Texture descriptor based on local combination adaptive ternary pattern

Faten Sandid *, Ali Douik

Research Laboratory of Automatic, Signal and Image Processing (LARATSI), National Engineering School of Monastir – ENIM, University of Monastir-Tunisia, Rue Ibn El Jazzar, Monastir 5000, Monastir, Tunisia
* E-mail: sandidfaten@hotmail.fr

Abstract: Material recognition has several applications, such as image retrieval, object recognition and robotic manipulation. To make the material classification more suitable for real-world applications, it is fundamental to satisfy two characteristics: robustness to scale and to pose variations. In this study, the authors propose a novel discriminant descriptor for texture classification based on a new operator called local combination adaptive ternary pattern (LCATP) descriptor used to encode both colour and local information. They start by building the LCATP descriptor using a combination of three different adaptive thresholding techniques. Moreover, they present a novel operator, mean histogram (MH), used jointly with the LCATP in order to incorporate colour information into the descriptor. This approach is then extended to four different colour spaces: LC1-C2, L/H2, LSHuv and O1-O2-O3. The final descriptor, LCATP fusion (LCATP_F), is produced by fusing the basic histogram (H) and MH extracted from the different colour spaces. Finally, the LCATP_F descriptor properties, such as the robustness to scale and pose changes are evaluated using the challenging KTH-textures under varying illumination, pose and scale (TIPS2b) dataset along with the least squares support vector machines classifier. The obtained experimental results, using the LCATP_F descriptor, show a significant improvement with respect to the state-of-the-art results.

1 Introduction

Material recognition has several applications, for instance, image retrieval, object recognition and robotic manipulation. It is often treated as a texture classification problem. To make the material classification more suitable for real-world applications, it is fundamental to verify the performance of the proposed approach towards different challenges in processing the real-world images, which bring out many difficulties. For instance, scale changes [1, 2], since material appearance varies significantly whether the image is presenting a zoom in (details are visible) or a zoom out (fine details may disappear), poses and illumination variation, depending on the point of view and illumination used to capture the material image [3–6], low quality etc. Recent researches have proven the efficiency of several local methods as texture descriptors such as the features based on local binary pattern (LBP) descriptors that have exhibited promising results [7, 8].

In this paper, we propose a simple yet discriminant and powerful approach for texture classification based on a new descriptor: local combination adaptive ternary pattern (LCATP) used to encode both colour and local structure information. This descriptor overcomes the low quality of an image by introducing a tolerant range, which makes it more robust and less sensitive to noise. In addition, the use of a combination of three different adaptive thresholding techniques leads to overcome the grey-scale changes present in the image, because of low quality, non-uniform regions, bad illumination and shadows etc., leading to a new operator more flexible to changes introduced by the image and even less sensitive to noise.

We start by building the LCATP descriptor using a combination of a three different local adaptive thresholding techniques. Then, a novel operator is presented, mean histogram (MH), that will be used jointly with the LCATP and the basic histogram (H) in order to incorporate colour information into the descriptor. This task can be achieved by computing the local means values and accumulating them into the MH bins. Since colour images exhibit more information than grey-scale images, and thus descriptors extracted from various colour spaces help to improve results, we extend our approach to four different colour spaces: LC1-C2, L/H2, LSHuv and O1-O2-O3. The final descriptor, LCATP fusion (LCATP_F), is produced by fusing the basic histogram (H) and MH extracted from the different colour spaces. Finally, the LCATP_F descriptor properties, such as the robustness to scale and pose changes are evaluated using the challenging KTH-textures under varying illumination, pose and scale (TIPS2b) dataset along with the least squares support vector machines classifier. The obtained experimental results, using the LCATP_F descriptor, show a significant improvement with respect to the state-of-the-art results.

2 Traditional LBP and its variants

2.1 Local binary pattern

LBPs, originally proposed by Ojala et al. [9] for texture analysis, are non-parametric descriptors that represent the local structure of an image efficiently by comparing each pixel with its neighbours [10]. It has proven to be a simple yet powerful operator to describe the local structures because of its various beneficial properties: robustness to illumination, computational simplicity and ability to encode texture details [10, 11].

Each pixel of the image is labelled with a decimal number, so-called ‘LBP code’: the central pixel (Pc) is compared with its neighbouring pixels (Pn) lying at a distance R from Pc, if the value of Pn is ≥ Pc, it will be coded with the value 1 otherwise it is set to 0. The binary numbers obtained are then multiplied by the corresponding values of a weight mask. The final decimal label of
the pixel $P_i$ is deducted by summing the different values computed, as given by (1) [10, 11]. Fig. 1 shows an example of an LBP code calculation

$$LBP(P_i) = \sum_{n=0}^{P_i-1} S(g_n - g_i)2^n$$

(1)

with $S$ as the sign function. $g_i$ and $g_n$ are, respectively, the grey values of $P_i$ and $P_n$.

Once the LBP’s codes of the whole image are calculated, the texture is represented by the histogram deduced from the LBP.

A rotation invariant LBP is obtained by choosing the smallest value from the $P-1$ bitwise shift operations applied on the binary pattern. This latter is also considered as uniform ($U(LBP_{P,R})$) if the number of transitions from 0 to 1, or conversely, in the binary number is $\leq 2$ (see [12] for more details). The LBP descriptor is then defined by (2)

$$LBP_{P,R}^{bin} = \begin{cases} \sum_{n=0}^{P_i-1} S(g_n - g_i), & \text{if } U(LBP_{P,R}) \\ P+1, & \text{otherwise} \end{cases}$$

(2)

Despite its advantages, the LBP has several drawbacks, inter alia, its sensitivity to noise [13]. Therefore many LBP variants descriptors were proposed in the past few years as local ternary pattern (LTP), local adaptive ternary pattern (LATP) etc.

### 2.2 Local ternary pattern

Tan and Triggs [13] suggest that the classical LBP tends to be sensitive to noise, mainly in quasi-uniform image areas as the used threshold is calculated based on the exact value of the centre pixel $P_c$. To overcome this problem, they proposed to use a new approach: LTPs [13]: the difference between $g_c$ and $g_n$ is then coded by three values, rather than two, using a threshold as given by (1) [10, 11]. Fig. 1 shows an example of an LBP code approach: LTPs [13]: the difference between

The resulting LTP is then split into two binary patterns: lower LTP ($c=−1$) and upper LTP ($c=1$) [5, 10, 13] using (4). Finally, the histograms computed from the two binary codes are concatenated to form the features vector

$$b_c(x) = \begin{cases} 1, & \text{if } x = c \\ 0, & \text{otherwise} \end{cases} \text{ with } c \in \{-1, 1\}$$

(4)

$$S = \begin{cases} 1, & \text{if } g_n \geq g_c + \tau \\ 0, & \text{if } g_c - \tau < g_n < g_c + \tau \\ -1, & \text{if } g_n \leq g_c - \tau \end{cases}$$

(3)

where $\mu_i$ is the local mean, $\sigma_i$ is the local standard deviation and $\kappa$ is a constant [14].

Akhlofi and Bendada use Niblack’s method [15] to compute the threshold value $\tau$. The local adaptive thresholding provides more robustness to illumination variations and less sensitivity to noise. There are various adaptive thresholding techniques. In the next section, four of the locally adaptive thresholding methods are discussed.

### 3 Local adaptive thresholding approaches

Unlike global thresholding techniques, which select a single threshold value to be applied for the entire image, the local adaptive thresholding methods allow the selection of different thresholds values specific to each pixel based on its neighbourhood statistics [16, 17]. Therefore they are more suitable for image thresholding. In this section, we present four different methods: Niblack, modified Niblack, Wolf and Yung.

#### 3.1 Niblack’s method

The threshold provided by Niblack’s method [15] is calculated pixel-wise using the local mean ($\mu_i$) and standard deviation ($\sigma_i$) computed over a $w \times w$ window around the central pixel. The local threshold for a given pixel is defined by (6)

$$T(i,j) = \mu_i(i,j) + k_N \sigma_i(i,j)$$

(6)

The parameter $k_N$ is used to control the impact of the standard deviation on the threshold value. It can be set to a negative or a positive value depending on the quality of the image [15, 16]. The value of $k_N$ is fixed to $-0.2$ by Niblack.

#### 3.2 Modified Niblack’s method

Unlike Niblack’s method, this approach incorporates both local and global characteristics of the image to deduce the threshold value. In addition, the parameter $k_N$ is no more a fixed value, but an adaptive one specific to each pixel depending on its local characteristics ($k_{MN}$) [18]. The weight $k_{MN}$ is computed using (7)

$$k_{MN} = k' \frac{\mu_g \cdot \sigma_g - \mu_i \cdot \sigma_i}{\max(\mu_g \cdot \sigma_g, \mu_i \cdot \sigma_i)}$$

(7)

where $\mu_i$ and $\sigma_i$ are, respectively, the global mean and standard deviation of the whole image, $\mu_g$ and $\sigma_g$ are, respectively, the local mean and standard deviation computed over a $w \times w$ window and $k' = -0.3$. 

---

**Fig. 1** LBP coding example
3.3 Wolf’s method

Wolf and Jolion [19] propose an adaptive thresholding algorithm designed to enhance the local contrast by normalising the various elements used to compute the threshold value for Niblack’s algorithm [18, 19]. The binarisation decision is then based on the contrast instead of the grey values of the different pixels. The threshold expression is given by (8)

\[ T = (1 - k_u) \times \mu + k_u M + \frac{\sigma}{R} (\mu - M_u) \]  

where \( R \) is the maximum of the local standard deviations of the whole image. We denote by \( M_u \) the minimum value of the grey level of the entire image. The parameter \( k_u \) is used in order to control the uncertainty around the mean value. \( k_u \) is fixed to 0.5 [19].

3.4 Yung’s method

Chiu et al. [20] make use of the local information by incorporating the local mean and the standard deviation computed from the gradient magnitude \( G \) [20] into the threshold formula given by (9)

\[ T = \mu + k_u \frac{\sigma}{l_G} (\mu - M_u) \]  

where \( M_u = \max(\mu_{Gx}, \mu_{Gy}) \), \( \mu_c \) denotes the local mean of the grey-level image and \( \mu_{Gx} \) is the local mean of the gradient magnitude. The parameter \( k_u \) is used to minimise the influence of the information deduced from the gradient for computing the thresholding value.

4 Proposed approach

We start this section by reviewing the four colour spaces in which we will define the new descriptor. Then, the LCATP is introduced.

4.1 Colour spaces

The colour images to be processed are usually represented in the red, green and blue (RGB) colour space. However, this colour representation has several drawbacks: the colour components are highly correlated, lack of human interpretation etc. [21]. Therefore different colour spaces were tested and we retained the most representative ones, that is, \( L_C C_2, L_I I_2, LSH_{uv} \) and \( O_1 O_2 O_3 \).

The transformation to the \( LSH_{uv} \) colour space from the \( (L^* u^* v^*) \) is defined by (11) [22].

\[ L^* = \begin{cases} \frac{29}{3} Y \sqrt{Y_p}, & \text{if } \frac{Y}{Y_p} \geq \left( \frac{6}{29} \right)^3 \\ \frac{116 Y^{1/3}}{Y_p} - 16, & \text{otherwise} \end{cases} \]

where \( u^* = 13L^* (u - u'_n) \)

\[ v^* = 13L^* (v - v'_n) \]

\[ X_n, Y_n \text{ and } Z_n \] are tristimulus values for a specified reference white point and \( X, Y \) and \( Z \) are defined by the linear transformation given by (12) [23].

\[ \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4306 & 0.3415 & 0.1784 \\ 0.222 & 0.7067 & 0.0713 \\ 0.0202 & 0.1295 & 0.9394 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \]

\[ \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} (R + G + B) \\ \frac{1}{2} (R - G) \\ \frac{1}{2} (2G - R - B) \end{bmatrix} \]

The \( I_1 I_2 I_3 \) colour space provides a stabilisation of the RGB one, achieved by a decorrelation of the RGB components by applying the linear transformation of (13) [22].

\[ \begin{bmatrix} L \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.5 & 0.5 & -1 \\ 0.866 & -0.866 & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \]

The study of the final colour space was motivated by the good correlation that it presents with the human visual system [25].

\[ L^* u^* v^* \] is defined by (11) [22]

\[ F. G. \]  

\[ (\frac{29}{3})^3 \]  

\[ \frac{116 Y^{1/3}}{Y_p} - 16, \] otherwise

\[ \text{if } \frac{Y}{Y_p} \geq \left( \frac{6}{29} \right)^3 \]

\[ \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.4306 & 0.3415 & 0.1784 \\ 0.222 & 0.7067 & 0.0713 \\ 0.0202 & 0.1295 & 0.9394 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \]

\[ \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} (R + G + B) \\ \frac{1}{2} (R - G) \\ \frac{1}{2} (2G - R - B) \end{bmatrix} \]

\[ \begin{bmatrix} L \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.5 & 0.5 & -1 \\ 0.866 & -0.866 & 0 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \]

Fig. 2 Sample image from the database and its colour components in \( L_C C_2, I_1 I_2 I_3, LSH_{uv} \) and \( O_1 O_2 O_3 \)
Equations (15) define the RGB to $O_1O_2O_3$ transformation [26]

$$
\begin{align*}
O_1 &= (R - G)/\sqrt{2} \\
O_2 &= ((R + G) - 2B)/\sqrt{6} \\
O_3 &= (R + G + B)/\sqrt{3}
\end{align*}
$$

(15)

Fig. 2 illustrates an example of an image coded into different colour spaces.

4.2 LCATP descriptor

As detailed in Section 2.3, to improve the LTP, the threshold is no longer a fixed value, but an adaptive one specific to each pixel [14]. Although the adaptive thresholding methods are more suitable for images with low quality, none of them is suitable in all cases [17, 27]. The performance of an approach differs from an image to another and so the resulting binarised image. Therefore combining various methods may improve the performance of the final results [17, 27].

Our approach incorporates the LTP with three different methods of local adaptive thresholding: modified Niblack ($T_N$) [18], Wolf ($T_W$) [19] and Yung ($T_Y$) [20] to create the LCATP descriptor. The $S$ function, given by (3), is then replaced by (16)

$$
S = \begin{cases} 
1, & \text{if } g_n \geq T_{N_1}, g_n \geq T_{N_2}, g_n \geq T_{Y_1} \\
-1, & \text{if } g_n \leq T_{N_1}, g_n \leq T_{N_2}, g_n \leq T_{Y_2} \\
0, & \text{otherwise}
\end{cases}
$$

(16)

The thresholds $T_{N_1}$, $T_{N_2}$ and $T_{Y_j}$, $j \in \{1, 2\}$, are defined, respectively, by (17)–(19), where $\mu_i$ and $\sigma_i$ are no longer defined on a $w \times w$ window, but their local values are calculated with the same neighbourhood topology used to calculate LCATP: a circle of radius $R$ and the number of neighbours $P_o$

$$
T_{N_1} = \mu_1 + k_{MN_1}\sigma_1
$$

(17)

$$
T_{N_2} = \mu_1 - k_{MN_1}\sigma_1
$$

(18)

$$
T_{Y_j} = \mu_j + k_{MN_2}(1 - \frac{\sigma_j}{R})(\mu_j - M_W)
$$

(19)

$k$, $k_n$ and $k_s$, defined, respectively, in (7)–(9) are user-defined parameters and sensitive to the contents of the images. Thus, to fix their values, a trial was conducted. We alternatively varied the value of one out of the three parameters in the interval $[0, 1]$ with a step of 0.05, as we verify the classification results, whereas the other two are fixed. Since the best results could be achieved when we set $k_s = 0.4$, $k_n = 0.15$ and computing $k_{MN}$ using $k' = 0.5$, we adopt the same settings in the following experiments.

The ternary pattern obtained is then split into two binary patterns according to (4): upper LCATP (ULCATP) and lower LCATP (LLCATP). Fig. 3 gives an example of these two patterns calculated on an image represented in the $O_1O_2O_3$ colour space, for two different $R(R = 1$ and 2). We then extract two histograms (basic histogram (H) and MH) from the ULCATP and LLCATP. Since H does not incorporate any information on the grey-level values of the image, it only helps encoding the texture information [6, 11], a new operator (MH) is then implemented to incorporate the colour information into the final descriptor: the local means of the grey-level values over the image are computed and then accumulated into the histogram bins according to (20). Finally, H and MH are concatenated to create the final descriptor

$$
MH(h) = \sum_{i=1}^{N} \sum_{j=1}^{M} l(\text{ULCATP}(i, j), h), h \in [0, K]
$$

(20)

where $l(\text{ULCATP}(i, j), h) = \begin{cases} 
\mu_i(i, j), & \text{if } \text{ULCATP}(i, j) = h \\
0, & \text{otherwise}
\end{cases}$

$K$ is the maximum value of LCATP patterns and $[M, N]$ are the image size.

4.3 Final fused descriptor LCATP_F

To calculate the proposed descriptor, we start by converting the RGB images into the four colour spaces: $L_C C_1$ ($L_{SHuv}$), $L_{SHuv}$ and $O_1O_2O_3$. The colour components of the produced images are then used to calculate the LCATP for $R = 1$ and 2, with, respectively, $P_o = 8$ and 16 (as detailed in Section 4.2) resulting into 24 different images, from which H and MH are computed. Next, the generated histograms are concatenated to form four different colour descriptors: $L_{C1}C_2$, $L_{SHuv}$, LSHuv and $O_1O_2O_3$. A fusion of these four vectors leads to the final descriptor LCATP_F as shown in (21). In the last step, we use the resulted descriptors along with the LS-SVM classifier [28, 29] in order to evaluate the performance of the resulted method for texture recognition. The different steps of our approach are depicted in Fig. 4

$$
\begin{align*}
\text{LCATP}_F &= [L_{C1}C_2, \text{LCATP} I_1, L_{SHuv}] \\
\text{LCATP}_F &= [L_{SHuv}, \text{LCATP} O_1O_2O_3] \\
\text{LCATP}_F &= [L_{SHuv}, \text{LCATP} O_1O_2O_3, \text{LCATP}]
\end{align*}
$$

(21)

5 Implementation details and results

In this section, we assess our proposed descriptor for material classification using the KTH-TIPS2b dataset. We start with a brief presentation of the dataset. Then, we discuss the different
experiments conducted in order to evaluate the performance of our approach for material categorisation.

5.1 Dataset

KTH-TIPS2b [30] is a challenging and publically accessible dataset [KTH-TIPS 2b: http://www.nada.kth.se/cvap/databases/kth-tips/] providing colour images with a total of nine different scales, three poses and four illumination conditions. It contains 11 different categories of materials, with four samples each, leading to a total number of 4752 images. The specificity of this database is that it contains samples that have a low-inter-classes dissimilarity (brown bread, white bread etc.) and low-intra-class similarity (cotton, linen etc.). Fig. 5 shows examples from the KTH-TIPS2b dataset.

5.2 Results

In this section, we introduce a detailed evaluation of the performance of the proposed method for material categorisation using the KTH-TIPS2b dataset, introduced in Section 5.1. The presentation of the experimental results is divided into three different steps with a corresponding objective to each one:

1. In experiment 1, presented in Section 5.2.1, we conduct a series of trials in order to test the influence of different parameters on the descriptors performance.
2. In experiment 2, presented in Section 5.2.2, an evaluation of the overall classification performance of LCATP, LCATP_R1, LCATP_R2, LSHav, LCATP, O1O2O3, LCATP and LCATP_F descriptors is conducted as well as an overview of class-wise results.
3. In experiment 3, presented in Section 5.2.3, we test the robustness of the proposed method to scale and pose invariance.

For both experiments 2 and 3, we provide a comparison with previous researchers’ results.

5.2.1 Parameters’ influence: We start the evaluation by testing the influence of each parameter of LCATP on the performance of each descriptor. For this purpose, four experimental setups are carried out; in each of them the descriptors were computed using different settings. Table 1 gives an overview of the parameters used for each vector. LCATP_R1 and LCATP_R2 descriptors were calculated using different configurations of (P_1, R_1)=(8, 1) and (P_2, R_2)=(16, 2). As for LCATP_H and LCATP_MH descriptors, the basic histogram (H) was used when computing the first vector, whereas the new operator MH was used to calculate the second descriptor.

In this trial, 72 images from each sample were used for training and the remaining 36 images for validation purpose.

The detailed experimental results are shown in Fig. 6. First, to evaluate the performance of each configuration, we start by a comparison of the results given by Figs. 6a and b. We note that using the LCATP_R2 with the configuration (P_2, R_2)=(16, 2) leads generally to better results than using the configuration (P_1, R_1)=(8, 1). Compared with the use of these two configurations alone, their combination provides an amelioration of the classification rates, as it can be noted in Fig. 7. As an example, the descriptors calculated in the \( O_1O_2O_3 \) colour space achieve a recognition rate of 92.487 and 96.023%, respectively, for (P_1, R_1) and (P_2, R_2) that leads generally to better results than using the configuration (P_1, R_1)=(8, 1). Compared with the use of these two configurations alone, their combination provides an amelioration of the classification rates, as it can be noted in Fig. 7. As an example, the descriptors calculated in the \( O_1O_2O_3 \) colour space achieve a recognition rate of 92.487 and 96.023%, respectively, for (P_1, R_1) and (P_2, R_2) compared with a classification rate of 97.538% obtained using the combination of both configurations in order to compute the LCATP.

Next, we evaluated the influence of the histogram type on the classification rates. The results of Figs. 6c and d demonstrate that the LCATP descriptor computed by using only the MH operator provides mainly better results than when using the basic histogram alone. It can also be noted that a combination of both types of

![Fig. 4 Overview on the stages of our approach](image-url)

<table>
<thead>
<tr>
<th>Classes</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>R=1</td>
<td>R=2</td>
</tr>
<tr>
<td>LCATP_R1</td>
<td>✓</td>
</tr>
<tr>
<td>LCATP_R2</td>
<td>✓</td>
</tr>
<tr>
<td>LCATP_H</td>
<td>✓</td>
</tr>
<tr>
<td>LCATP_MH</td>
<td>✓</td>
</tr>
</tbody>
</table>
Fig. 5 Example from each sample of the 11 classes materials
histograms improves the final performance of the descriptor as shown in the results of Fig. 7. Taking the case of the descriptor extracted from the \( O_1O_2O_3 \) colour space, we note a classification rate of 93.561 and 94.949% when using the original histogram and the MH histogram against a recognition rate of 97.538% when the LCATP descriptor is computed using a combination of \( H \) and \( MH \).

These results reinforce the effectiveness of the proposed descriptor.

5.2.2 Success rate evaluation: The second part of our experiments is conducted to evaluate the performance related to the fusion of the descriptors extracted from different colour spaces. As shown in Section 5.2.1., computing the LCATP descriptor using a combination of the different parameters: \((P_1, R_1), (P_2, R_2)\), \( H \) and \( MH \) lead to better results. Therefore this configuration is adopted to carry out the comparison between the results given by each descriptor and the recognition rate achieved by the fusion of the four of them. As for the training settings, we used the same protocol adopted in the first experiment. As seen in Fig. 7, the best success rate is given by the descriptor extracted from the \( I_1I_2I_3 \) colour space with 99.37%. The remaining colour spaces \( LC_1C_2 \), \( LSHuv \) and \( O_1O_2O_3 \) achieve a classification rate, respectively, of 98.9, 97.54 and 97.54%. It can be noted that the fusion of the four colour spaces descriptors, to form the final descriptor LCATP_F, outperforms the classification results of the individual ones with the rate of 99.81%, which prove the effectiveness of the final descriptor over the ones using one colour space.

In the second part of this experiment, we evaluate the performance of LCATP_F on each individual class of the dataset. Table 2 summarises the classification rates specific to each category for the colour descriptors and LCATP_F as well as a comparison with the results obtained by Banerji et al. [7]. As it can be noted, for LCATP_F, eight out of eleven categories achieve a success rate of 100% and the remaining three accomplish a classification rate of 99.3% which represents only one miss classification by class, therefore the total number of miss classification images are three out of 1584.

5.2.3 Invariance property to scale and pose: Since robustness to changes of scale and poses can be crucial for material categorisation, we conducted a set of trials to evaluate the performance of the proposed descriptor with different training sets specific to each objective.

Robustness to scale variation: To test the scale invariance property of the proposed method, we follow the training and validation scheme used in [1].

Table 2 Category wise descriptors success rates (%)

<table>
<thead>
<tr>
<th>Classes</th>
<th>( LC_1C_2 )</th>
<th>( h_{BC} )</th>
<th>( LSHuv )</th>
<th>( O_1O_2O_3 )</th>
<th>LCATP_F</th>
<th>Banerji [7]</th>
</tr>
</thead>
<tbody>
<tr>
<td>brown bread</td>
<td>100</td>
<td>100</td>
<td>99.306</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>white bread</td>
<td>100</td>
<td>100</td>
<td>99.306</td>
<td>97.917</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>cork</td>
<td>100</td>
<td>99.306</td>
<td>97.22</td>
<td>97.22</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>wood</td>
<td>99.306</td>
<td>99.306</td>
<td>98.611</td>
<td>97.22</td>
<td>99.306</td>
<td>100</td>
</tr>
<tr>
<td>aluminium foil</td>
<td>96.528</td>
<td>97.917</td>
<td>93.056</td>
<td>88.89</td>
<td>99.306</td>
<td>100</td>
</tr>
<tr>
<td>corduroy</td>
<td>97.22</td>
<td>98.611</td>
<td>88.19</td>
<td>97.917</td>
<td>99.306</td>
<td>98</td>
</tr>
<tr>
<td>cotton</td>
<td>99.306</td>
<td>99.306</td>
<td>100</td>
<td>99.306</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>lettuce leaf</td>
<td>97.917</td>
<td>98.611</td>
<td>97.22</td>
<td>97.22</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>wool</td>
<td>99.05</td>
<td>99.18</td>
<td>97.22</td>
<td>97.16</td>
<td>99.81</td>
<td>99.8</td>
</tr>
</tbody>
</table>
In each test, we alternatively use one out of the nine different scales for training while the remaining eight scales are reserved for validation. The results of this test are depicted in Table 3 along with the success rates of previous methods quoted directly from [1]. As it can be noted, a significant improvement in performance of our proposed method is achieved in comparison with previous researches’ results: the average classification rate reached by the LCATP_F is 82.3% (vs. 72.96 and 76.16%, respectively, for Harris-Laplace detector and SIFT descriptor (SI-SRC) [1]) which represent an average success rate of 97.66%.

Fig. 8 summarises the class-wise results of this trial. Going over the classification rates depicted in this figure, we note that materials, such as cotton, with a classification rate of 53.91%, show hardly any robustness to scale variations, whereas others, such as in the case of aluminium_foil present more relevant results with 99.05% proving robustness to changes of scale. This difference in performance from one class to another can be attributed to two main factors. The first related to intra-class properties: materials with a greatly regular pattern, such as cotton, present a clear characteristic scale, whereas other classes with irregular pattern, such as aluminium_foil, present similar characteristics over a range of scales. Thus, the characteristic vector of the former material could be severely changed from one scale to another, whereas the descriptor of the latter is expected to present more robustness scale variations.

Table 3 Success rates (%) obtained for the evaluation of scale robustness trial by selecting one scale for training and the remaining eight for test

<table>
<thead>
<tr>
<th>Training set</th>
<th>HL-SIFT</th>
<th>SI-SRC</th>
<th>LCATP_F</th>
</tr>
</thead>
<tbody>
<tr>
<td>scale 1</td>
<td>62.47</td>
<td>67.71</td>
<td>80.87</td>
</tr>
<tr>
<td>scale 2</td>
<td>71.96</td>
<td>73.96</td>
<td>85.89</td>
</tr>
<tr>
<td>scale 3</td>
<td>78.92</td>
<td>76.62</td>
<td>86.67</td>
</tr>
<tr>
<td>scale 4</td>
<td>83.21</td>
<td>81.34</td>
<td>89.18</td>
</tr>
<tr>
<td>scale 5</td>
<td>83.99</td>
<td>82.88</td>
<td>87.48</td>
</tr>
<tr>
<td>scale 6</td>
<td>80.37</td>
<td>82.03</td>
<td>84.71</td>
</tr>
<tr>
<td>scale 7</td>
<td>73.39</td>
<td>80.66</td>
<td>77.13</td>
</tr>
<tr>
<td>scale 8</td>
<td>64.6</td>
<td>73.27</td>
<td>75.05</td>
</tr>
<tr>
<td>scale 9</td>
<td>57.41</td>
<td>64.96</td>
<td>70.69</td>
</tr>
<tr>
<td>average</td>
<td>72.92</td>
<td>76.16</td>
<td>82.3</td>
</tr>
</tbody>
</table>

6 Conclusions

In this paper, we proposed a new texture descriptor (LCATP_F), based on the local patterns of the image with a fusion of three different local adaptive thresholding techniques: modified Niblack’s method, Wolf’s method and Yung’s method. We also introduced another new descriptor, MH, in order to improve the performance of the proposed method by incorporating the contrast information. A series of experiments have been conducted on the challenging KTH-TIPS2b database to evaluate the performance of our proposed method on material categorisation. The obtained results of the first set of trials show a significant improvement of the classification rate (99.81%) over the previous researches results. The recognition rates achieved on the second test (82.30%) prove the scale invariant property of the LCATP_F descriptor, with an improvement of more than 6% over the previous state-of-the-art results. Finally, the classification results of the last experiment (97.66%) prove the robustness of the proposed method to pose changes.

On the basis of the obtained results, we believe that the proposed approach could be applicable in different computer vision tasks involving scale and pose variation, such as object recognition in complex scene etc.

7 References


Table 4 Classification results (%) for robustness to poses’ changes

<table>
<thead>
<tr>
<th>Training set</th>
<th>LCATP_F</th>
</tr>
</thead>
<tbody>
<tr>
<td>pose 1</td>
<td>97.38</td>
</tr>
<tr>
<td>pose 2</td>
<td>98.35</td>
</tr>
<tr>
<td>pose 3</td>
<td>97.26</td>
</tr>
<tr>
<td>average</td>
<td>97.66</td>
</tr>
</tbody>
</table>

The second factor is related to the inter-class variation. Indeed, the classification performance depends on the degree of distraction caused by other classes. As shown in Fig. 9, it is possible that an image of a material, taken at a certain scale, closely resembles another material of a different scale, which leads to a confusion between classes.

Robustness to poses changes: As mentioned in Section 5.1, the images of the KTH-TIPS2b dataset were captured using three different points of view. To test the robustness of our proposed method to poses’ changes, a different training scheme is used: in each test, we alternatively selected one out of the three poses for training while the remaining two are dedicated to the validation. The results of this experiment are recorded in Table 4. As it can be seen in Table 4, the poses invariant property of the proposed descriptor can be proved by the experimental results with an average success rate of 97.66%.

Fig. 9 Example of similar materials of different scales

Fig. 8 Class-wise results obtained for the evaluation of scale robustness trial by selecting one scale for training and the remaining eight for test

<table>
<thead>
<tr>
<th>Training set</th>
<th>Average</th>
<th>CI_margin</th>
<th>HLT_sift</th>
<th>SI-sift</th>
<th>LCATP_F</th>
</tr>
</thead>
<tbody>
<tr>
<td>scale 1</td>
<td>78.96</td>
<td>64.70</td>
<td>79.91</td>
<td>98.78</td>
<td></td>
</tr>
<tr>
<td>scale 2</td>
<td>78.39</td>
<td>64.70</td>
<td>80.66</td>
<td>77.13</td>
<td></td>
</tr>
<tr>
<td>scale 3</td>
<td>64.6</td>
<td>73.27</td>
<td>75.05</td>
<td>70.69</td>
<td></td>
</tr>
<tr>
<td>scale 4</td>
<td>57.41</td>
<td>64.96</td>
<td>70.25</td>
<td>99.05</td>
<td></td>
</tr>
<tr>
<td>scale 5</td>
<td>62.47</td>
<td>67.71</td>
<td>80.87</td>
<td>97.66</td>
<td></td>
</tr>
<tr>
<td>scale 6</td>
<td>71.96</td>
<td>73.96</td>
<td>85.89</td>
<td>97.66</td>
<td></td>
</tr>
<tr>
<td>scale 7</td>
<td>78.92</td>
<td>76.62</td>
<td>86.67</td>
<td>97.66</td>
<td></td>
</tr>
<tr>
<td>scale 8</td>
<td>83.21</td>
<td>81.34</td>
<td>89.18</td>
<td>97.66</td>
<td></td>
</tr>
<tr>
<td>scale 9</td>
<td>83.99</td>
<td>82.88</td>
<td>87.48</td>
<td>97.66</td>
<td></td>
</tr>
<tr>
<td>scale 10</td>
<td>80.37</td>
<td>82.03</td>
<td>84.71</td>
<td>97.66</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pose 1</th>
<th>Pose 2</th>
<th>Pose 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>97.38</td>
<td>98.35</td>
<td>97.26</td>
</tr>
<tr>
<td>97.66</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Classification results (%) for robustness to poses’ changes

Training set | LCATP_F |
-------------|---------|
pose 1       | 97.38   |
pos 2        | 98.35   |
pos 3        | 97.26   |
average      | 97.66   |
26 Giang, B.T.: ‘Combining of text-based semantics and vision-based Semantics’. Master thesis, Faculty of Mathematics and Physics, Charles University in Prague and Faculty of Computer Science, Free University of Bozen Bolzano, 2011