



**HAL**  
open science

## **Spatio-spectral binary patterns based on multispectral filter arrays for texture classification**

Sofiane Mihoubi, Olivier Losson, Benjamin Mathon, Ludovic Macaire

► **To cite this version:**

Sofiane Mihoubi, Olivier Losson, Benjamin Mathon, Ludovic Macaire. Spatio-spectral binary patterns based on multispectral filter arrays for texture classification. *Journal of the Optical Society of America. A Optics, Image Science, and Vision*, 2018, 35 (9), pp.1532-1542. <10.1364/JOSAA.35.001532>. <hal-01839207>

**HAL Id: hal-01839207**

**<https://hal.science/hal-01839207v1>**

Submitted on 8 Dec 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Spatio-spectral binary patterns based on multispectral filter arrays for texture classification

SOFIANE MIHOUBI<sup>1</sup>, OLIVIER LOSSON<sup>1,\*</sup>, BENJAMIN MATHON<sup>1</sup>, AND LUDOVIC MACAIRE<sup>1</sup>

<sup>1</sup>Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRISTAL - Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France.

\*Corresponding author: olivier.losson@univ-lille.fr

Compiled July 11, 2018

To discriminate gray-level texture images, spatial texture descriptors can be extracted using the local binary pattern operator. This operator has been extended to color images at the expense of increased memory and computation requirements. Some authors propose to compute texture descriptors directly from raw images provided through a Bayer color filter array, which both avoids the demosaicing step and reduces the descriptor size. Recently, multispectral snapshot cameras have emerged to sample more than three wavelength bands using a multispectral filter array. Such cameras provide a raw image in which a single spectral channel value is available at each pixel. In this paper we design a local binary pattern operator that jointly extracts the spatial and spectral texture information directly from a raw image. Extensive experiments on a large dataset show that the proposed descriptor has both reduced computation cost and high discriminative power with regard to classical LBP descriptors applied to demosaiced images. © 2018

Optical Society of America

**OCIS codes:** (110.4234) Multispectral and hyperspectral imaging; (100.0100) Image processing.

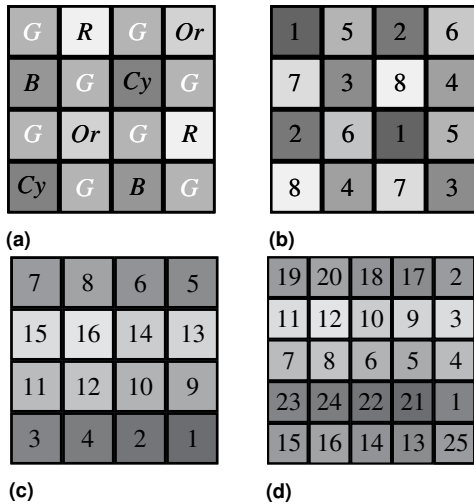
<http://dx.doi.org/10.1364/ao.XX.XXXXXX>

## 1. INTRODUCTION

Among multispectral imaging technologies, “linescan” devices acquire frames of narrow spatial slices for several spectral bands then, once the whole scene has been scanned, form the multispectral image by stacking the acquired frames. “Snapshot” devices can oppositely acquire the multispectral image from a single shot. Multi-sensor snapshot systems straight form a fully-defined multispectral image thanks to dichroic beam splitters that selectively redirect the incoming light by wavelength onto the sensors. Because these systems are expensive and are sensitive to a limited number of bands, single-sensor snapshot systems have been recently developed [1, 2]. Most of them use a multispectral filter array (MSFA) laid over the sensor that spectrally samples the incident light, like the widely-used Bayer color filter array (CFA) in color imaging. The MSFA is defined by a basic periodic pattern in which each filter is sensitive to a narrow spectral band. Each pixel of the resulting raw image is then characterized by one single band according to the MSFA. Such technology achieves a compromise between spatial and spectral samplings (see Sec. A). The fully-defined multispectral image is estimated by a demosaicing process (see Sec. B). This process may be greedy and is prone to generate spatio-spectral estimation artifacts. We therefore propose to avoid it for texture classification and design a texture descriptor that is directly computed from the raw image (see Sec. C).

## A. Multispectral filter arrays

Like the Bayer CFA in which the green (G) band is over-represented with regard to red (R) and blue (B) bands, some MSFA patterns in the literature include one or several dominant bands [3]. For instance, Monno *et al.* [4] propose the  $4 \times 4$  basic pattern shown in Fig. 1a that samples 5 bands in the visible (Vis) domain. The dominant G band characterizes half of the pixels and the other four bands evenly characterize the other half. This is a special case of the generic pattern generated by a binary tree [5], where the prior probability of the G band is  $\frac{1}{2}$  and that of the other bands is  $\frac{1}{8}$ . Thomas *et al.* [6] propose the basic redundant pattern shown in Fig. 1b that samples 8 bands in the visible and near infrared (NIR) domains. The two MSFAs based on the above basic patterns, called Vis5 and VisNIR8 in the following, own the desirable properties of spectral consistency and spatial uniformity [5]. An MSFA is spectrally consistent if the same bands are sampled the same number of times in the neighborhood of all filters associated to any given band. Spatial uniformity requires that an MSFA spatially samples each band as evenly as possible. Both requirements are related to the demosaicing process that is applied to the raw image. Indeed, demosaicing independently scans all the pixels associated to a given band and considers pixels in their neighborhood. This neighborhood layout should then be same whatever the pixel considered in the raw image.



**Fig. 1.** Basic patterns of different MSFAs: (a) Vis5 [4], (b) Vis-NIR8 [6], (c) IMEC16 and (d) IMEC25 [7]. Numbers are band indices. Vis5 band labels ( $Or=Orange$ ,  $Cy=Cyan$ ) are those of [4] but could also be replaced by indices. **Dominant G band is marked as white.**

Increasing the number of bands to enhance spectral resolution is a goal of multispectral imaging. Some MSFAs are then defined by a basic pattern with no dominant band although this conflicts with a dense spatial sampling. The two MSFAs whose square basic patterns are shown in Figs. 1c and 1d are manufactured by IMEC [7]. They are embedded in a few off-the-shelf MSFA-based devices available on the market, like XIMEA xiSpec and IMEC “snapshot mosaic” multispectral cameras, with applications in medical imaging [8] or terrain classification [9]. The  $4 \times 4$  basic pattern samples 16 bands centered at wavelengths  $\lambda^1 = 469 \text{ nm}$ ,  $\dots$ ,  $\lambda^{16} = 633 \text{ nm}$  (visible domain) and the  $5 \times 5$  one samples 25 bands centered at  $\lambda^1 = 678 \text{ nm}$ ,  $\dots$ ,  $\lambda^{25} = 960 \text{ nm}$  (visible and NIR domains). **Note that the band centers are not equally spaced and that they are not ascending in the classical pixel readout order** (see Figs. 1c and 1d). The MSFAs defined by these patterns (or the corresponding cameras) are shortly called IMEC16 and IMEC25 in the following.

## B. MSFA demosaicing

In the raw image acquired through an MSFA that samples  $K$  bands, a single band value is available at each pixel. A procedure called demosaicing must be applied to this raw image in order to obtain a multispectral image with  $K$  fully-defined channels, among which  $K - 1$  ones are estimated at each pixel.

The simplest way to demosaic the raw image is to use a weighted bilinear (WB) interpolation filter that estimates each missing band value at a given pixel by using the available values of the same band at its neighbors. Each neighbor is associated with a weight that depends on its spatial distance to the considered pixel [10]. This method only exploits spatial correlation within each channel to estimate the missing values.

To improve the demosaicing performances, some authors also use spectral (inter-channel) correlation. Brauers and Aach [10] compute the channel differences and Mizutani *et al.* [11] iterate this scheme a number of times given by the spectral distance between band centers of the considered pair of bands. Wang *et al.* [12] extend the CFA demosaicing method based on the discrete wavelet transform to MSFA demosaicing. Mihoubi *et*

*al.* [13] estimate the pseudo-panchromatic image (PPI) from the raw image and propose a demosaicing scheme based on PPI differences (PPID).

When the MSFA includes a dominant band (as does Vis5 in Fig. 1a), a standard approach is to pre-estimate the dominant channel, then to rely on it to estimate the other missing channels. The binary tree-based edge-sensing method of Miao *et al.* [5] estimates missing values from the edge information of the pre-estimated dominant channel. However, this method is only applicable when the prior probability of each band is  $\frac{1}{2^n}$ ,  $n \in \mathbb{N}$ , which notably excludes IMEC25 MSFA. Monno *et al.* propose three demosaicing methods that exploit the dominant G band of their Vis5 MSFA. The original algorithm [4] that uses an adaptive kernel up-sampling is improved by guided filtering [14], and further on by residual interpolation [15].

## C. Our contribution to texture classification

To classify multispectral texture images acquired by single-sensor snapshot cameras, the classical approach is to demosaic the raw images, extract texture features from the estimated images, then compare features that are computed from images thanks to a similarity measure.

In this paper, we focus on texture features based on local binary patterns (LBPs). The many variants of LBP operators have indeed proved to be very efficient for a wide variety of applications [16]. LBP-based texture classification has first been performed on gray-level images since the original operator only uses the spatial information of texture [17]. Later, Palm [18] has shown that classification based on a color analysis outperforms that based on the spatial information only. Texture feature extraction is then extended to the color domain by taking both spatial and spectral textural information into account.

Recent advances in multispectral imaging lead us to extend the color texture features to the multispectral domain (see Sec. 2). But the computational cost significantly increases with the number of channels due to demosaicing and feature extraction. Thus we propose a new computationally-efficient LBP-based descriptor that is directly computed from raw images, which allows us to avoid the demosaicing step (see Sec. 3). Extensive experiments on a large dataset of multispectral texture images prove the relevance of our approach (see Sec. 4).

## 2. SPECTRAL TEXTURE FEATURES BASED ON LOCAL BINARY PATTERNS

Below we formulate several LBP-based texture features for any fully-defined  $K$ -channel image, that we generically denote as  $\mathbf{I} = \{I^k\}_{k=1}^K$  for simplicity whatever the value of  $K$  and even it has been estimated by demosaicing. For a given pixel  $p$ , we consider the neighborhood  $N_p$  defined by its support  $\mathcal{N}^{p,d}$  made of  $P$  pixels at spatial uniform distance  $d$  from  $p$ . We now detail how the LBP operator compares the pixel values in the neighborhood  $N_p$  with respect to the value of  $p$ .

### A. Marginal LBPs

The marginal LBP operator (see Fig. 2) is defined for each band (index)  $k \in \{1, \dots, K\}$  at each pixel  $p$  as

$$LBP^k[\mathbf{I}](p) = \sum_{q \in N_p} s(I_q^k, I_p^k) \cdot 2^{\epsilon(q)}, \quad (1)$$

where  $I_p^k$  is the value of channel  $I^k$  at  $p$ ,  $\epsilon(q) \in \{0, \dots, P - 1\}$  is the index of each neighboring pixel  $q$  in  $N_p$ , and  $s(\cdot)$  is the unit

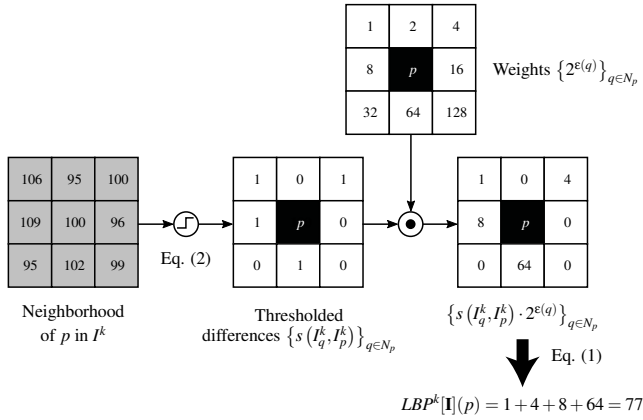


Fig. 2. Marginal LBP operator applied to a pixel  $p$  of channel  $I^k$ .

step function:

$$s(\alpha, \beta) = \begin{cases} 1 & \text{if } \alpha \geq \beta, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Each channel  $I^k$  is characterized by the  $2^P$ -bin un-normalized histogram of its LBP values. The multispectral texture image  $\mathbf{I}$  is then described by the concatenation of the  $K$  histograms of  $\{LBP^k[\mathbf{I}]\}_{k=1}^K$ . This feature, whose size is  $K \cdot 2^P$ , represents the spatial interaction between neighboring pixels within each channel independently.

Note that we only consider the basic LBP operator in this paper even though many LBP variants have been described in the literature [16]. Also note that the definition of Eq. (1) ignores border effects for readability sake and that only those pixels at which  $N_p$  is fully enclosed in the image are actually taken into account to compute the LBP histogram.

### B. Moment LBPs

Mirhashemi [19] proposes an LBP-based spectral feature using mathematical moments to characterize the reflectance spectrum shape. The LBP operator of Eq. (1) is no longer applied to pixel values but to moment values of the pixel spectral signatures.

Different moments can be extracted from the reflectance  $\{R_p(\lambda^k)\}_{k=1}^K$  sampled over  $K$  bands at each pixel  $p$ . Raw and central type-I moments of order  $n \in \mathbb{N}$  are defined as

$$M_n(p) = \sum_{k=1}^K (\lambda^k)^n R_p(\lambda^k) \quad (3)$$

and

$$\mu_n(p) = \sum_{k=1}^K \left( \lambda^k - \frac{M_1(p)}{M_0(p)} \right)^n R_p(\lambda^k).$$

Type-II moments are estimated moments of the probability density function from which reflectance values are sampled. Raw and central type-II moments are expressed as

$$\hat{M}_n(p) = \frac{1}{K} \sum_{k=1}^K (R_p(\lambda^k))^n \quad (4)$$

and

$$\hat{\mu}_n(p) = \frac{1}{K} \sum_{k=1}^K \left( R_p(\lambda^k) - \frac{\hat{M}_1(p)}{\hat{M}_0(p)} \right)^n.$$

Alternatively, these moments can be computed from the reflectance normalized by its  $L_1$ -norm at each pixel  $p$ :  $r_p(\lambda^k) =$

$\frac{R_p(\lambda^k)}{\sum_{i=1}^K R_p(\lambda^i)}$ . We then denote type-I and type-II raw moments as  $m_n(p)$  and  $\hat{m}_n(p)$ .

Mirhashemi assesses the texture classification performance of all the possible moment-based features (namely the 38 moment LBP histograms obtained for  $n = 1..6$ ), either considered alone or concatenated in 2- or 3-feature combinations. The most powerful combinations use three features based on the following moments:  $m_1(p)$  or  $\hat{m}_1(p)$ ,  $M_1(p)$  or  $\hat{M}_1(p)$ , and  $\mu_3(p)$ ,  $\mu_5(p)$ ,  $\hat{\mu}_3(p)$ , or  $\hat{\mu}_5(p)$  [19]. The texture feature is then a concatenated histogram with  $3 \cdot 2^P$  bins.

### C. Map-based LBPs

Dubey *et al.* [20] propose two kinds of LBP operators that can theoretically be applied to any  $K$ -channel image. These operators use the spectral information in the encoding scheme by testing the sum of the marginal comparison patterns between each pixel  $p$  and its neighbors over all channels. The adder-based LBPs  $\{maLBP^m\}_{m=0}^K$  are defined as

$$maLBP^m[\mathbf{I}](p) = \sum_{q \in N_p} \begin{cases} 2^{\epsilon(q)} & \text{if } \sum_{k=1}^K s(I_q^k, I_p^k) = m, \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

The decoder-based LBPs  $\{mdLBP^n\}_{n=0}^{2^K-1}$  are defined as

$$mdLBP^n[\mathbf{I}](p) = \sum_{q \in N_p} \begin{cases} 2^{\epsilon(q)} & \text{if } \sum_{k=1}^K s(I_q^k, I_p^k) \cdot 2^{(K-k)} = n, \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

The concatenation of the operator's histograms provides the final feature of size  $(K+1) \cdot 2^P$  (maLBPs) or  $2^K \cdot 2^P$  (mdLBPs).

### D. Luminance-spectral LBPs

By analogy with the luminance-chrominance model for a color image, a multispectral image can be represented as both a panchromatic channel and the joint information computed from two or more channels. The PPI that carries the spatial information of the luminance is computed as the average value over all channels at each pixel  $p$  [13]:

$$\bar{I}_p = \frac{1}{K} \sum_{k=1}^K I_p^k. \quad (7)$$

To form the final feature, the histogram of the output of the LBP operator applied to  $\bar{I}$  is concatenated with a histogram based on the spectral content according to one of the following propositions that we extend here to the multispectral domain:

- Cusano *et al.* [21] define the local color contrast (LCC) operator that depends on the angle between the value of a pixel  $p$  and the average value  $\bar{I}_p = \frac{1}{P} \sum_{q \in N_p} I_q$  of its neighbors in the spectral domain:

$$LCC[\mathbf{I}](p) = \cos^{-1} \left( \frac{\langle \mathbf{I}_p, \bar{\mathbf{I}}_p \rangle}{\|\mathbf{I}_p\| \cdot \|\bar{\mathbf{I}}_p\|} \right), \quad (8)$$

where  $\langle \cdot, \cdot \rangle$  and  $\|\cdot\|$  denote the inner product and the Euclidean norm. The histogram of  $LBP[\bar{I}]$  is concatenated to that of  $LCC[\mathbf{I}]$  quantized on  $2^P$  bins to provide the final feature of size  $2 \cdot 2^P$ .

- Lee *et al.* [22] consider  $\mathbf{I}$  in a  $K$ -dimensional space and compute spectral angular patterns between bands at each pixel. Specifically, for each pair of bands  $(k, l) \in \{1, \dots, K\}^2, k \neq l$ ,

the authors apply the LBP operator to the image  $\tilde{I}^{k,l}$  defined at each pixel  $p$  as the angle between the axis of the band  $k$  and the projection of  $\mathbf{I}_p$  onto the plane associated with bands  $k$  and  $l$ :

$$\tilde{I}_p^{k,l} = \tan^{-1} \left( \frac{I_p^k}{I_p^l + \eta} \right), \quad (9)$$

where  $\eta$  is a small-valued constant to avoid division by zero. The histogram of  $LBP[\tilde{I}]$  is concatenated to the  $K(K-1)$  histograms of  $\{LBP[\tilde{I}^{k,l}]\}_{k,l=1,k \neq l}^K$  to provide the final feature of size  $(1 + K(K-1)) \cdot 2^P$ .

### E. Opponent band LBPs

To fully take spectral correlation into account, Mäenpää *et al.* [23] apply the opponent color LBP (OCLBP) operator to each pair of channels of a color image. This operator can be directly generalized as the opponent band LBP (OBLBP) applied to each pair of channels  $(I^k, I^l)$ ,  $(k, l) \in \{1, \dots, K\}^2$ , of a multispectral image:

$$OBLBP^{k,l}[\mathbf{I}](p) = \sum_{q \in N_p} s(I_q^l, I_p^k) \cdot 2^{\epsilon(q)}. \quad (10)$$

Bianconi *et al.* [24] similarly consider both intra- and inter-channel information but with a different thresholding scheme. Their improved OBLBP (IOBLBP) operator uses a local average value rather than the sole central pixel value as threshold:

$$IOBLBP^{k,l}[\mathbf{I}](p) = \sum_{q \in \{p\} \cup N_p} s(I_q^l, \bar{I}_p^k) \cdot 2^{\epsilon(q)}, \quad (11)$$

where  $\bar{I}_p^k = \frac{1}{p+1} \sum_{r \in \{p\} \cup N_p} I_r^k$  and  $\epsilon(p) = P$ .

In both cases, the texture feature is the concatenation of the  $K^2 2^P$ -bin histograms of  $\{(I)OBLBP^{k,l}[\mathbf{I}]\}_{k,l=1}^K$ .

## 3. MSFA TEXTURE FEATURE BASED ON LOCAL BINARY PATTERNS

We intend to design an LBP-like operator to characterize multispectral texture images directly from the raw image acquired by snapshot cameras. A similar approach was proposed by Losson and Macaire [25] for color texture representation from raw CFA images. Rather than a straightforward extension that would neglect spectral correlation, we here propose a new operator dedicated to raw MSFA images and inspired by OBLBPs.

We first present the raw image in details in this section, including the specific neighborhoods defined by the MSFA. Then we describe our operator based on these MSFA neighborhoods, and we explain how it is related to OBLBPs.

### A. MSFA neighborhoods

An MSFA associates a single spectral band with each pixel. It can be defined as a function  $MSFA: S \rightarrow \{1, \dots, K\}$  over the set  $S$  of all pixels. Let  $S^k = \{p \in S, MSFA(p) = k\}$  be the pixel subset where the MSFA samples the band  $k$ , such that  $S = \bigcup_{k=1}^K S^k$ . In other words,  $S^k$  is the subset of pixels that are associated with the band  $k$ . Figure 3 shows the example of the IMEC16 MSFA and one among its 16 pixel subsets.

For a given pixel  $p \in S^k$ ,  $k \in \{1, \dots, K\}$ , let  $B^k = \{l \in \{1, \dots, K\}, MSFA(q) = l\}_{q \in N_p}$  be the set of bands that are associated with the neighboring pixels in  $N_p$  according to the MSFA. Note that  $N_p$  is always composed of pixels with the same associated bands whatever the location of  $p$ . Moreover, we assume that

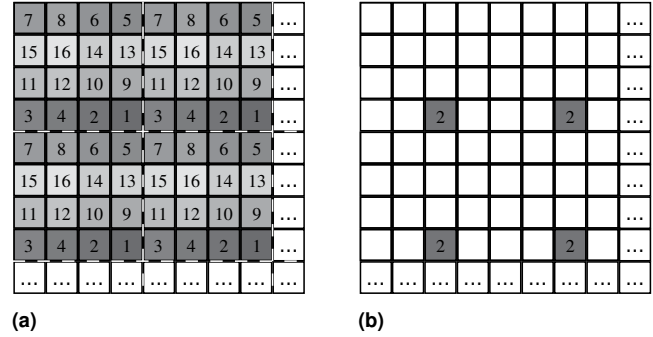


Fig. 3. (a) IMEC16 MSFA and (b) its  $S^2$  pixel subset. Dashes on (a) bound the basic pattern of Fig. 1c.

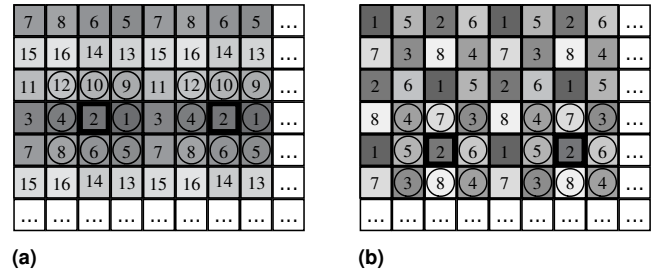


Fig. 4. Neighborhood  $N_p$  defined by the support  $\mathcal{N}^{8,1}$  for two pixels  $p \in S^2$  (bold squares) in (a) IMEC16 and (b) VisNIR8 MSFAs, with associated bands  $B^2$  shown in solid circles.

any neighbor  $q \in N_p$  is always associated with the same band for a given relative position with respect to  $p$  in the MSFA pattern. A necessary but not sufficient condition for this assumption to be fulfilled is spectral consistency. Then, the neighborhood of  $p \in S^k$  can be decomposed into

$$N_p = \bigcup_{l \in B^k} N_p^{k,l}, \quad (12)$$

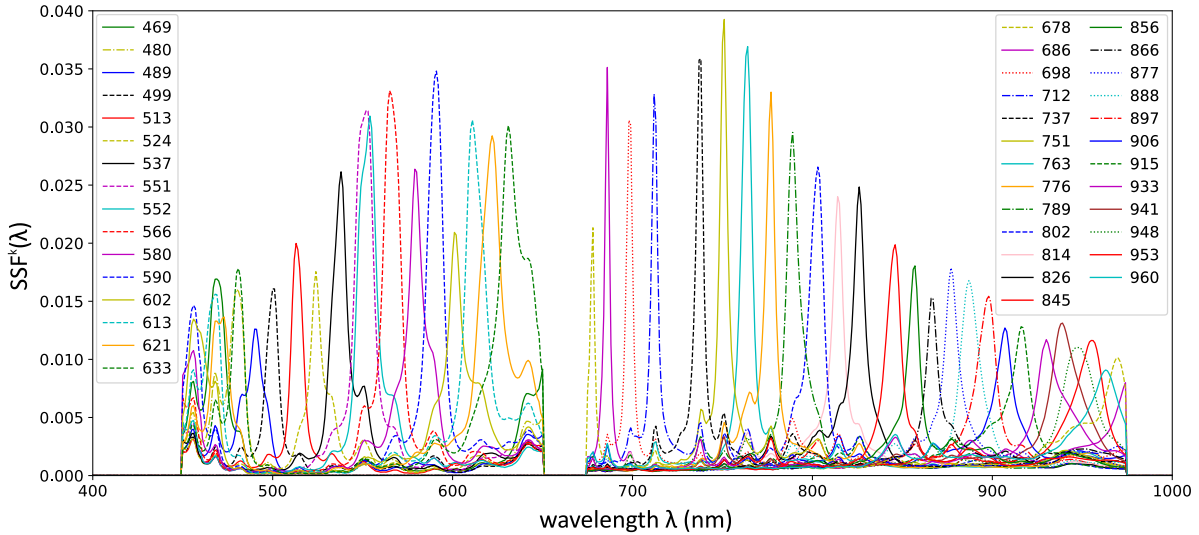
where  $N_p^{k,l} = N_p \cap S^l$  is the MSFA-based neighborhood made of the neighboring pixels of  $p$  that belong to  $S^l$ . Let us notice that  $N_p^{k,l} \neq \emptyset \iff l \in B^k$  and stress out that  $B^k$  and  $N_p^{k,l}$  both depend on  $\mathcal{N}^{P,d}$  and on the basic MSFA pattern.

For illustration purposes, let us consider the IMEC16 and VisNIR8 MSFAs and focus on the  $3 \times 3$  neighborhood defined by the support  $\mathcal{N}^{8,1}$  as shown in Fig. 4. In the IMEC16 MSFA of Fig. 4a, the neighbors of any pixel  $p \in S^2$  are associated with the bands  $B^2 = \{12, 10, 9, 4, 1, 8, 6, 5\}$  and  $|N_p^{2,l}| = 1$  for all  $l \in B^2$ , where  $|\cdot|$  is the cardinal operator. In the VisNIR8 MSFA of Fig. 4b, we have  $B^2 = \{4, 7, 3, 5, 6, 8\}$  and  $|N_p^{2,l}| = 1$  for  $l \in \{5, 6, 7, 8\}$ , but  $|N_p^{2,3}| = |N_p^{2,4}| = 2$ .

### B. MSFA-based LBPs

A snapshot multispectral camera provides a raw image  $I^{raw}$  in which a single band is associated with each pixel according to the MSFA. Then,  $I^{raw}$  can be seen as a spectrally-sampled version of the reference fully-defined image  $\mathbf{I} = \{I^k\}_{k=1}^K$  (that is unavailable in practice) according to the MSFA:

$$\forall p \in S, I_p^{raw} = I_p^{MSFA(p)}. \quad (13)$$



**Fig. 5.** Normalized SSFs of IMEC16 (left) and IMEC25 (right) cameras. Captions: band center wavelengths  $\{\lambda^k\}_{k=1}^K$  in ascending order.

To design a texture feature dedicated to the raw image, let us first consider applying the basic LBP operator of Eq. (1) directly to  $I^{raw}$  considered as a gray-level image:

$$MLBP[I^{raw}](p) = \sum_{q \in N_p} s(I_q^{raw}, I_p^{raw}) \cdot 2^{\epsilon(q)}. \quad (14)$$

The LBP operator is here renamed as MSFA-based LBP (MLBP) to make clear the key difference introduced by its application to  $I^{raw}$  and its dependency upon the considered MSFA. Unlike Eq. (1), Eq. (14) combines the spectral information of  $B^{MSFA(p)}$ , i.e, the different bands that are associated with the neighbors of  $p$ .

Because this set of bands depends on the band  $MSFA(p)$  associated with  $p$ , we separately consider each pixel subset  $S^k$  to compute the LBP histogram. Specifically, we compute the histogram of  $MLBP[I^{raw}]$  for each band  $k \in \{1, \dots, K\}$ :

$$h^k[MLBP[I^{raw}]]: [0, 2^P - 1] \rightarrow \{0, \dots, |S^k|\} \\ j \mapsto \left| \left\{ p \in S^k, MLBP[I^{raw}](p) = j \right\} \right|. \quad (15)$$

Let us point out that only pixels in  $S^k$  are considered to compute the  $k$ -th histogram. The concatenation of all the  $K$  histograms provides the final feature of size  $K \cdot 2^P$ .

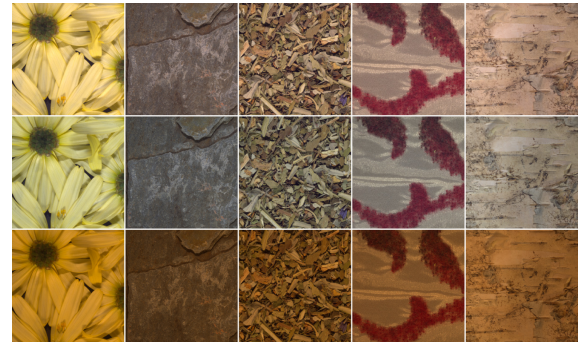
### C. Relation between MSFA-based and opponent band LBPs

To show that the MSFA-based LBP defined by Eq. (14) bears an analogy to OBLBP (see Eq. (10)), let us consider its output as the direct sum of the sparse outputs of the same operator restrictively applied to each pixel subset  $S^k$ :

$$\text{im} \left\{ MLBP[I^{raw}] \right\} = \bigoplus_{k=1}^K \text{im} \left\{ MLBP|_{S^k}[I^{raw}] \right\}, \quad (16)$$

where  $\text{im}\{\cdot\}$  is a function output. According to the definition of  $S^k$ , we have  $I_p^{raw} = I_p^k$  for each pixel  $p \in S^k$ . From Eq. (14) and the decomposition of the neighborhood  $N_p$  according to Eq. (12), we can then express  $MLBP|_{S^k}$  from  $\{I^k\}_{k=1}^K$  as

$$MLBP|_{S^k}[I^{raw}](p) = \sum_{l \in B^k} \sum_{q \in N_p^{k,l}} s(I_q^l, I_p^k) \cdot 2^{\epsilon(q)}. \quad (17)$$



**Fig. 6.** Samples of HyTexiLa texture images rendered in sRGB under E (top), D65 (middle) and A (bottom row) illuminants, from left to right: vegetation, stone, food, textile, and wood.

Therefore, MLBP is related to OBLBP since both operators take opponent bands into account. But unlike OBLBP that considers any band  $l$  at all the neighbors of  $p$ , each MLBP code combines opponent band information from the  $|B^k|$  bands that are available at the neighbors of  $p \in S^k$ .

## 4. EXPERIMENTS

### A. Experimental settings

#### A.1. Texture dataset and classification scheme

In order to perform texture classification, we use HyTexiLa database that contains the reflectance images of 112 textured materials (see Fig. 6) and is currently the most suitable database in the VisNIR domain [26]. We consider three CIE standard illuminants (E, D65, and A) to simulate three fully-defined images from the reflectance of each texture. Finally, we sample these images according to several MSFAs to simulate the raw images that would be acquired by various snapshot multispectral cameras.

We consider the four MSFAs of Fig. 1 since they are either related to research works with detailed publications (Vis5, VisNIR8) or available in consumer cameras (IMEC). The spectral sensitivity functions (SSFs) of Vis5 and VisNIR8 can be found in the respective papers [27] and [6]. Figure 5 shows the normal-

ized SSFs of IMEC16 and IMEC25 cameras [7] that respectively sample the visible light around  $\lambda^1 = 469 \text{ nm}, \dots, \lambda^{16} = 633 \text{ nm}$  and NIR light around  $\lambda^1 = 678 \text{ nm}, \dots, \lambda^{25} = 960 \text{ nm}$ . This figure shows that the SSFs of bands with close centers may overlap to a large extent. Note that the SSFs include the optics and sensor sensitivities and the optical band-pass filter that removes spectral artifacts.

HyTexiLa reflectances are defined on 186 bands whose centers range from  $\lambda_{\min} = 405.37 \text{ nm}$  to  $\lambda_{\max} = 995.83 \text{ nm}$  with a step of  $3.19 \text{ nm}$ . From these data and the three illuminations provided by Thomas *et al.* [6] (that extend A and E illuminants to the NIR domain and simulate D65), we compute the images of illuminated textures as [13]

$$I_p^k = Q \left( \sum_{\lambda=\lambda_{\min}}^{\lambda_{\max}} E(\lambda) R_p(\lambda) \text{SSF}^k(\lambda) \right). \quad (18)$$

In this equation,  $Q(\cdot)$  is the radiance quantization function on 8 bits, and  $E(\lambda)$ ,  $R_p(\lambda)$ , and  $\text{SSF}^k(\lambda)$  are the illumination, the reflectance of the surface elements captured at pixel  $p$ , and the camera-specific SSF of the band  $k$ , respectively. The values of these discrete functions of  $\lambda$  are obtained at  $1\text{-nm}$  resolution by linear interpolation of the original values. To prevent saturation in the simulated images, illuminations and SSFs are scaled so that  $\max_{\lambda} E(\lambda) = 255$  and  $\max_k \sum_{\lambda} \text{SSF}^k(\lambda) = 1$ , and reflectance values above 1 (that correspond to specular surfaces) are clipped to 1. **This normalization practically corresponds to setting the integration time of the camera as the limit before saturation when a white patch is observed. This has an effect on the dynamic range of channels [28].**

For each of the considered four cameras and three illuminations, these simulations provide 112 raw images of size  $1024 \times 1024$  pixels. Each image is then split (leaving out the last four columns and rows) into 25 sub-images of size  $204 \times 204$  pixels, among which 12 are randomly picked for training and the 13 others for testing.

In order to determine the most discriminative texture feature, we retain the 1-nearest neighbor classifier coupled with the similarity measure based on intersection between histograms [29] since this classification scheme requires no parameter.

## A.2. MSFA neighborhoods

As explained in Sec. 3A, the neighbors of any pixel  $p$  are associated with different bands according to the MSFA. It is thus impossible to consider interpolated values in a circular neighborhood of  $p$  as is usually done for LBP-like operators. To avoid interpolation, we therefore consider the uniform spatial distance (hence square neighborhoods) rather than the Euclidean one. Moreover, LBP operators classically use neighborhoods with  $P = 8, 16$ , or  $24$  pixels. But  $P = 16$  with  $d = 3$  does not match the image lattice and requires interpolation, and  $P = 24$  would yield extremely large features. We therefore set  $P = 8$  and consider the three supports  $\mathcal{N}^{8,d}$  with uniform distance  $d \in \{1, 2, 3\}$  as shown in Fig. 7.

Figure 7 also shows that the number of bands available in the neighborhood of a pixel  $p$  generally depends on the distance  $d$  for a given MSFA. This number is formalized by  $|B^k|$ , where  $k = \text{MSFA}(p) \in \{1, \dots, K\}$  is the band associated with  $p$  (i.e.,  $p \in S^k$ ), and its dependency upon  $d$  is summarized in Tab. 1. In VisNIR8 for instance (see Fig. 7b), the neighborhood of  $p \in S^3$  contains eight different bands for  $d = 1$  and  $d = 3$ , but only the bands 3 and 4 for  $d = 2$ .  $|B^k|$  is also lower for  $d = 2$  with Vis5 and IMEC16 MSFAs but is constant to eight whatever  $d \in \{1, 2, 3\}$

$G_1$	R	$G_2$	Or	$G_1$	R	$G_2$	Or	...
B	$G_3$	Cy	$G_4$	B	$G_3$	Cy	$G_4$	...
$G_2$	Or	$G_1$	R	$G_2$	Or	$G_1$	R	...
Cy	$G_4$	B	$G_3$	Cy	$G_4$	B	$G_3$	...
$G_1$	R	$G_2$	Or	$G_1$	R	$G_2$	Or	...
B	$G_3$	Cy	$G_4$	B	$G_3$	Cy	$G_4$	...
$G_2$	Or	$G_1$	R	$G_2$	Or	$G_1$	R	...
Cy	$G_4$	B	$G_3$	Cy	$G_4$	B	$G_3$	...
...	...	...	...	...	...	...	...	...

(a)

1	5	2	6	1	5	2	6	...
7	3	8	4	7	3	8	4	...
2	6	1	5	2	6	1	5	...
8	4	7	3	8	4	7	3	...
1	5	2	6	1	5	2	6	...
7	3	8	4	7	3	8	4	...
2	6	1	5	2	6	1	5	...
8	4	7	3	8	4	7	3	...
...	...	...	...	...	...	...	...	...

(b)

7	8	6	5	7	8	6	5	...
15	16	14	13	15	16	14	13	...
11	12	10	9	11	12	10	9	...
3	4	2	1	3	4	2	1	...
7	8	6	5	7	8	6	5	...
15	16	14	13	15	16	14	13	...
11	12	10	9	11	12	10	9	...
3	4	2	1	3	4	2	1	...
...	...	...	...	...	...	...	...	...

(c)

19	20	18	17	2	19	20	18	17	2	...
11	12	10	9	3	11	12	10	9	3	...
7	8	6	5	4	7	8	6	5	4	...
23	24	22	21	1	23	24	22	21	1	...
15	16	14	13	25	15	16	14	13	25	...
19	20	18	17	2	19	20	18	17	2	...
11	12	10	9	3	11	12	10	9	3	...
7	8	6	5	4	7	8	6	5	4	...
23	24	22	21	1	23	24	22	21	1	...
15	16	14	13	25	15	16	14	13	25	...
...	...	...	...	...	...	...	...	...	...	...

(d)

19	20	18	17	2	19	20	18	17	2	...
11	12	10	9	3	11	12	10	9	3	...
7	8	6	5	4	7	8	6	5	4	...
23	24	22	21	1	23	24	22	21	1	...
15	16	14	13	25	15	16	14	13	25	...
...	...	...	...	...	...	...	...	...	...	...

**Fig. 7.** Neighborhood  $N_p$  of a pixel  $p$  (bold square) in (a) Vis5, (b) VisNIR8, (c) IMEC16, and (d) IMEC25 MSFAs, considering the supports  $\mathcal{N}^{8,1}$  (solid circles),  $\mathcal{N}^{8,2}$  (dashed), and  $\mathcal{N}^{8,3}$  (dotted).

**Table 1.** Number of available bands  $|B^k|$ ,  $k \in \{1, \dots, K\}$ , in the neighborhood of any pixel according to each MSFA and each distance.

MSFA	$d = 1$	$d = 2$	$d = 3$
Vis5	3 or 5	1 or 2	3 or 5
VisNIR8	6	2	6
IMEC16	8	3	8
IMEC25	8	8	8

with IMEC25 due to the large  $5 \times 5$  basic pattern of this MSFA. Note that  $|B^k|$  reflects the degree to which spectral correlation is taken into account by an MSFA neighborhood.

The basic pattern of the Vis5 MSFA (see Fig. 1a) is particular because of its single dominant  $G$  band. Unlike in the other considered MSFAs,  $|B^k|$  in Vis5 depends on  $d$  but also on  $p$  (i.e., on  $k$ ) (see Tab.1). Considering  $d = 1$  for instance, the neighbors of a pixel  $p_1 \in S^G$  belong to all the bands (hence  $|B^G| = 5$ ) while those of a pixel  $p_2 \in S^{Cy}$  belong to  $B^{Cy} = \{R, G, Or\}$ . Moreover, Vis5 contradicts our assumption that a neighbor of  $p \in S^k$  is always associated with the same band for a given relative position whatever the location of  $p$ . Indeed, for two pixels associated with the  $G$  band, vertical neighbors may either be associated with  $R$  and  $Or$  or with  $B$  and  $Cy$ . To fulfill our assumption and compute MLBP with Vis5, we therefore split  $S^G$  into four pixel subsets  $\{S_i^G\}_{i=1}^4$  as shown in Fig. 7a. The information of  $\{MLBP|_{S_i^G}, [I^{raw}]_{i=1}^4\}$  is then merged into a single histogram  $h^G [MLBP [I^{raw}]]$  for the  $G$  band.

### A.3. Feature extraction

Table 2 summarizes the sizes of the texture features described in Sec. 2 as the size of each histogram (that depends on  $P$ ) and the number of histograms (that depends on  $K$ ). Setting  $P = 8$  makes the histogram size to be 256 but the mdLBP operator provides a prohibitively large number of histograms when  $K \geq 16$ . All approaches but mdLBP are hence tested against our MSFA-based LBP in the experiments. Besides, among the 16 moment combinations from the format  $\{m_1 | \hat{m}_1\} \{M_1 | \hat{M}_1\} \{\mu_{3|5} | \hat{\mu}_{3|5}\}$  (see Sec. 2B), we only retain  $m_1 M_1 \mu_3$  whose LBP histogram provides the best classification result on average over all the experiments.

The number of histograms may impact classification accuracy and computational burden. The approaches can then be divided into three groups, depending on whether this number is constant (Cusano and Moment LBPs), proportional to  $K$  (Marginal LBPs, maLBP, and MLBP), or to  $K^2$  (OBLBP, IOBLBP, and Lee LBPs).

Computation cost also deserves some attention as an implementation-independent indication of the required processing time. The last two columns of Tab. 2 show this cost as the number of elementary operations per pixel required to compute a feature. This estimation includes all arithmetic operations at the same cost of 1 but excludes array indexing.

All features but ours first require to estimate a fully-defined multispectral image by demosaicing. In our experiments, we only consider WB [10] that is both the most simple and generic method (see Sec. 1B) and PPID [13] that provides the best demosaicing results in most cases. We have adapted PPID to the Vis5 MSFA and we retain it because it globally yields better classification results than the dedicated guided filtering method provided by [27]. Since demosaicing is neither our main concern nor the greedier feature computation step, we minimally evaluate its

number of operations as that of the weighted average of two values required by WB to estimate each missing value at a pixel, namely  $4(K - 1)$ .

The feature computation costs given in the last column of Tab. 2 result from the equation(s) of each feature recalled in the second column. As previously stated, the computation cost of mdLBP is prohibitive. Our MLBP-based feature requires 24 operations per pixel. In contrast, the cost of Marginal LBPs, maLBP, and Cusano LBPs grows with  $K$ , and that of other approaches with  $K^2$ . By considering both the feature size and computation cost, Lee LBPs, OBLBP and IOBLBP are the most greedy features. MLBP is the most efficient and is represented by the same number of histograms as Marginal LBPs and maLBP. This should be kept in mind while analyzing the classification results.

## B. Experimental results

### B.1. Accuracy vs. computation cost

We first propose a study to highlight the above remark about feature computation costs. Let us consider the case  $d = 1$  and the D65 illuminant, and assess the classification accuracy provided by each feature with regard to its cost. For all approaches except MLBP, WB is chosen to demosaic the raw image.

Figure 8 separately shows the results for the four considered MSFAs. OBLBP globally outperforms other features but at a very high computation cost. MLBP provides only slightly lower results than OBLBP for Vis5 and VisNIR8 MSFAs and similar results for IMEC MSFAs, at about a  $K^2$  times smaller cost. Considering features with comparable costs, Lee LBPs and IOBLBP perform worse than OBLBP. Moment LBPs generally provide fair results with regard to the other three features with moderate costs (marginal LBPs, maLBP, and Cusano LBPs). MLBP clearly outperforms the latter four features in all cases with the benefit of reduced computation requirements.

### B.2. Classification results and discussion

We now extensively assess the performances of our MLBP-based feature with respect to those of other features in various experimental conditions. Table 3 shows the classification accuracies provided by the different features for each of the four MSFAs using three illuminants (E, D65, and A), three distance values ( $d \in \{1, 2, 3\}$ ), and two demosaicing methods (WB and PPID). For each combination of experimental settings (illuminant, distance, and demosaicing method), the better classification accuracies than that of our descriptor are highlighted as bold.

Let us first study the behavior of our MLBP-based feature with respect to the different settings. The neighborhood distance  $d$  slightly influences MLBP performances for IMEC25 MSFA (see Tab. 3d) whereas  $d = 2$  provides clearly lower performances than  $d \in \{1, 3\}$  for IMEC16 (see Tab. 3c). This was expected since  $d$  determines the number  $|B^k|$  of available bands in a pixel neighborhood, that does not depend on  $d$  for IMEC25 but is reduced from 8 to 3 bands when  $d = 2$  for IMEC16 (see Tab. 1). Regarding Vis5 and VisNIR8, there is no systematic performance loss for  $d = 2$  though  $|B^k|$  is reduced. This is because the poor information about spectral (inter-channel) correlation in these cases is completed by taking spatial (intra-channel) correlation into account. When  $d = 2$  indeed,  $k \in B^k$  (see Figs. 7a and 7b), which means that MLBP takes both intra- and inter-channel correlation into account. No such case occurs with IMEC16 and IMEC25 (see Figs. 7c and 7d) with which MLBP only takes inter-channel correlation into account because  $k \notin B^k$  whatever the value of  $d$ . Two other outstanding results are obtained with Vis5 MSFA and illuminant A when  $d \in \{1, 3\}$ . We explain these

**Table 2.** Feature size (histogram size and number of concatenated histograms) and required number of operations per pixel for each approach according to the number  $K$  of spectral bands.

Approach	Eq.	Feature size		Number of operations	
		Hist. size	Number of hist.	Demosaicing	Feature computation
Marginal LBPs [17]	Eq. (1)	$2^P$	$K$	$4K - 4$	$24K$
Moment LBPs [19]	Eq. (3), Eq. (4)		3		$K^2 + 10K + 69$
maLBP [20]	Eq. (5)		$K + 1$		$17K + 8$
mdLBP [20]	Eq. (6)		$2^K$		$2^K + 24K + 7$
Cusano LBPs [21]	Eq. (8)		2		$15K + 27$
Lee LBPs [22]	Eq. (9)		$1 + K(K - 1)$		$27K^2 - 26K + 24$
OBLBP [23]	Eq. (10)		$K^2$		$24K^2$
IOBLBP [24]	Eq. (11)	$2^{P+1}$	$K^2$		$27K^2 + 10K$
MLBP (ours)	Eq. (14), Eq. (15)	$2^P$	$K$	0	24

lower accuracies by the shape of illuminant A that increases with respect to the wavelength, which makes pixel values in  $B$  and  $C_y$  channels significantly lower than in  $R$  and  $O_r$ . Besides, values of pixels in  $S^B \cup S^{C_y}$  are always compared to neighboring pixels in  $S^{O_r} \cup S^R$  (or vice versa) when  $d \in \{1, 3\}$ , whereas  $R$  values are compared to  $O_r$  values and  $B$  values to  $C_y$  values when  $d = 2$  (see Fig. 7a). Therefore, due to the particular band arrangement in Vis5 MSFA, MLBP is less discriminative and performances are notably reduced when  $d \in \{1, 3\}$  with A illuminant.

Let us now compare the performances reached by our descriptor with those of other approaches. Table 3 shows that, except with Vis5 MSFA images simulated under A illuminant, our MLBP-based feature always outperforms approaches with either smaller (Cusano and Moment LBPs) or similar-size features (Marginal LBPs and maLBP). Moreover, our lightweight approach obtains close results to greedy ones (Lee LBP, OBLBP and IOLBP), especially with IMEC MSFAs, and even performs better than them in 95 out of the 216 tested cases. The best accuracy reached by MLBP is 97.32% (with IMEC16 MSFA under D65 illuminant using  $d = 1$ ) while the best descriptor (OBLBP) reaches 97.60% (with the same settings and PPID demosaicing).

## 5. CONCLUSION

In this paper, we introduce a conceptually simple and highly-discriminative LBP-based feature for multispectral raw images. In addition to its algorithmic simplicity, our operator is directly applied to raw images, which allows it to avoid the demosaicing step and keeps its computational cost low. We perform extensive experiments of texture classification on simulated multispectral images with four well-referenced MSFAs. The results show that the proposed approach outperforms existing ones using features of similar sizes, and provides comparable results to that of features with large size and high computational cost.

Future works will further study how our feature embeds spatial and spectral correlations according to the MSFA and neighborhood parameters. Since MLBP is a small-size feature, there is room for additional correlation information that could still improve its classification results. For instance, it could be made more robust to the neighborhood distance by concatenating several MLBP histograms computed with different distances. Another approach to improve how our feature accounts for spatial correlation would be to use a demosaiced dominant channel. Other investigations could at last focus on the spectral distance

among the considered neighbors.

## FUNDING

National French platform ANR-11-EQPX-23 IrDIVE; National French project CNRS-INS2I-JCJC-2017 MOSAIC.

## REFERENCES

- N. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," *Optical Engineering* **52**(9), 090901 (2013).
- M. A. Martínez, E. M. Valero, J. Hernández-Andrés, J. Romero, and G. Langfelder, "Combining transverse field detectors and color filter arrays to improve multispectral imaging systems," *Applied Optics* **53**(13), C14–C24 (2014).
- P.-J. Lapray, X. Wang, J.-B. Thomas, and P. Gouton, "Multispectral Filter Arrays: Recent Advances and Practical Implementation," *Sensors* **14**(11), 21626-21659 (2014).
- Y. Monno, M. Tanaka, and M. Okutomi, "Multispectral demosaicking using adaptive kernel upsampling," in *Procs. 18th IEEE International Conference on Image Processing (ICIP 2011)*, pp. 3157-3160 (Brussels, Belgium, 2011).
- L. Miao, H. Qi, R. Ramanath, and W. Snyder, "Binary Tree-based Generic Demosaicking Algorithm for Multispectral Filter Arrays," *IEEE Transactions on Image Processing* **15**(11), 3550-3558 (2006).
- J.-B. Thomas, P.-J. Lapray, P. Gouton, and C. Clerc, "Spectral Characterization of a Prototype SFA Camera for Joint Visible and NIR Acquisition," *Sensors* **16**(7), 993 (2016).
- B. Geelen, N. Tack, and A. Lambrechts, "A compact snapshot multispectral imager with a monolithically integrated per-pixel filter mosaic," in *Procs. SPIE: Advanced Fabrication Technologies for Micro/Nano Optics and Photonics VII*, vol. 8974, p. 89740L (San Francisco, CA, USA, 2014).
- J. Pichette, A. Laurence, L. Angulo, F. Lesage, A. Bouthillier, D. Nguyen, and F. Leblond, "Intraoperative video-rate hemodynamic response assessment in human cortex using snapshot hyperspectral optical imaging," *Neurophotonics* **3**(4), 045003 (2016).
- C. Winkens, V. Kobelt, and D. Paulus, "Robust features for snapshot hyperspectral terrain-classification," in *Procs. 17th International Conference on Computer Analysis of Images and Patterns (CAIP 2017)*, vol. 10424 of *Lecture Notes in Computer Science*, pp. 16-27 (Ystad, Sweden, 2017).
- J. Brauers and T. Aach, "A Color Filter Array Based Multispectral Camera," in *12. Workshop Farbbildverarbeitung*, pp. 55-64 (Illmenau, Germany, 2006).
- J. Mizutani, S. Ogawa, K. Shinoda, M. Hasegawa, and S. Kato, "Multispectral demosaicking algorithm based on inter-channel correlation,"

**Table 3.** Classification accuracy (%) of the different approaches for each experimental setting (illuminant, neighborhood distance, demosaicing method) and each MSFA: (a) Vis5 ( $K = 5$ ), (b) VisNIR8 ( $K = 8$ ), (c) IMEC16 ( $K = 16$ ), and (d) IMEC25 ( $K = 25$ ).

(a)

Approach	E						D65						A					
	$d = 1$		$d = 2$		$d = 3$		$d = 1$		$d = 2$		$d = 3$		$d = 1$		$d = 2$		$d = 3$	
	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID
Marginal LBPs [17]	87.71	86.95	88.39	87.50	88.67	88.46	87.36	87.23	88.60	88.26	88.46	88.60	<b>84.89</b>	<b>84.62</b>	87.71	87.84	88.05	88.19
Moment LBPs [19]	90.11	88.05	88.26	88.67	86.74	86.13	90.32	88.53	89.29	89.01	88.53	87.36	<b>87.98</b>	<b>87.16</b>	86.74	86.74	85.99	85.71
maLBP [20]	82.83	84.55	86.06	86.40	85.65	86.95	81.11	82.01	84.34	84.89	85.30	86.13	81.66	82.21	85.99	86.20	86.88	88.32
Cusano LBPs [21]	87.64	87.36	88.60	88.26	88.80	89.15	86.88	87.36	88.46	88.39	89.29	89.22	<b>88.80</b>	<b>89.01</b>	90.52	90.25	<b>90.32</b>	<b>90.87</b>
Lee LBPs [22]	93.41	93.75	<b>95.40</b>	<b>95.12</b>	<b>95.60</b>	94.85	92.99	93.13	<b>94.57</b>	<b>94.44</b>	95.47	95.47	<b>91.48</b>	<b>91.96</b>	<b>94.23</b>	<b>93.82</b>	<b>94.64</b>	<b>94.37</b>
OBLBP [23]	<b>97.39</b>	<b>97.46</b>	<b>97.39</b>	<b>97.53</b>	<b>97.12</b>	<b>97.46</b>	<b>96.84</b>	<b>96.70</b>	<b>97.18</b>	<b>97.39</b>	<b>97.12</b>	<b>97.39</b>	<b>94.64</b>	<b>94.51</b>	<b>95.67</b>	<b>95.95</b>	<b>96.29</b>	<b>96.57</b>
IOBLBP [24]	<b>96.02</b>	<b>95.95</b>	<b>95.60</b>	<b>96.29</b>	<b>96.70</b>	<b>96.09</b>	<b>95.05</b>	94.57	<b>95.81</b>	<b>96.15</b>	96.36	96.15	<b>91.55</b>	<b>92.24</b>	<b>93.06</b>	<b>93.41</b>	<b>94.57</b>	<b>94.09</b>
MLBP (ours)	93.82		92.86		94.99		94.92		92.45		96.63		84.48		92.99		89.90	

(b)

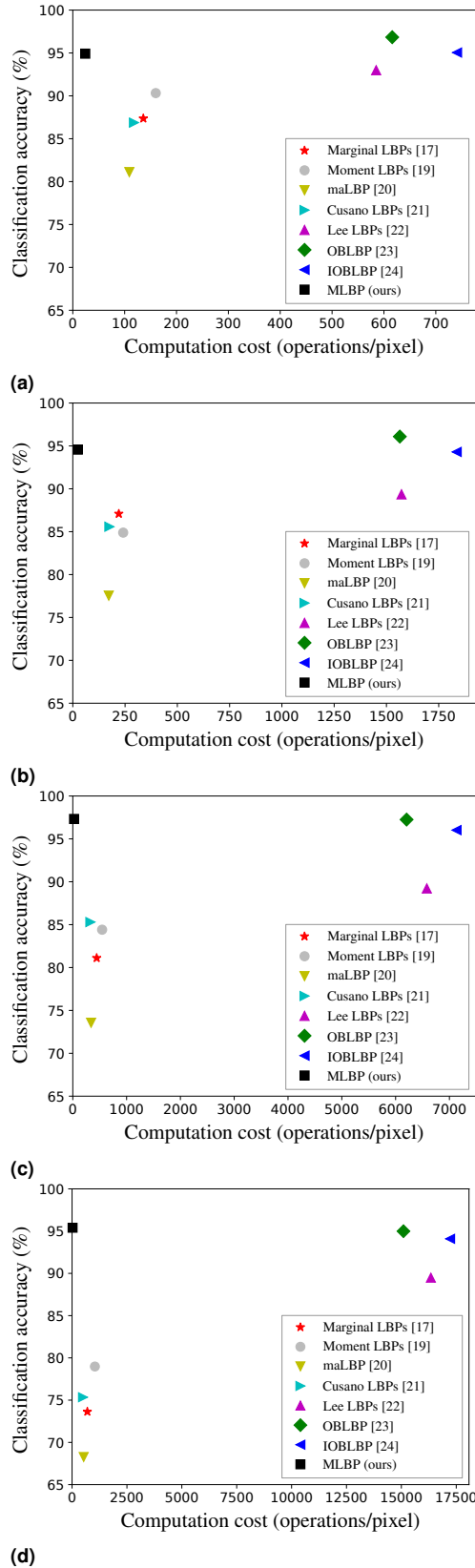
Approach	E						D65						A					
	$d = 1$		$d = 2$		$d = 3$		$d = 1$		$d = 2$		$d = 3$		$d = 1$		$d = 2$		$d = 3$	
	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID
Marginal LBPs [17]	84.89	86.40	85.44	85.78	86.06	86.13	87.09	88.94	87.50	88.67	87.64	88.53	84.00	85.51	85.03	85.37	85.78	85.51
Moment LBPs [19]	86.26	86.40	85.85	85.37	86.06	84.75	84.89	84.62	85.71	84.55	85.51	85.23	82.76	83.04	83.93	83.79	84.20	83.79
maLBP [20]	77.75	77.88	79.46	79.33	82.55	79.95	77.54	79.33	80.22	81.11	79.95	79.95	78.23	76.99	79.12	78.78	81.46	80.29
Cusano LBPs [21]	82.49	82.55	83.93	83.72	84.34	85.03	85.58	85.51	87.16	87.09	87.16	87.57	80.77	80.43	82.42	82.01	82.49	83.17
Lee LBPs [22]	91.28	91.69	93.96	93.89	<b>95.05</b>	<b>95.26</b>	89.35	91.41	91.83	91.69	93.48	93.41	91.62	92.86	93.82	93.89	<b>95.26</b>	<b>94.16</b>
OBLBP [23]	<b>96.29</b>	<b>96.91</b>	<b>96.63</b>	<b>96.57</b>	<b>96.57</b>	<b>96.63</b>	<b>96.09</b>	<b>96.22</b>	<b>96.15</b>	<b>95.67</b>	<b>96.09</b>	<b>95.54</b>	<b>95.19</b>	<b>95.40</b>	<b>95.47</b>	<b>95.12</b>	<b>95.12</b>	<b>95.33</b>
IOBLBP [24]	<b>94.09</b>	93.61	94.64	94.30	94.44	<b>94.57</b>	94.30	94.02	93.82	93.61	94.23	93.48	92.10	<b>93.34</b>	92.65	92.31	92.86	92.51
MLBP (ours)	93.75		95.67		94.57		94.57		94.71		94.64		93.06		94.02		93.20	

(c)

Approach	E						D65						A					
	$d = 1$		$d = 2$		$d = 3$		$d = 1$		$d = 2$		$d = 3$		$d = 1$		$d = 2$		$d = 3$	
	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID
Marginal LBPs [17]	81.59	85.10	83.79	86.54	83.86	86.20	81.11	84.89	83.72	86.95	85.30	87.16	80.70	84.89	83.10	85.92	85.71	87.50
Moment LBPs [19]	85.30	89.56	87.50	90.32	87.71	90.04	84.41	88.67	88.12	89.42	87.02	89.90	84.34	88.05	88.05	89.84	88.05	89.15
maLBP [20]	74.93	78.57	75.21	76.17	76.51	78.57	73.56	79.12	75.41	77.13	75.34	80.08	75.07	77.13	74.52	73.97	76.51	74.86
Cusano LBPs [21]	84.07	86.40	86.20	87.98	85.92	87.98	85.30	86.20	86.33	87.91	86.54	87.02	85.44	86.61	87.02	88.19	87.16	88.67
Lee LBPs [22]	89.77	95.40	<b>92.45</b>	<b>96.02</b>	93.68	95.95	89.22	95.26	92.58	<b>95.95</b>	93.41	95.81	87.09	94.44	<b>90.93</b>	<b>94.09</b>	92.93	94.64
OBLBP [23]	<b>96.91</b>	<b>96.57</b>	<b>96.22</b>	<b>96.77</b>	96.09	<b>96.50</b>	97.25	<b>97.60</b>	<b>97.05</b>	<b>97.05</b>	96.84	96.98	<b>95.67</b>	<b>96.09</b>	<b>95.47</b>	<b>95.81</b>	95.05	<b>95.95</b>
IOBLBP [24]	95.40	<b>96.36</b>	<b>95.19</b>	<b>95.88</b>	95.33	95.47	96.02	96.91	<b>95.88</b>	<b>96.57</b>	95.74	96.29	93.54	94.71	<b>94.44</b>	<b>94.78</b>	93.61	94.64
MLBP (ours)	96.22		90.87		96.29		97.32		94.92		97.05		95.33		90.32		95.19	

(d)

Approach	E						D65						A					
	$d = 1$		$d = 2$		$d = 3$		$d = 1$		$d = 2$		$d = 3$		$d = 1$		$d = 2$		$d = 3$	
	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID	WB	PPID
Marginal LBPs [17]	73.15	78.57	74.93	78.71	75.62	78.50	73.63	78.71	75.76	80.01	77.47	79.81	72.87	78.09	74.45	78.43	75.82	77.54
Moment LBPs [19]	79.46	83.17	80.56	83.72	80.77	84.27	78.98	84.20	81.25	85.51	81.46	85.16	78.71	83.31	80.15	83.93	80.36	84.07
maLBP [20]	63.67	78.64	61.26	78.50	67.03	76.79	68.27	78.09	62.77	77.06	65.73	75.14	64.08	77.88	60.37	78.37	67.17	75.55
Cusano LBPs [21]	77.13	81.80	78.57	82.69	77.68	81.39	75.34	80.15	76.51	82.35	78.98	82.14	76.65	81.18	78.64	81.94	77.75	80.70
Lee LBPs [22]	87.77	92.58	92.03	<b>93.89</b>	<b>92.79</b>	<b>94.57</b>	89.49	94.16	91.14	94.78	92.31	<b>94.92</b>	88.12	91.96	<b>91.96</b>	<b>93.54</b>	<b>92.65</b>	<b>94.23</b>
OBLBP [23]	94.02	<b>95.05</b>	<b>93.41</b>	<b>94.92</b>	<b>93.48</b>	<b>94.71</b>	94.99	<b>95.88</b>	94.51	<b>95.95</b>	94.51	<b>95.60</b>	93.75	94.51	<b>93.13</b>	<b>94.71</b>	<b>93.06</b>	<b>94.16</b>
IOBLBP [24]	92.93	<b>94.37</b>	91.83	<b>93.68</b>	91.21	<b>92.58</b>	94.09	94.92	93.82	<b>95.33</b>	92.79	<b>94.85</b>	92.51	93.48	91.35	<b>93.54</b>	90.73	<b>92.24</b>
MLBP (ours)	94.30		93.13		92.03		95.40		95.33		94.78		94.71		91.76		92.24	



**Fig. 8.** Classification accuracy (%) vs. computation cost (number of operations per pixel) of the different approaches (with D65 illuminant,  $d = 1$ , and WB demosaicing) for each MSFA: (a) Vis5 ( $K = 5$ ), (b) VisNIR8 ( $K = 8$ ), (c) IMEC16 ( $K = 16$ ), and (d) IMEC25 ( $K = 25$ ).

in *Procs. Visual Communications and Image Processing Conference (VCIP 2014)*, pp. 474-477 (Valletta, Malta, 2014).

- X. Wang, J.-B. Thomas, J. Y. Hardeberg, and P. Gouton, "Discrete wavelet transform based multispectral filter array demosaicking," in *Procs. 2013 Colour and Visual Computing Symposium (CVCS 2013)*, pp. 1-6 (Gjøvik, Norway, 2013).
- S. Mihoubi, O. Losson, B. Mathon, and L. Macaire, "Multispectral demosaicing using pseudo-panchromatic image," *IEEE Transactions on Computational Imaging* **3**(4), 982-995 (2017).
- Y. Monno, M. Tanaka, and M. Okutomi, "Multispectral demosaicking using guided filter," in *Procs. SPIE Electronic Imaging Annual Symposium: Digital Photography VII*, vol. 8299, p. 82990O (Burlingame, CA, USA, 2012).
- Y. Monno, D. Kiku, S. Kikuchi, M. Tanaka, and M. Okutomi, "Multispectral Demosaicking with Novel Guide Image Generation and Residual Interpolation," in *Procs. International Conference on Image Processing (ICIP 2014)*, pp. 645-649 (Paris, France, 2014).
- M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen, *Computer vision using local binary patterns*, vol. 40 of *Computational Imaging and Vision* (Springer-Verlag London, 2011).
- T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(7), 971-987 (2002).
- C. Palm, "Color texture classification by integrative Co-occurrence matrices," *Pattern Recognition* **37**(5), 965-976 (2004).
- A. Mirhashemi, "Introducing spectral moment features in analyzing the SpecTex hyperspectral texture database," *Machine Vision and Applications* **29**(3), 415-432 (2018). <http://www.uef.fi/web/spectral/spectex>.
- S. R. Dubey, S. K. Singh, and R. K. Singh, "Multichannel Decoded Local Binary Patterns for Content-Based Image Retrieval," *IEEE Transactions on Image Processing* **25**(9), 4018-4032 (2016).
- C. Cusano, P. Napoletano, and R. Schettini, "Combining local binary patterns and local color contrast for texture classification under varying illumination," *Journal of the Optical Society of America A* **31**(7), 1453-1461 (2014).
- S. H. Lee, J. Y. Choi, Y. M. Ro, and K. N. Plataniotis, "Local Color Vector Binary Patterns From Multichannel Face Images for Face Recognition," *IEEE Transactions on Image Processing* **21**(4), 2347-2353 (2012).
- T. Mäenpää, M. Pietikäinen, and J. Viertola, "Separating color and pattern information for color texture discrimination," in *Procs. 16th International Conference on Pattern Recognition (ICPR'02)*, vol. 1, pp. 668-671 (Québec, Canada, 2002).
- F. Bianconi, R. Bello-Cerezo, and P. Napoletano, "Improved opponent color local binary patterns: an effective local image descriptor for color texture classification," *Journal of Electronic Imaging* **27**(1), 011002 (2017).
- O. Losson and L. Macaire, "CFA local binary patterns for fast illuminant-invariant color texture classification," *Journal of Real-Time Image Processing* **10**(2), 387-401 (2015).
- H. A. Khan, S. Mihoubi, B. Mathon, J.-B. Thomas, and J. Y. Hardeberg, "HyTexiLa: High Resolution Visible and Near Infrared Hyperspectral Texture Images," *Sensors* **18**(7), 2045 (2018). <http://color.univ-lille.fr/datasets/hytextila>.
- 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* **24**(10), 3048-3059 (2015).
- P.-J. Lapray, J.-B. Thomas, P. Gouton, and Y. Ruichek, "Energy balance in Spectral Filter Array camera design," *Journal of the European Optical Society-Rapid Publications* **13**(1), 1 (2017).
- M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision* **7**(1), 11-32 (1991).