,



2004. ICIP '04., Oct. 2004, pp. 3311-3314 Vol. 5. doi: 10.1109/ICIP.2004.1421822.
- [26] G. Spampinato, A. Bruna, G. Sanguedolce, E. Ardizzone, and M. L. Cascia, "Improved color interpolation using discrete wavelet transform," in *Image and Video Communications and Processing 2005*, SPIE, Mar. 2005, pp. 753-760. doi: 10.1117/12.585200.
- [27] J. Aelterman, B. Goossens, A. Pižurica, and W. Philips, "Locally adaptive complex wavelet-based demosaicing for color filter array images," in *Wavelet Applications in Industrial Processing VI*, SPIE, Jan. 2009, pp. 115-126. doi: 10.1117/12.807347.
- [28] M. C. Kim, J. H. Park, and M. H. Sunwoo, "Multilevel Feature Extraction Using Wavelet Attention for Deep Joint Demosaicking and Denoising," *IEEE Access*, vol. 10, pp. 77099-77109, 2022, doi: 10.1109/ACCESS.2022.3192451.
- [29] B. K. Gunturk, Y. Altunbasak, and R. Mersereau, "Color plane interpolation using alternating projections," in *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, May 2002, p. IV-3333-IV-3336. doi: 10.1109/ICASSP.2002.5745367.
- [30] M. R. Gupta and T. Chen, "Vector color filter array demosaicing," presented at the *Photonics West 2001 - Electronic Imaging*, M. M. Blouke, J. Canosa, and N. Sampat, Eds., San Jose, CA, May 2001, p. 374. doi: 10.1117/12.426974.
- [31] A. Horé and D. Ziou, "Image Quality Metrics: PSNR vs. SSIM," in *2010 20th International Conference on Pattern Recognition*, Aug. 2010, pp. 2366-2369. doi: 10.1109/ICPR.2010.579.
- [32] Z. Wang and A. C. Bovik, "Reduced- and No-Reference Image Quality Assessment," *IEEE Signal Processing Magazine*, vol. 28, no. 6, pp. 29-40, Nov. 2011, doi: 10.1109/MSP.2011.942471.
- [33] M. Siniukov, D. Kulikov, and D. Vatolin, "Applicability limitations of differentiable full-reference image-quality metrics," in *2023 Data Compression Conference (DCC)*, Mar. 2023, pp. 1-1. doi: 10.1109/DCC55655.2023.00082.
- [34] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'Completely Blind' Image Quality Analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209-212, Mar. 2013, doi: 10.1109/LSP.2012.2227726.
- [35] C. Kwan, B. Chou, and J. F. Bell III, "Comparison of Deep Learning and Conventional Demosaicing Algorithms for Mastcam Images," *Electronics*, vol. 8, no. 3, Art. no. 3, Mar. 2019, doi: 10.3390/electronics8030308.
- [36] A. G. Asuero, A. Sayago, and A. G. González, "The Correlation Coefficient: An Overview," *Critical Reviews in Analytical Chemistry*, vol. 36, no. 1, pp. 41-59, Jan. 2006, doi: 10.1080/10408340500526766.
- [37] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv:1409.1556 [cs], Apr. 2015, Accessed: Oct. 13, 2021. [Online]. Available: <http://arxiv.org/abs/1409.1556>