

# Motifs locaux binaires pour la classification d’images de textures multispectrales

Sofiane MIHOUBI, Olivier LOSSON\*, Benjamin MATHON, Ludovic MACAIRE

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

\*olivier.losson@univ-lille.fr

**Résumé** – Les motifs locaux binaires (LBP) permettent d’extraire des descripteurs de texture spatiaux pour discriminer les images de textures en niveaux de gris. Dans cet article, nous proposons un opérateur LBP qui extrait conjointement les informations de texture spatiales et spectrales à partir de l’image brute fournie par une caméra équipée d’une matrice de filtres multispectraux. Nous montrons expérimentalement que le descripteur proposé présente à la fois un faible coût de calcul et un haut pouvoir discriminant par rapport aux descripteurs LBP classiques appliqués aux images multispectrales.

**Abstract** – To discriminate gray-level texture images, spatial texture descriptors can be extracted using the local binary pattern (LBP) operator. In this paper we design an LBP operator that jointly extracts the spatial and spectral texture information directly from a raw image provided by a camera equipped with a multispectral filter array. Extensive experiments on a large dataset show that the proposed descriptor has both low computation cost and high discriminative power with regard to classical LBP descriptors applied to multispectral images.

## 1 Introduction

Une image multispectrale est composée de canaux représentant la radiance de la scène dans les différentes bandes spectrales auxquelles la caméra est sensible. Une telle image peut être obtenue à partir d’une seule acquisition à l’aide de la technologie « snapshot » [4]. La plupart des dispositifs de ce type utilisent une mosaïque de filtres multispectraux (MSFA) qui échantillonne spectralement la lumière incidente, sur le même principe que la mosaïque de Bayer très utilisée en imagerie couleur. Un MSFA est défini par un motif périodique de base dans lequel chaque filtre est sensible à une bande spectrale étroite. Chaque pixel de l’image brute ainsi acquise est associé à une seule bande selon le MSFA. L’image multispectrale entièrement définie, pour laquelle chaque pixel est associé à l’ensemble des bandes auxquelles la caméra est sensible, doit ensuite être estimée par dématricage.

Dans cet article, nous nous concentrons sur les descripteurs de texture basés sur les motifs locaux binaires (LBP). Les LBP ont été initialement calculés sur des images en niveaux de gris pour la classification de textures car l’opérateur originel utilise uniquement l’information spatiale de texture [12]. Les progrès récents en imagerie multispectrale nous amènent à étendre ces descripteurs de texture au domaine multispectral (voir Sec. 2).

Pour classifier des images de textures multispectrales acquises par une caméra équipée d’un MSFA, l’approche classique consiste à dématricer les images brutes, à extraire les caractéristiques de texture des images estimées, puis à comparer ces caractéristiques en utilisant une mesure de simila-

rité. Cependant, en plus d’être gourmand en temps de calcul, le dématricage est souvent susceptible de générer des artefacts d’estimation spatio-spectraux. Nous proposons donc de l’éviter en concevant un descripteur de textures calculé directement à partir de l’image brute (voir Sec. 3). Nous comparons enfin notre approche aux approches classiques sur un ensemble d’images de textures multispectrales (voir Sec. 4). Cet article résume notre proposition parue dans la revue JOSA A [9].

## 2 LBP sur une image multispectrale

### 2.1 Motifs MSFA et dématricage

Considérons les quatre motifs de base MSFA présentés sur la Fig. 1, qui sont liés à des travaux de recherche publiés (Vis5, VisNIR8), voire disponibles dans des caméras du marché (IMEC16, IMEC25).

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

(a) Vis5 [11] (b) VisNIR8 [13] (c) IMEC16 [3] (d) IMEC25 [3]

FIGURE 1 – Motifs de base de différents MSFA. Les numéros/lettres sont les indices/domaines des bandes.

Monno *et al.* [11] proposent le motif  $4 \times 4$  de la Fig. 1a qui

échantillonne cinq bandes dans le domaine du visible (Vis5). La bande dominante  $G$  caractérise la moitié des pixels et les quatre autres bandes caractérisent l'autre moitié. Thomas *et al.* [13] proposent le motif de la Fig. 1b qui échantillonne huit bandes dans les domaines du visible et du proche infrarouge (VisNIR8). La société IMEC commercialise des caméras dotées de MSFA basés sur un motif  $4 \times 4$  (cf. Fig. 1c) échantillonnant seize bandes dans le domaine du visible (IMEC16) et un motif  $5 \times 5$  (cf. Fig. 1d) échantillonnant 25 bandes dans le visible et le proche infrarouge (IMEC25).

Dans l'image brute acquise via un MSFA échantillonnant  $K$  bandes, la valeur de chaque pixel est associé à l'unique bande qui l'échantillonne. Une procédure de dématricage doit être appliquée à cette image brute afin d'obtenir une image multispectrale avec  $K$  canaux entièrement définis, parmi lesquels  $K-1$  sont estimés en chaque pixel. Dans nos expériences, nous considérons uniquement la méthode de dématricage basée sur les différences par rapport à l'image pseudo-panchromatique (PPID), qui est applicable à tous les motifs MSFA considérés et fournit les meilleurs performances [8].

## 2.2 LPBs appliqués à une image multispectrale

Soit  $\mathbf{I} = \{I^k\}_{k=1}^K$  une image multispectrale dématricée à  $K$  canaux. En un pixel donné  $p$ , nous considérons le voisinage  $N_p$  défini par son support  $\mathcal{N}^{P,d}$  constitué des  $P$  pixels situés à distance  $d$  de  $p$ . L'opérateur LBP marginal est défini pour chaque (indice de) bande  $k \in \{1, \dots, K\}$  en chaque pixel  $p$  par :

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

où  $I_p^k$  est la valeur du canal  $I^k$  en  $p$ ,  $\epsilon(q) \in \{0, \dots, P-1\}$  est l'indice de chaque pixel voisin  $q$  dans  $N_p$ , et  $s(\cdot)$  est la fonction échelon unité définie par :

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

Chaque canal  $I^k$  est caractérisé par l'histogramme des valeurs de LBP de ses pixels. L'image de texture multispectrale  $\mathbf{I}$  est alors décrite par la concaténation des  $K$  histogrammes des valeurs  $\{\text{LBP}^k[\mathbf{I}]\}_{k=1}^K$ . Ce descripteur représente l'interaction spatiale entre pixels voisins dans chaque canal indépendamment.

Plusieurs approches basées sur les LBP ont été proposées pour combiner l'information spatiale et l'information spectrale des images couleur. Parmi les méthodes extensibles au cas multispectral, Mirhashemi [10] applique l'opérateur LBP originel sur différents moments des signatures spectrales des pixels. Dubey *et al.* [2] proposent un opérateur spécifique (maLBP) utilisant les informations spectrales dans le schéma de codage du LBP. D'autres auteurs combinent l'histogramme du LBP appliqué à l'image pseudo-panchromatique portant l'information spatiale avec un histogramme basé sur le contenu spectral. Ce dernier peut être déduit, comme suggéré par Cusano [1], de l'angle entre la valeur d'un pixel et la valeur moyenne de ses

voisins dans le domaine spectral ou, comme suggéré par Lee [6], de l'angle entre l'axe d'une bande et la projection spectrale du pixel d'analyse sur le plan associé à chaque couple de bandes. Pour prendre pleinement en compte la corrélation spectrale, Mäenpää *et al.* [7] proposent le LBP des bandes opposées (OBLBP) appliqué à chaque couple de canaux  $(I^k, I^l)$ ,  $(k, l) \in \{1, \dots, K\}^2$ , d'une image multispectrale :

$$\text{OBLBP}^{k,l}[\mathbf{I}](p) = \sum_{q \in N_p} s(I_q^l, I_p^k) \cdot 2^{\epsilon(q)} \quad (3)$$

## 3 LBP sur une image brute

### 3.1 Voisinage dans l'image brute

Afin de correspondre aux spécifications des MSFA, nous fixons  $P = 8$  et considérons les trois supports  $\mathcal{N}^{8,d}$  avec une distance  $d \in \{1, 2, 3\}$  liée à la norme  $\mathcal{L}_\infty$ , comme illustré sur la Fig. 2a pour le MSFA VisNIR8.

Un MSFA associe une seule bande spectrale à chaque pixel. Il peut être défini sur l'ensemble  $S$  des pixels par la fonction MSFA :  $S \rightarrow \{1, \dots, K\}$ . Soit  $S^k = \{p \in S, \text{MSFA}(p) = k\}$  le sous-ensemble de pixels où le MSFA échantillonne la bande  $k$ , de sorte que  $S = \bigcup_{k=1}^K S^k$ . Pour l'ensemble des pixels  $p$  appartenant à  $S^k$ , on suppose que le voisinage  $N_p$  est toujours composé des mêmes bandes aux mêmes positions relatives par rapport à  $p$  dans le MSFA (cf. Fig. 2b pour le cas de VisNIR8).

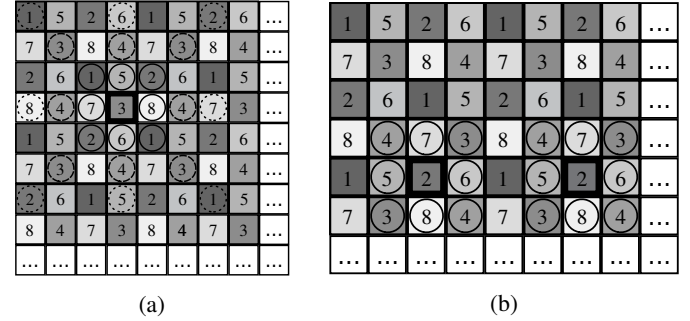


FIGURE 2 – Exemples de voisinages pour le MSFA VisNIR8. (a) : voisinage  $N_p$  d'un pixel  $p$  (carré noir), en considérant le support  $\mathcal{N}^{8,1}$  (cercles pleins),  $\mathcal{N}^{8,2}$  (tirets) et  $\mathcal{N}^{8,3}$  (pointillés). (b) : voisinage  $N_p$  de support  $\mathcal{N}^{8,1}$  pour deux pixels de  $S^2$  (carrés noirs).

Soit  $B^k$  l'ensemble des bandes disponibles dans le voisinage  $N_p$ ,  $p \in S^k$ , défini par le support  $\mathcal{N}^{8,d}$ . Comme on peut le voir sur la Fig. 2a, le nombre de bandes  $|B^k|$  varie en fonction de  $d$ . En effet, le voisinage de  $p \in S^3$  contient six bandes différentes pour  $d = 1$  et  $d = 3$ , mais seulement les bandes 3 et 4 pour  $d = 2$ .

### 3.2 LBP appliqué sur une image brute

Soit  $\underline{I}$  l'image brute fournie par une caméra équipée d'un MSFA. Afin de calculer notre descripteur de texture, nous ap-

pliquons d'abord l'opérateur LBP défini par l'Éq. (1) directement à l'image  $\underline{I}$  considérée comme une image à niveaux de gris :

$$\text{MLBP}[\underline{I}](p) = \sum_{q \in N_p} s(I_q, I_p) \cdot 2^{\epsilon(q)} \quad (4)$$

L'opérateur LBP est ici renommé MSFA LBP (MLBP) afin d'indiquer clairement la différence essentielle introduite par son application à  $\underline{I}$  et sa dépendance vis-à-vis du MSFA. Contrairement à l'Éq. (1), (4) combine les informations spectrales de  $B^{\text{MSFA}(p)}$ , c'est-à-dire des différentes bandes associées aux voisins de  $p$ . Puisque cet ensemble de bandes dépend de la bande MSFA( $p$ ) associée à  $p$ , nous calculons  $K$  histogrammes  $h^k$  en considérant chaque sous-ensemble de pixels  $S^k$  séparément :

$$\begin{aligned} h^k [\text{MLBP}[\underline{I}]] : [0, 2^P - 1] &\rightarrow \{0, \dots, |S^k|\} \\ j &\mapsto |\{p \in S^k, \text{MLBP}[\underline{I}](p) = j\}| \end{aligned} \quad (5)$$

Notons que seuls les pixels de  $S^k$  sont considérés pour calculer le  $k$ -ième histogramme. Le descripteur MLBP est défini par la concaténation des  $K$  histogrammes.

Le cas particulier de Vis5 contredit notre hypothèse selon laquelle un voisin de  $p \in S^k$  est toujours associé à la même bande pour une position relative donnée. En effet, pour deux pixels associés à la bande  $G$ , les voisins verticaux peuvent être associés à  $R$  et  $Or$  ou à  $B$  et  $Cy$ . Pour valider notre hypothèse sur Vis5, nous proposons de diviser  $S^G$  en quatre sous-ensembles de pixels  $\{S^{G_i}\}_{i=1}^4$ , dont les valeurs de MLBP seront finalement combinées en un seul histogramme.

### 3.3 Relation entre MLBP et OBLBP

Le MLBP défini par l'Éq. (4) est lié à OBLBP (cf. Éq. (3)). En effet, sa sortie peut être considérée comme la somme directe des sorties du même opérateur appliqué de manière restreinte à chaque sous-ensemble de pixels  $S^k$  :

$$\text{im} \{ \text{MLBP}[\underline{I}] \} = \bigoplus_{k=1}^K \text{im} \{ \text{MLBP}|_{S^k}[\underline{I}] \} \quad (6)$$

où  $\text{im}\{\cdot\}$  est l'image d'une application. Selon la définition de  $S^k$ , nous avons  $\underline{I}_p = I_p^k$  pour chaque pixel  $p \in S^k$ . À partir de l'Éq. (4), nous pouvons exprimer  $\text{MLBP}|_{S^k}$  à partir de  $\{I^k\}_{k=1}^K$  :

$$\text{MLBP}|_{S^k}[\underline{I}](p) = \sum_{l \in B^k} \sum_{q \in N_p \cap S^l} s(I_q^l, I_p^k) \cdot 2^{\epsilon(q)} \quad (7)$$

Par conséquent, MLBP est lié à OBLBP puisque les deux opérateurs prennent en compte les bandes opposées. Mais contrairement à OBLBP qui considère toutes les bandes  $l \in \{1, \dots, K\}$  pour tous les voisins de  $p$ , chaque code MLBP combine les informations de la bande  $k$  disponible en  $p$  et des  $|B^k|$  bandes disponibles aux voisins de  $p \in S^k$ .  $|B^k|$  reflète alors dans quelle mesure MLBP prend en compte la corrélation spectrale dans le voisinage MSFA.

## 4 Résultats expérimentaux

### 4.1 Base de textures et schéma de classification

Pour effectuer une classification expérimentale de textures, nous utilisons la base de données HyTexiLa [5] qui contient 112 images de réflectance de textures dont six sont représentées sur la Fig. 3. Nous simulons chaque image multispectrale de radiance pleinement définie à partir de la réflectance de chaque texture en utilisant l'illuminant D65 et la fonction de sensibilité spectrale relative associée à chaque bande des différents MSFA. Enfin, nous échantillons les images ainsi obtenues selon les différents MSFA afin de simuler les images brutes qui seraient acquises par les caméras associées. Pour chacune des quatre caméras considérées, ces simulations fournissent 112 images brutes de taille  $1024 \times 1024$  pixels. Chaque image est ensuite divisée en 25 sous-images de taille  $204 \times 204$  pixels, parmi lesquelles douze sont choisies au hasard pour l'apprentissage et les treize autres sont retenues pour les tests. Afin de déterminer le descripteur de texture le plus discriminant, nous retenons le classifieur « plus proche voisin » associé à la mesure de similarité basée sur l'intersection d'histogrammes car ce schéma de classification ne nécessite aucun paramètre.



FIGURE 3 – Échantillons d'images de la base de textures HyTexiLa. Pour une meilleure visualisation, les images multispectrales sont représentées en sRGB sous l'illuminant D65.

### 4.2 Taille et coût de calcul des descripteurs

La seconde colonne du Tab. 1 contient les tailles des descripteurs décrits dans la Sec. 2 et le nombre d'opérations élémentaires par pixel nécessaires à leur calcul. Notons que toutes les approches, à l'exception de la nôtre, nécessitent d'estimer une image multispectrale pleinement définie par dématricage, ce qui est aussi coûteux en temps de calcul. Notre approche MLBP nécessite 24 opérations par pixel. Le coût des approches marginale, maLBP et de Cusano augmente proportionnellement à  $K$ , et celui des approches de Lee et OBLBP

TABLE 1 – Taille du descripteur, nombre d’opérations et taux de bonne classification (%) des différentes approches.

Approche	Descripteur		Vis5 ( $K = 5$ )			VisNIR8 ( $K = 8$ )			IMEC16 ( $K = 16$ )			IMEC25 ( $K = 25$ )		
	taille ( $\times 2^P$ )	opérations	$d = 1$	$d = 2$	$d = 3$	$d = 1$	$d = 2$	$d = 3$	$d = 1$	$d = 2$	$d = 3$	$d = 1$	$d = 2$	$d = 3$
LBP marginal [12]	$K$	$24K$	87,23	88,26	88,60	88,94	88,67	88,53	84,89	86,95	87,16	78,71	80,01	79,81
LBP de moments [10]	3	$K^2+10K+69$	88,53	89,01	87,36	84,62	84,55	85,23	88,67	89,42	89,90	84,20	85,51	85,16
maLBP [2]	$K + 1$	$17K+8$	82,01	84,89	86,13	79,33	81,11	79,95	79,12	77,13	80,08	78,09	77,06	75,14
Cusano [1]	2	$15K+27$	87,36	88,39	89,22	85,51	87,09	87,57	86,20	87,91	87,02	80,15	82,35	82,14
Lee [6]	$1+K(K-1)$	$27K^2-26K+24$	93,13	<b>94,44</b>	95,47	91,41	91,69	93,41	95,26	<b>95,95</b>	95,81	94,16	94,78	<b>94,92</b>
OBLBP [7]	$K^2$	$24K^2$	<b>96,70</b>	<b>97,39</b>	<b>97,39</b>	<b>96,22</b>	<b>95,67</b>	<b>95,54</b>	<b>97,60</b>	<b>97,05</b>	96,98	<b>95,88</b>	<b>95,95</b>	<b>95,60</b>
MLBP (notre proposition)	$K$	24	94,92	92,45	96,63	94,57	94,71	94,64	97,32	94,92	97,05	95,40	95,33	94,78

à  $K^2$ . En tenant compte à la fois de la taille du descripteur et du coût de calcul, notre approche est la moins coûteuse tout en étant représentée par le même nombre d’histogrammes que les LBP marginaux.

### 4.3 Résultats de classification

Pour chaque combinaison de paramètres expérimentaux (MSFA et distance de voisinage  $d$ ), les taux de classification meilleurs que ceux de notre approche sont mis en évidence en gras dans le Tab. 1. Nous notons d’abord que  $d$  influe sur les performances de notre approche de façon différente selon le MSFA considéré. Cela est dû à la variation du nombre  $|B^k|$  de bandes disponibles en fonction de la distance  $d$  pour chaque MSFA. Ce nombre détermine en effet comment notre descripteur prend en compte la corrélation spectrale. Comparons maintenant les performances de notre descripteur à celles des autres. Le Tab. 1 montre que notre descripteur MLBP surpasse toujours les descripteurs de taille plus petite (LBP de moments et Cusano) ou équivalente (LBP marginal et maLBP). De plus, notre approche permet d’obtenir des résultats proches des descripteurs gourmands (Lee et OBLBP), en particulier avec IMEC16 et IMEC25. Le meilleur taux de bonne classification atteint par notre descripteur MLBP (tous MSFA et toutes distances confondues) est de 97,32 %, tandis que le meilleur descripteur (OBLBP) atteint 97,60 %.

## 5 Conclusion

Dans cet article, nous présentons un descripteur basé sur l’opérateur LBP conceptuellement simple et hautement discriminant pour les images brutes multispectrales. En plus de sa simplicité algorithmique, notre opérateur s’applique directement aux images brutes, ce qui lui permet d’éviter l’étape de dématricage et de réduire les coûts de calcul. Nous effectuons des expériences de classification de textures approfondies sur des images multispectrales simulées avec quatre MSFA référencés et très différents. Les résultats montrent que l’approche proposée surpasse les approches existantes utilisant des descripteurs de tailles similaires et fournit des résultats comparables à ceux des descripteurs de grande taille et de coût de calcul élevé.

## Références

- [1] C. Cusano, P. Napoletano et R. Schettini. *Combining local binary patterns and local color contrast for texture classification under varying illumination*. JOSA A, 31(7) :1453–1461, 2014.
- [2] S. R. Dubey, S. K. Singh et R. K. Singh. *Multichannel decoded local binary patterns for content-based image retrieval*. IEEE TIP, 25(9) :4018–4032, 2016.
- [3] B. Geelen, N. Tack et A. Lambrechts. *A compact snapshot multispectral imager with a monolithically integrated per-pixel filter mosaic*. Procs. SPIE, 8974 : 89740L, 2014.
- [4] N. Hagen et M. W. Kudenov. *Review of snapshot spectral imaging technologies*. Optical Engineering, 52(9) :090901, 2013.
- [5] H. A. Khan, S. Mihoubi, B. Mathon, J.-B. Thomas et 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>
- [6] S. H. Lee, J. Y. Choi, Y. M. Ro et K. N. Plataniotis. *Local color vector binary patterns from multichannel face images for face recognition*. IEEE TIP, 21(4) :2347–2353, 2012.
- [7] T. Mäenpää, M. Pietikäinen et J. Viertola. *Separating color and pattern information for color texture discrimination*. Procs. ICPR, 1, 668–671, 2002.
- [8] S. Mihoubi, O. Losson, B. Mathon et L. Macaire. *Multispectral demosaicing using pseudo-panchromatic image*. IEEE TCI, 3(4) :982–995, 2017.
- [9] S. Mihoubi, O. Losson, B. Mathon et L. Macaire. *Spatio-spectral binary patterns based on multispectral filter arrays for texture classification*. JOSA A, 35(9) :1532–1542, 2018.
- [10] A. Mirhashemi. *Introducing spectral moment features in analyzing the SpecTex hyperspectral texture database*. Machine Vision and Applications, 29(3) :415–432, 2018.
- [11] Y. Monno, M. Tanaka et M. Okutomi. *Multispectral demosaicing using adaptive kernel upsampling*. Procs. ICIP, 3157–3160, 2011.
- [12] T. Ojala, M. Pietikäinen et T. Mäenpää. *Multiresolution gray-scale and rotation invariant texture classification with local binary patterns*. IEEE TPAMI, 24(7) :971–987, 2002.
- [13] J.-B. Thomas, P.-J. Lapray, P. Gouton et C. Clerc. *Spectral characterization of a prototype SFA camera for joint visible and NIR acquisition*. Sensors, 16(7) :993, 2016.