Journal of Electronic Imaging

JElectronicImaging.org

Extended color local mapped pattern for color texture classification under varying illumination

Tamiris Trevisan Negri
Fang Zhou
Zoran Obradovic
Adilson Gonzaga

SPIE. IS&T imaging.org

Tamiris Trevisan Negri, Fang Zhou, Zoran Obradovic, Adilson Gonzaga, "Extended color local mapped pattern for color texture classification under varying illumination," *J. Electron. Imaging* **27**(1), 011008 (2018), doi: 10.1117/1.JEI.27.1.011008.

Extended color local mapped pattern for color texture classification under varying illumination

Tamiris Trevisan Negri,[a,b,*] Fang Zhou,[c] Zoran Obradovic,[c] and Adilson Gonzaga[a]
[a]University of Sao Paulo, Laboratory of Computer Vision, Department of Electrical and Computer Engineering, Sao Carlos, Brazil
[b]Federal Institute of Education, Science and Technology of Sao Paulo, Araraquara, Brazil
[c]Temple University, Center for Data Analytics and Biomedical Informatics, Philadelphia, Pennsylvania, United States

Abstract. This paper presents a color–texture descriptor based on the local mapped pattern approach for color–texture classification under different lighting conditions. The proposed descriptor, namely extended color local mapped pattern (ECLMP), considers the magnitude of the color vectors inside the RGB cube to extract color–texture information from the images. These features are combined with texture information from the luminance image in a multiresolution fashion to get the ECLMP feature vector. The robustness of the proposed method is evaluated using the RawFooT, KTH-TIPS-2b, and USPtex databases. The experimental results show that the proposed descriptor is more robust to changes in the illumination condition than 22 alternative commonly used descriptors. *© 2018 SPIE and IS&T* [DOI: 10.1117/1.JEI.27.1.011008]

Keywords: color texture; local mapped pattern; local descriptors; texture description; illumination; texture classification.
Paper 170852SS received Sep. 29, 2017; accepted for publication Jan. 10, 2018; published online Feb. 5, 2018.

1 Introduction

In recent studies, the importance of incorporating color information into texture description has been investigated extensively.[1–4] The findings demonstrate that combining color and texture information improves the accuracy of color–texture classification. This approach has been explored by a variety of descriptors.[5–12]

Color and texture information can be combined in different ways. Different taxonomies have been proposed to group the color–texture descriptors into classes.[4,13,14] Palm[4] classified the approaches into parallel, sequential, and integrative. The parallel approach measures texture by calculating the relationship of neighboring pixels intensities ignoring their color, and color is measured globally ignoring local neighborhood pixels. In the sequential concept, the color image is quantized to produce a single channel image, which is subsequently processed as a gray-scale texture. Integrative methods process color and texture information jointly, considering either a single channel or a multichannel approach.

Integrative methods have been widely adopted to combine color and texture features. Bianconi et al.[15] combined color and texture information by computing the ranklet transform intrachannel and interchannel. The local color contrast descriptor (LCC)[10] captures information from the image color contrast and associate them with the local binary patterns (LBP) histogram.[16] The opponent color local binary pattern (OCLBP)[3] and the improved opponent color local binary pattern (IOCLBP)[12] combine opponent color and monochrome features, both extracted from the image using the LBP approach, to obtain a color–texture descriptor. Opponent color features were also investigated together with Gabor wavelets for texture recognition.[17] Ledoux et al.[18] proposed the mixed color order LBPs, which uses the vector information of color.

Texture descriptors, designed originally to deal with gray-scale images, were also extended to be applied to color textures. For instance, Maenppa and Pietikainen[3] investigated the use of concatenated LBP histograms, obtained from each color channel, as a color–texture descriptor. The same strategy was performed using Gabor wavelets.[3] In both experiments, different color spaces were considered, such as RGB, Lab, HSV, and $I_1I_2I_3$.

One of the main challenges in color–texture description is the robustness to changes in the scene illumination. Variations in the light intensity, direction, and temperature can result in changing the color of the observed texture and further interfering the descriptor capability. One way of overcoming this issue is to apply preprocessing steps to normalize the texture color. However, these techniques require extra computational time and they only improve the texture description under some particular conditions.[19] Therefore, an ideal scenario to deal with varying illumination is to use color–texture descriptors that are robust to possible changes in the scene illumination.

In a previous work,[20] a color–texture descriptor, namely color intensity local mapped pattern (CILMP), which incorporates color information using the magnitude of the color vectors inside the RGB cube has been proposed. The method captures small transitions of the pixels in the image by comparing the information of a set of sampled pixels in a local circular neighborhood to the neighborhood central pixel. We demonstrate that concatenating color and texture information in a multiresolution fashion improves the descriptor performance.

In the current paper, we propose an improvement over the CILMP descriptor. The proposed color–texture descriptor, namely extended color local mapped pattern (ECLMP), incorporates color and texture information, from multiple texture resolutions, in two different ways. First, both color

*Address all correspondence to: Tamiris Trevisan Negri, E-mail: tamirisnegri@usp.br

and texture (color–texture) features are extracted from the RGB channels in an integrative approach by considering the relationship between the neighboring color pixels. Then, the color–texture information is combined with luminance texture information in a parallel approach. The contributions of the ECLMP descriptor include a neighborhood comparison approach, which considers the average of the neighborhood values, including the central pixel rather than the central pixel alone. Furthermore, the descriptor adopts the comparison between pixels located at circular neighborhoods with different radii, thus capturing subtle transitions among different resolutions.

The ECLMP was evaluated over three datasets consisting of images taken under controlled and uncontrolled variations of illumination conditions. The experimental results indicated that the approach is robust to changes in the light source, thus increasing the classification accuracy.

The paper is organized as follows: Sec. 2 presents the proposed descriptor, explaining how color–texture and texture information are combined. The classification set up and experimental results are described in Sec. 3. Finally, we conclude the paper in Sec. 4.

2 Extended Color Local Mapped Pattern

The ECLMP is a parametric LMP-based descriptor[21] designed to be robust to changes in the scene illumination. The proposed descriptor uses two operators, ECLMP_c and ECLMP_r, to extract color–texture and texture information from the image. The ECLMP_c explores an approach to extract information from neighboring pixels: the pixels in a circular neighborhood are compared with the neighborhood pixel average, including the central pixel. The ECLMP_r captures the relationship between pixels from different neighboring resolutions of an image. The texture information is extracted from the image luminance and the color–texture information is incorporated using the magnitude of the color vectors inside the RGB cube.

In this section, we first explain how the ECLMP incorporates texture information from the image luminance using the ECLMP_c and ECLMP_r operators. Then, we extend the method for color textures in the RGB space. Finally, we present how the ECLMP combines both texture and color–texture information in a multiresolution fashion.

2.1 *Texture Information Extraction*

Consider a 5×5 local pattern, taken from an image as shown in Fig. 1. The ECLMP extracts information from such local pattern by considering the luminance values located in concentric circular neighborhoods as shown in Fig. 2. Each neighborhood is defined by P pixels $g_{p,r}$ $(p = 1, \ldots, P)$ equally spaced in a circle of radius $r, r > 0$. Assuming the central pixel g_c is located at $(x, y) = (0,0)$, the coordinates of the neighboring pixels $g_{p,r}$ are given by Eq. (1). When the pixel coordinates are not integer values, the pixel gray levels are estimated through interpolation, as suggested in Ref. 16

$$(x, y) = [-r\sin(2\pi p/P), r\cos(2\pi p)/P], \quad p = 1, \ldots, P. \tag{1}$$

The ECLMP texture feature extraction is performed by two operators—ECLMP_c and ECLMP_r—applied to the luminance image.

The ECLMP_c operator considers the relationship between the local pixel average and circular neighborhood with larger radius r_{out} (Fig. 2). This information is captured by the sigmoid function defined as

$$f(g_{p,r_{\mathrm{out}}}) = \frac{1}{1 + \exp\left[\frac{-(g_{p,r_{\mathrm{out}}} - m_1)}{\beta_{c1}}\right]}, \tag{2}$$

where β_{c1} is the steepness of the curve and m_1 is the average of the gray level values of the pixels in the circular neighborhood, including the central pixel, as presented in

$$m_1 = \frac{1}{P+1}\left[\left(\sum_{p=1}^{P} g_{p,r_{\mathrm{out}}}\right) + g_c\right]. \tag{3}$$

Then, each local pattern is represented by a code generated using Eq. (4), where f is the mapping function defined in Eq. (2). The operator round(.) rounds a noninteger number to the nearest integer; in the case of a tie, the number is rounded to the nearest integer greater than it. The image codes are stored into a matrix named coding map. These codes are uniformly quantized in Q elements

$$\mathrm{ECLMP}_c = \mathrm{round}\left\{\frac{[\sum_{p=1}^{P} f(g_{p,r_{\mathrm{out}}})] + f(g_c)}{P+1}(Q-1)\right\}. \tag{4}$$

Inspired by Liu et al.,[22,23] the ECLMP_r operator captures the relationship between the pixels $g_{p,r_{\mathrm{in}}}$ located in the circular neighborhood of smaller radius r_{in} and the pixels $g_{p,r_{\mathrm{out}}}$ located in the circular neighborhood of larger radius r_{out}. Comparisons are made using the differences between the gray levels of $g_{p,r_{\mathrm{in}}}$ and $g_{p,r_{\mathrm{out}}}$ as the input of a sigmoid

g_1	g_2	g_3	g_4	g_5
g_6	g_7	g_8	g_9	g_{10}
g_{11}	g_{12}	$\boldsymbol{g_c}$	g_{13}	g_{14}
g_{15}	g_{16}	g_{17}	g_{18}	g_{19}
g_{20}	g_{21}	g_{22}	g_{23}	g_{24}

Fig. 1 Local pattern 5×5.

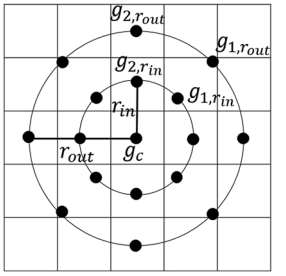

Fig. 2 Concentric circular neighborhoods.

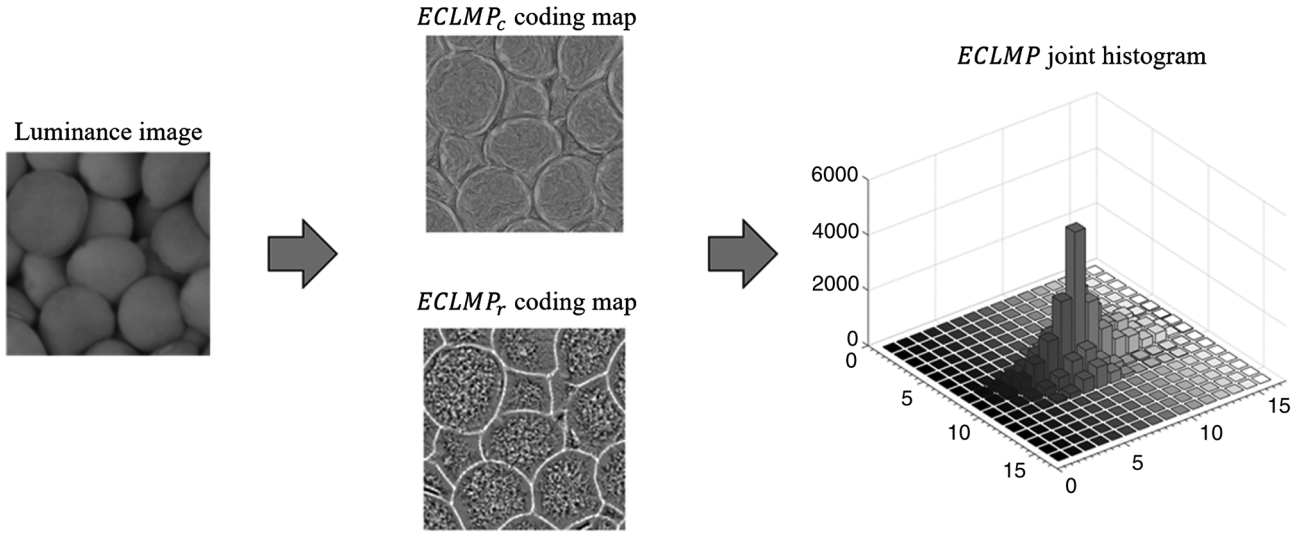

Fig. 3 ECLMP feature vector generation of a luminance image, considering $P = 8$, $r_{\text{out}} = 2$, $r_{\text{in}} = 1$, $Q = 16$, and $\beta_{c1} = \beta_{r1} = 0.1$.

function as defined in Eq. (5), where β_{r1} is the steepness of the curve

$$z(g_{p,r_{\text{out}}}, g_{p,r_{\text{in}}}) = \frac{1}{1 + \exp\left[\frac{-(g_{p,r_{\text{out}}} - g_{p,r_{\text{in}}})}{\beta_{r1}}\right]}. \tag{5}$$

The differences are then averaged, generating the ECLMP_r codes through Eq. (6), where z is the mapping function defined in Eq. (5). As in the ECLMP_c, the codes are uniformly quantized in Q elements and stored in a coding map

$$\text{ECLMP}_r = \text{round}\left[\frac{\sum_{p=1}^{P} z(g_{p,r_{\text{out}}}, g_{p,r_{\text{in}}})}{P}(Q-1)\right]. \tag{6}$$

The ECLMP feature vector is generated by combining the ECLMP_c and ECLMP_r information into a joint histogram,[16] which is built using both feature location and value in the coding maps, as shown in Fig. 3.

2.2 *Color — Texture Information Extraction*

The ECLMP_c and ECLMP_r operators can be easily extended for color textures using the approach presented in our previous work.[20] On the RGB color space, each color is represented as a point (or a vector) by its primary components red, green, and blue, respectively. So, a local $W \times W$ pattern of a color texture is a three-dimensional (3-D) matrix, where each pixel is considered as a vector $g_p = (R_p, G_p, B_p)$ in the RGB cube (Fig. 4).

Color–texture features are extracted from the RGB channels using the integrative concept: color and texture information are extracted jointly by considering the squared magnitudes of the vectors in the RGB space.

To incorporate the color–texture information, the ECLMP_c operator uses the sigmoid function as defined in Eq. (7), where β_{c2} is the steepness of the curve and m_2 is the average of the squared magnitudes of the pixels in the circular neighborhood including the central pixel [presented in Eq. (8)]. The ECLMP_c coding map is generated using Eq. (4), and the codes are uniformly quantized into Q elements

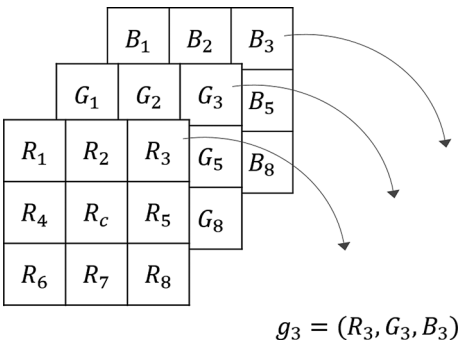

Fig. 4 Local pattern 3 × 3 of an RGB image.

$$f(g_{p,r_{\text{out}}}) = \frac{1}{1 + \exp\left[\frac{-(\|g_{p,r_{\text{out}}}\|^2 - m_2)}{\beta_{c2}}\right]}, \tag{7}$$

$$m_2 = \frac{1}{P+1}\left[\left(\sum_{p=1}^{P} \|g_{p,r_{\text{out}}}\|^2\right) + \|g_c\|^2\right]. \tag{8}$$

The ECLMP_r coding map of the color textures is also obtained using Eq. (6), following the same procedure as for the texture information extraction. However, the color–texture differences between the $g_{p,r_{\text{in}}}$ and $g_{p,r_{\text{out}}}$ pixels are captured by the sigmoid function defined in Eq. (9), where β_{r2} is the steepness of the curve

$$z(g_{p,r_{\text{out}}}, g_{p,r_{\text{in}}}) = \frac{1}{1 + \exp\left[\frac{-(\|g_{p,r_{\text{out}}}\|^2 - \|g_{p,r_{\text{in}}}\|^2)}{\beta_{r2}}\right]}, \tag{9}$$

Figure 5 shows ECLMP_c and ECLMP_r coding maps of a color texture and the corresponding ECLMP joint histogram. Note that to measure the differences between the pixels located on circular neighborhoods of different radius, the number of samples P should be the same. For example, considering the neighborhood presented in Fig. 2, we can

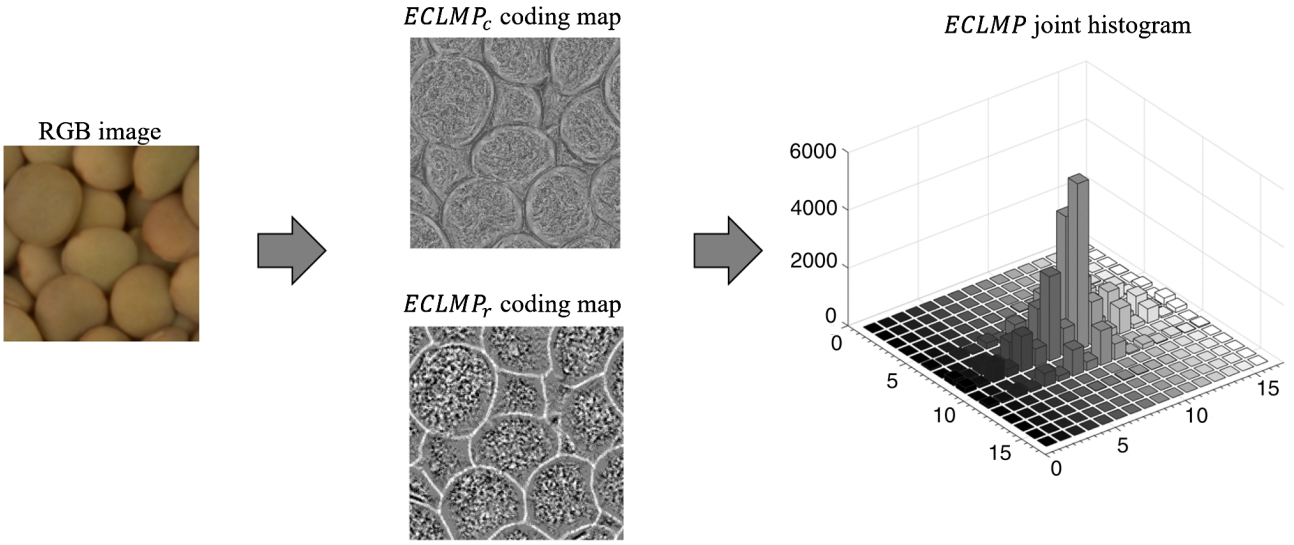

Fig. 5 ECLMP feature vector generation of an RGB image, considering $P = 8$, $r_{out} = 2$, $r_{in} = 1$, $Q = 16$, and $\beta = 0.1$.

set $r_{out} = 2$, $r_{in} = 1$, and $P = 8$ for both radii. To make the notation more understandable, we represent such configuration as $\text{ECLMP}(P, r_{out}, r_{in}) = \text{ECLMP}(8,2,1)$.

2.3 *Combining Texture and Color–Texture Information*

Taking advantage of the multiresolution analysis, we perform the feature extraction using different radii (r_{out}, r_{in}) and number of neighbors (P). Then, the features from different configurations of (P, r_{out}, r_{in}) are concatenated in a parallel approach to get the image feature vector.

In this work, the ECLMP descriptor uses two configurations, empirically tested and selected, for color–texture information extraction: $(P, r_{out}, r_{in}) = (8,2,1)$ and $(P, r_{out}, r_{in}) = (16,3,2)$; the same two configurations for texture information extraction, totalizing four configurations.

To get the image feature vector, the joint histograms are reshaped into one-dimensional (1-D) features and concatenated as shown in Fig. 6. The parameter Q was set to 16, so each joint histogram has $Q \times Q = 16 \times 16 = 256$ bins, and the image feature vector has $256 \times 4 = 1024$ features.

3 Experiments

3.1 *Datasets*

We evaluated the performance of the ECLMP over three databases: RawFooT (Raw Food Texture database),[19] KTH-TIPS-2b (textures under varying illumination, pose, and scale),[24] and USPtex dataset.[6]

The RawFooT database was specially designed to analyze the robustness of texture descriptors under several lighting conditions. The same evaluation protocol, as presented in Ref. 19, was used to allow the comparison of the results. The database contains 68 images of raw food textures, such as different types of fruits, vegetables, fish, cereals, and meat. Each image was acquired under 46 different lighting conditions, which differ in light intensity, direction, temperature, color, and combination of these settings, as shown in Fig. 7.

The original textures of 800×800 pixels were divided into 16 nonoverlapping samples of 200×200 pixels. For each class, eight samples were assigned for training and eight for test. So, the dataset has $68 \times (8 + 8) = 1088$ samples for each one of the 46 lighting conditions. Training and test images were grouped into subsets as suggested in Ref. 19 to perform the same nine classification tasks proposed by the authors.

Each task involves the samples taken under certain variation in the light source; for example, task 2 investigates the robustness of the descriptor to changes in the light intensity. So, we form 12 subsets by combining the samples acquired under four intensity levels (100%, 75%, 50%, and 25%), where the intensity of the training samples is different from the intensity of the test samples. The nine tasks are briefly described below and follow the terminology presented in Fig. 7.

1. *No variation:* 46 subsets, each one composed of training and test samples taken under the same light conditions.
2. *Light intensity*: 12 subsets, obtained by combining samples acquired under four intensity levels (D65, $I = 100\%$, 75%, 50%, and 25%). For each subset, training and test samples were taken under different light intensities.
3. *Light direction*: 72 subsets. Each subset is composed of training and test samples taken under different light directions (D65, $\theta = 24$ deg, 30 deg, 36 deg, 42 deg, 48 deg, 54 deg, 60 deg, 66 deg, and 90 deg).
4. *Daylight temperature*: 132 subsets, each one composed of training and test sets with images taken under different daylight temperatures (D40, D45, D50, ..., D95).
5. *LED temperature*: 30 subsets. For each subset, training and test sets were taken under different LED light temperatures (L27, L30, L40, L50, L57, and L65).
6. *Daylight versus LED*: 72 subsets obtained by combining the twelve daylight temperatures (D40,

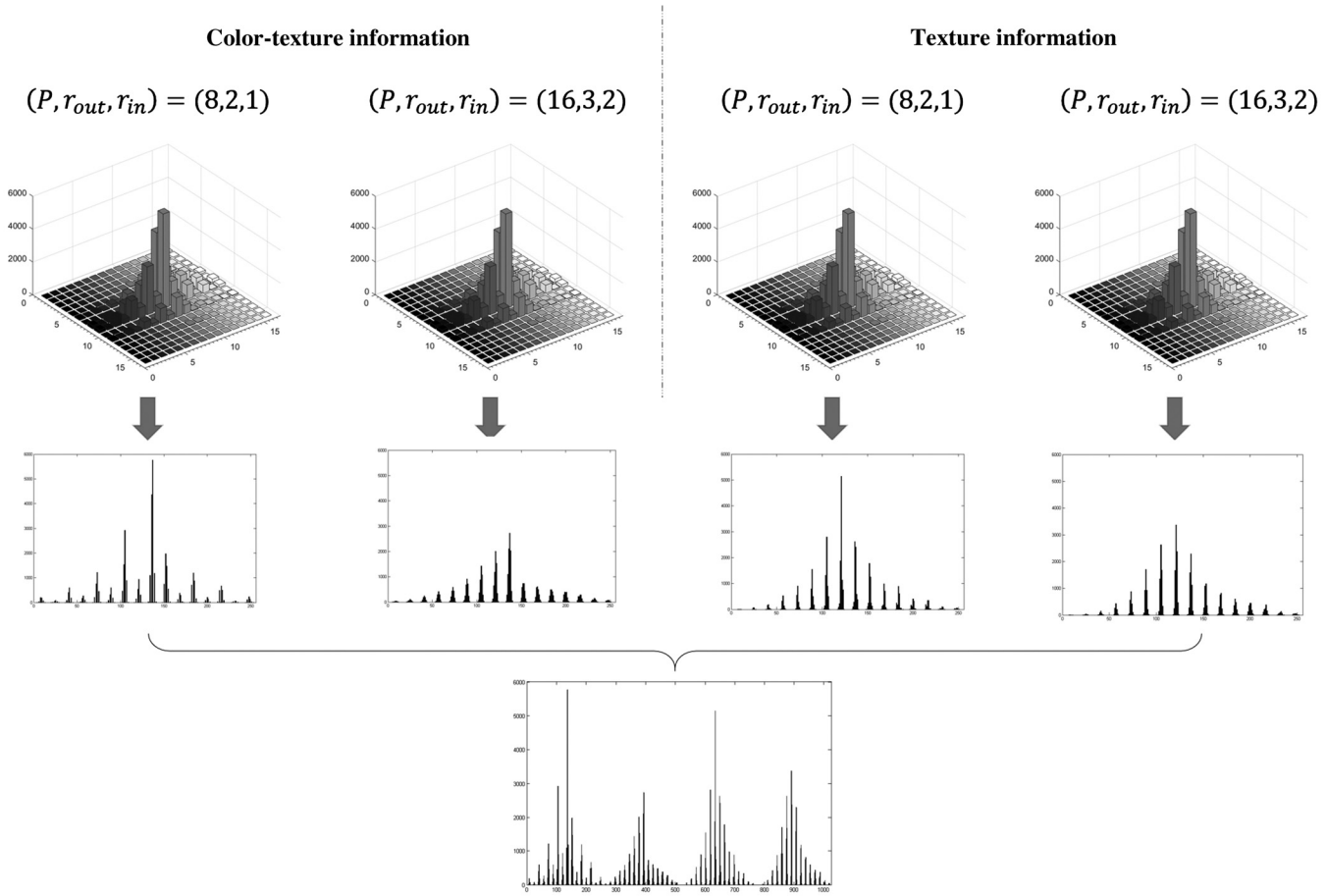

Fig. 6 ECLMP multiresolution feature vector.

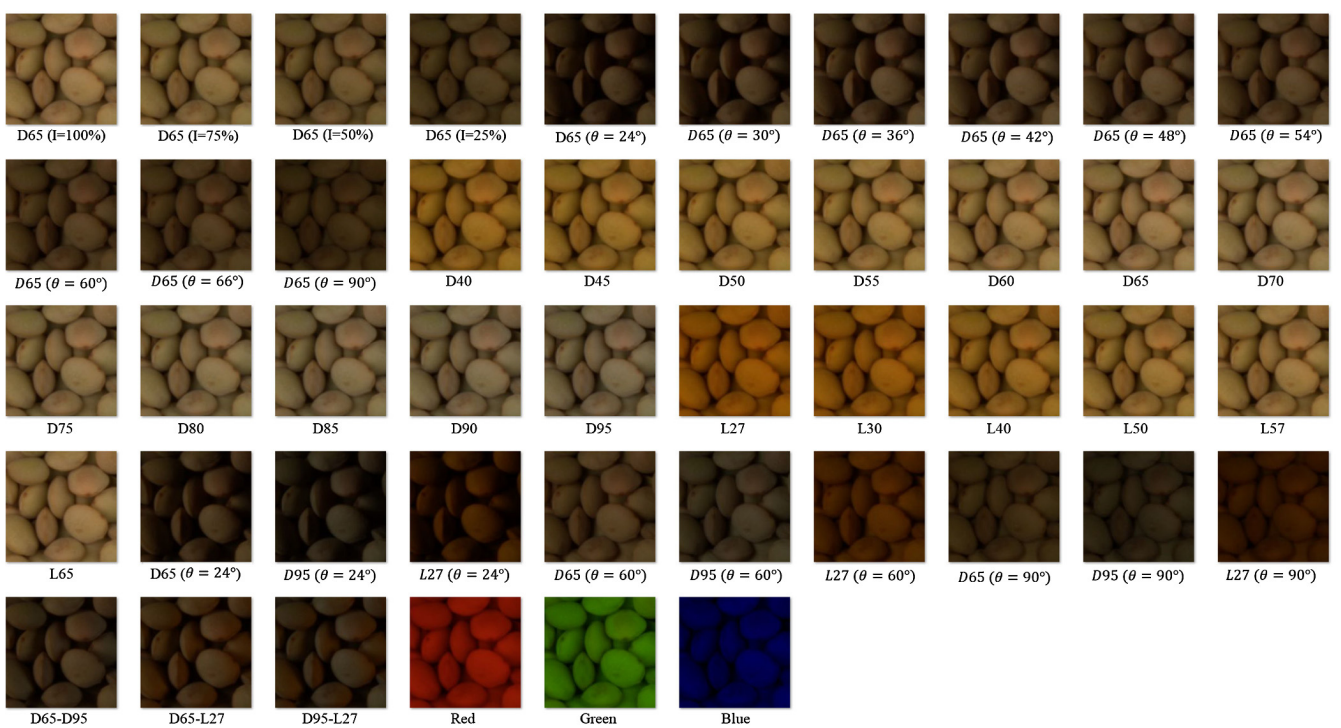

Fig. 7 Example of one of the textures (lentils) in the raw food texture database imaged under the 46 lighting conditions.

Fig. 8 One example image from each of four samples of a category in the KTH-TIPS-2b dataset. The presented samples were imaged at scale 4, frontal pose and frontal illumination.

D45, D50, ..., D95) with the six LED temperatures (L27, L30, L40, L50, L57, and L65).

7. *Temperature or direction*: 72 subsets. Each subset is composed of training and test samples that differ in color (D65, D95, and L27), or direction ($\theta = 24$ deg, 60 deg, 90 deg), or both, color and direction.
8. *Temperature and direction*: 36 subsets, each one composed of training and test samples that differ in both color (D65, D95, and L27) and direction.
9. *Multiple illuminants*: six subsets obtained by combining the three acquisitions with multiple illuminants (D65 to D95, D65 to L27, and D95 to L27).

The KTH-TIPS-2b database presents controlled variations of illumination conditions but also viewpoint and scale. The database includes 11 texture categories and 4 samples per category (Fig. 8). Each sample was acquired considering nine scales, three poses, and four illumination conditions, summing up 4742 samples of 200×200 pixels. Because our goal is to investigate the robustness of the proposed descriptor under different illumination conditions, we did not use the standard evaluation protocol for KHT-TIPS-2b. We selected the images taken under the frontal pose, at the scales 2, 3, 4, 5, and 6 and three illumination conditions: frontal, 45 deg from the top and 45 deg from the side. The images at scales 3, 4, and 6 were used as training set and those at scales 2 and 5 were used as test set.

Note that the image scale acquisition varies between training and test sets. This approach was necessary due to the low number of samples (4) at each illumination at a given scale. As a result, the robustness against changes in illumination is not being assessed isolated. The results will reflect the robustness to changes in illumination and scale combined.

The USPtex dataset is composed of natural color textures such as rice, seeds, fabric, walls, and vegetation. The images were acquired "in the wild," therefore, under uncontrolled illumination and viewing conditions.[25] The dataset contains 2292 samples of 128×128 pixels, grouped into 191 texture classes (12 samples per class). The odd samples were used for training and the even samples for testing. Figure 9 shows texture examples of the USPtex dataset.

3.2 *Experimental Setup*

In all experiments, the descriptors were applied to the training and test images to obtain the feature vectors. The distance between test and training feature vectors was measured using

Fig. 9 Texture samples of the USPtex dataset.

the $L1$ distance [Eq. (10)], where S and M stand for the feature vectors of training and test samples, respectively, and B_t is the number of bins in the feature vector

$$D(S,M)=\sum_{b=1}^{B_t}|S_b-M_b|. \tag{10}$$

In this work, our focus is to evaluate the performance of the color texture descriptors, so we chose the simple nearest neighbor classifier (1-NN) to estimate the label of the test samples. The performance of the descriptor is evaluated using the classification accuracy.

3.3 *Parameter Tuning*

The sigmoid mapping functions (f and z) used to extract texture information [Eqs. (2) and (5)] or color–texture information [Eqs. (7) and (9)] are responsible for capturing the relationship between neighboring pixels. The β parameter presented in those functions define the steepness of the curve.

As show in Sec. 2.3, the ECLMP descriptor is formed by two configurations of radii and number of neighbors for both color–texture and texture information extraction, totalizing four operators (Fig. 6).

For each one of the four operators, the descriptor uses two sigmoid functions (f and z) as presented in Sec. 2. We choose to make the β parameter the same for both functions, i.e., $\beta_{c1}=\beta_{r1}$ and $\beta_{c2}=\beta_{r2}$, since the information extracted by them are related to the same configuration, and they are merged into a joint histogram. Thus, there are four β parameters to be tuned, one for each operator.

These parameters can be tuned separately or simultaneously. Previous works[20,26] have shown the descriptors based on the LMP approach perform better when the parameters are tuned simultaneously since they influence each other when the features are combined in a single feature vector.

In this work, we propose to tune the parameters simultaneously using genetic algorithm. The classification accuracy is used as the objective of the fitness function, which is maximized. The MATLAB Global Optimization Toolbox was used for this purpose. For each dataset, a tuning set was generated by randomly selecting two samples per class from the total of training images. Thus, the RawFooT database tuning set contains $68\times 2=136$ images, the KTH-TIPS-2b database tuning set contains $11\times 2=22$ images, and the USPtex database tuning set is composed of $191\times 2=382$ images.

Table 1 Tuned parameters of the CILMP and ELMP descriptors for the RawFooT database.

	CILMP		ECLMP	
	Configuration (P,R)	Parameter (β)	Configuration $(P,r_{\text{out}},r_{\text{in}})$	Parameter (β)
Color–texture information	(8,2)	0.1712	(8,2,1)	0.3401
	(16,3)	0.1386	(16,3,2)	0.9204
Texture information	(8,1)	0.0844	(8,2,1)	0.0131
	(16,3)	0.1320	(16,3,2)	0.1176
	(24,5)	0.9797	—	—

3.4 *Experimental Results for RawFooT Database*

To evaluate the proposed descriptor over the RawFooT database, we reproduced the same nine classification tasks suggested by Cusano et al.[19] and described in Sec. 3.1. In that work, the authors compared the performance of several methods over the database, including not only traditional texture descriptors but also descriptors designed for object recognition and convolutional neural networks. Because the proposed descriptor belongs to the first category, we

Table 2 Average accuracy over the nine tasks obtained by ECLMP and other texture descriptors.

Descriptor	Avg. accuracy (%)
ECLMP	80.39
CILMP[20]	77.79
LCC[10,19]	65.60
OCLBP[16,19]	59.05
LBP lab[3,19]	58.61
LBP L[16,19]	58.61
LBP $I_1I_2I_3$[16,19]	57.44
LBP RGB[16,19]	57.05
Gabor RGB[14,19]	53.49
Gabor L[14,19]	52.25
Granulometry[17,19]	51.75
Opp Gabor RGB[19,27]	46.76
DT-CWT[14,19]	46.69
Gist RGB[19]	45.05
Hist. H V[19]	44.62
Hist. rgb[19,28]	42.98
Hist. RGB[19]	40.96
Chrom. mom.[19,29]	38.05
DT-CWT L[14,19]	36.67
HoG[19,30]	32.98
Hist. L[19]	31.89
Coocc. matr.[19]	14.50
Coocc. matr. L[19]	08.77

compare the performance obtained by the ECLMP to the results provided in Ref. 19 using the traditional texture descriptors. We also compare the ECLMP to our previous proposed method CILMP.[20]

Both ECLMP and CILMP descriptors are parametric models and require a tuning set. The same tuning set was used for both descriptors. Table 1 shows the tuned parameters for the RawFooT dataset.

Table 2 shows the average accuracy over the nine tasks for each descriptor. The proposed ECLMP descriptor obtained the best average accuracy, followed by the CILMP and LCC descriptors. The ECLMP increases the average accuracy in 2.60 percentage points (pp) over the CILMP and 14.79 pp over the LCC.

Figure 10 shows a comparison among the best three descriptors presented in Table 2 considering each one of the nine tasks proposed in Ref. 19. Because the descriptors are evaluated over several image subsets, average accuracy and minimum accuracy are reported in the charts.

For tasks 1 to 8, the proposed ECLMP descriptor obtained higher average accuracy than the CILMP and LCC. Only for task 9, the CILMP performed slightly better than ECLMP. Considering the minimum accuracy, the ECLMP outperforms the other descriptors for tasks 2 to 9.

In task 2, the descriptors are evaluated over image subsets where training and test samples were taken under different light intensities. For this specific task, we notice that ECLMP reaches an improvement of 7.89 pp over CILMP and 27.71 pp over LCC. Furthermore, the minimum accuracy provided by ECLMP in this task is close to the average accuracy reported by CILMP and 14.93 pp higher than the average accuracy obtained using LCC.

From tasks 3, 7, and 8, we notice that changes in the light source direction are the most challenging variations,

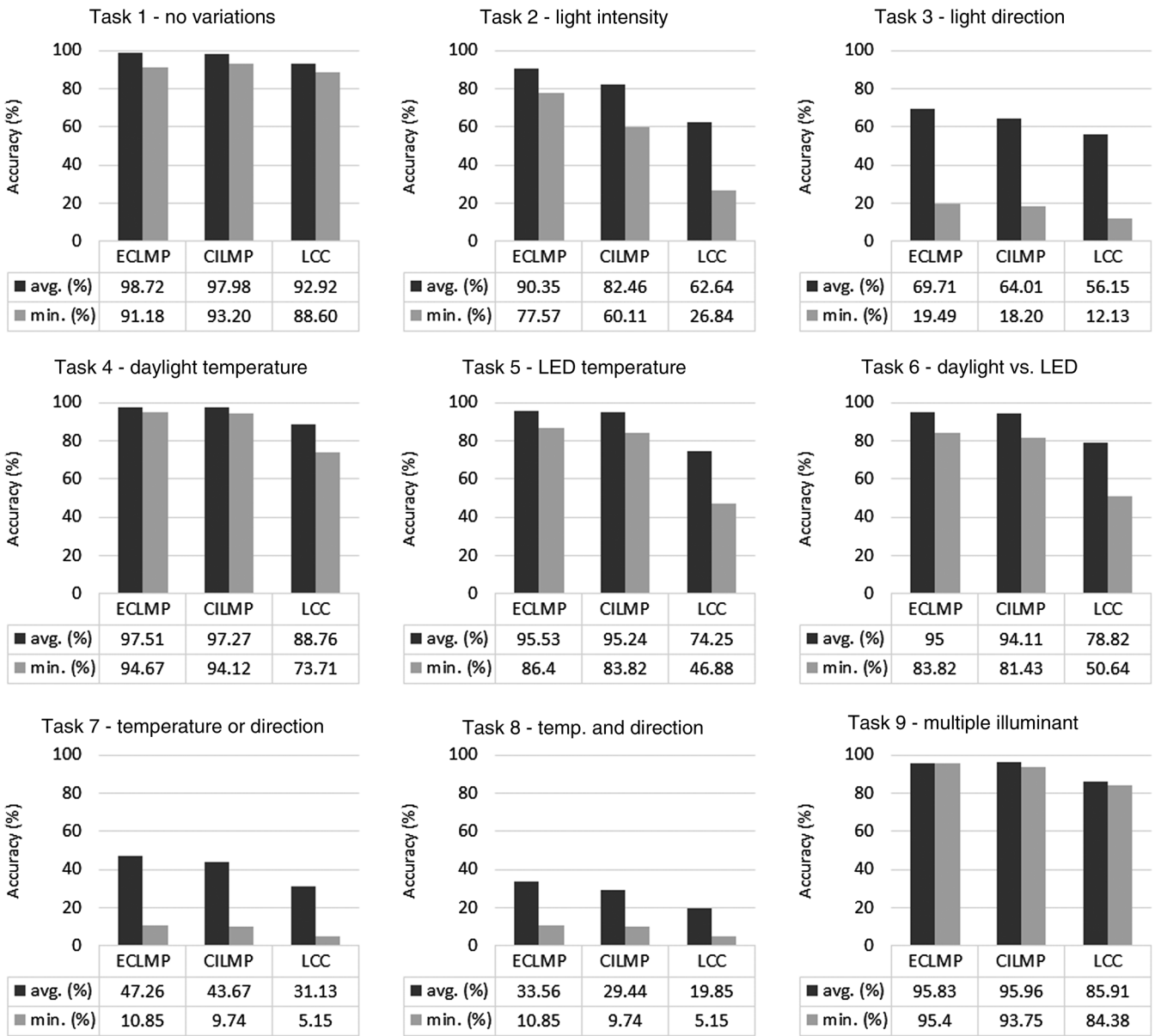

Fig. 10 Accuracy of the ECLMP, CILMP, and LCC descriptors over the RawFoot database.

especially when temperature and direction changes simultaneously (task 8). In task 3, which involves changes only in the light direction, the ECLMP outperforms the average accuracy reported by CILMP in 5.7 pp. For task 7, which evaluates the descriptor considering changes in the light source temperature *or* direction, the improvement achieved by ECLMP is of 3.59 pp over CILMP. Although all the descriptors obtained low classification accuracies for task 8, ECLMP increased the average accuracy in 4.12 and 13.71 pp compared with CILMP and LCC, respectively.

The experiments in the RawFooT dataset have shown that the texture description under variations in the light source depends on the type of variability. Although changes in the light source temperate alone do not affect the performance of the proposed descriptor, changes in the light direction cause a significant decrease in accuracy.

3.5 *Experimental Results for KTH-TIPS-2b Database*

As described in Sec. 3.1, the KTH-TIPS-2b database provides images taken under three different illumination directions: frontal, 45 deg from the top (45 deg top), and 45 deg from the side (45 deg side). We evaluate the robustness of the descriptor performance to changes in the light direction, which is the most challenge condition, as shown in Sec. 3.4. Since ECLMP, CILMP, and LCC provided the best results for RawFooT database, they were chosen for comparison using KTH-TIPS-2b database.

It is important to reinforce that the images in the training and test sets also vary in the acquisition scale. Training set includes images taken under scales 3, 4, and 6, and test set includes images taken under scales 2 and 5. Although scale invariance is not the focus of this work, the experiments performed in KTH-TIPS-2b database also provide some information about the robustness of the descriptors to changes in scale.

Table 3 shows the CILMP and ECLMP tuned parameters for KTH-TIPS-2b database. Table 4 shows the classification accuracies performed by ECLMP, CILMP, and LCC over the combinations of the three illumination conditions.

Table 4 shows that the proposed descriptor achieved the best classification accuracies for experiments 2 and 3. In experiment 3, the ECLMP achieved an improvement of 1.7 and 20.45 pp in the average accuracy over CILMP and LCC, respectively. When the illumination changes between frontal and 45 deg top directions (experiment 1), the LCC descriptor outperformed the descriptors CILMP and ECLMP.

However, notice that the LCC performance varies greatly between experiments. The best LCC performance (experiment 1) achieved 82.95% of accuracy, whereas the worst obtained accuracy was 55.14%, meaning 27.81 pp of variation. Such variation is of 5.69% for the ECLMP and 11.37% for the CILMP.

Overall, in conducted experiments ECLMP was more robust to changes in the light direction than CILMP and LCC.

Table 4 Classification accuracies (%) obtained by ECLMP, CILMP, and LCC over the KTH-TIPS-2b database.

	Illumination direction				
Experiment	Training images	Test images	ECLMP	CILMP	LCC
1	Frontal	45-deg top	77.27	75.00	82.95
	45-deg top	Frontal	76.14	79.55	81.82
2	Frontal	45-deg side	71.59	68.18	60.23
	45-deg side	Frontal	70.45	70.45	61.36
3	45-deg top	45-deg side	75.00	72.73	51.14
	45-deg side	45-deg top	72.73	71.59	55.68
Average accuracy			73.86	72.92	65.53

3.6 *Experimental Results for USPtex Database*

The USPtex database contains color textures acquired "in the wild," under uncontrolled illumination conditions. Therefore, the experiments performed on the USPtex database better simulates a real-life application under uncontrolled lighting conditions. The ECLMP, CILMP, and LCC descriptors were compared in the experiments.

Table 5 shows the tuned parameters of the CILMP and ECLMP descriptors on the USPtex database. The classification accuracies are compared in Fig. 11.

Table 3 Tuned parameters of the CILMP and ECLMP descriptors for the KTH-TIPS-2b database.

	CILMP		ECLMP	
	Configuration (P, R)	Parameter (β)	Configuration (P, r_{out}, r_{in})	Parameter (β)
Color–texture information	(8,2)	0.9502	(8,2,1)	66.7373
	(16,3)	0.7537	(16,3,2)	32.5745
Texture information	(8,1)	0.8693	(8,2,1)	5.1321
	(16,3)	0.1835	(16,3,2)	2.6885
	(24,5)	0.2217	—	—

Table 5 Tuned parameters of the CILMP and ELMP descriptors for the USPtex database.

	CILMP		ECLMP	
	Configuration (P, R)	Parameter (β)	Configuration (P, r_{out}, r_{in})	Parameter (β)
Color–texture information	(8,2)	0.7266	(8,2,1)	18.3727
	(16,3)	0.7121	(16,3,2)	10.0444
Texture information	(8,1)	0.0210	(8,2,1)	1.4005
	(16,3)	0.1971	(16,3,2)	0.0004
	(24,5)	0.3711	—	—

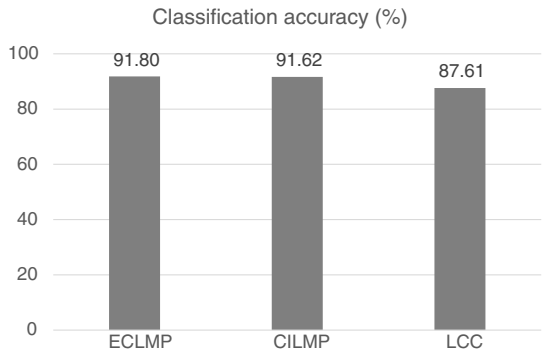

Fig. 11 Accuracy of the ECLMP, CILMP, and LCC descriptors over the USPtex database.

Figure 11 shows that the best classification accuracy was obtained by ECLMP (91.80%), followed by CILMP (91.62%). Moreover, the proposed descriptor achieved an improvement of 4.19 pp over the LCC descriptor.

The results obtained over the USPtex database show that the ECLMP descriptor is more suitable for applications with uncontrolled illumination than the CILMP and LCC.

4 Discussion and Conclusions

In this study, we proposed a descriptor based on the LMP approach, namely extended color local mapped pattern (ECLMP), to address the problem of color–texture classification under different illumination conditions. The proposed descriptor uses two operators, ECLMP_c and ECLMP_r, to incorporates color–texture and texture information from the images to the feature vectors.

The descriptor was evaluated over three databases: RawFooT and KTH-TIPS-2b, which present images taken under controlled variations of illumination, and USPtex, which contains images acquired under uncontrolled illumination conditions. The experimental results showed that the ECLMP outperforms other commonly used descriptors in the three databases. The ECLMP achieved a great improvement in the classification accuracy, especially when the light intensity varies: 7.89 pp over the second best method. Our findings also indicate that changes in the light source direction are the most challenge variations for the texture descriptors. Nevertheless, the ECLMP obtained the best classification accuracies for tasks that include these kinds of variations. The proposed descriptor also proved to be suitable for applications, in which the illumination conditions are not controlled.

In the ECLMP feature vector generation, the codes mapped by ECLMP_r and ECLMP_c are quantized into Q elements. Because the number of features in the ECLMP feature vector is Q^2, increasing Q leads to increased processing time. In this paper, the choice of Q was based on experimental tests using $Q = 2^n$, $n = 3$, 4, 5, 6, 7, and 8. The best trade-off between accuracy and histogram size was obtained with $n = 4$ ($Q = 16$). On a further note, the effect of Q on the accuracy was not too expressive, causing a maximum difference of 4.33, 2.27, and 6.20 pp between the best and worse accuracy for the RawFooT, KTH-TIPS-2b, and USPtex, respectively (see Appendix A).

The choice of the color space used to extract color–texture features can also influence the descriptors performance. This topic has been extensively discussed in Refs. 31 and 32. There is no color space that is well adapted to all databases. Furthermore, the best color space for a given database also depends on the texture descriptor adopted to extract the image features.[31] In the current paper, the RGB color space was used to incorporate color–texture information using the ECLMP descriptor. However, for future work, we suggest the evaluation of other color spaces for such task.

Appendix: The Influence of the Quantization Parameter Q Over the Classification Accuracy

The ECLMP descriptor uses two operators ECLMP_c and ECLMP_r to extract texture or color–texture information. The codes generated by such operators are uniformly quantized in Q elements and stored in a coding map. Afterward, the ECLMP feature vector is obtained by building a joint histogram from the ECLMP_c and ECLMP_r coding maps. Therefore, the Q parameter determines the ECLMP histogram size (Q^2), which can affect the classification accuracy.

The influence of the Q value over the classification accuracy is presented in Tables 6–8, considering the RawFooT, KTH-TIPS-2b, and USPtex databases, respectively.

Table 6 Classification accuracies (%) obtained by the ECLMP descriptor on the RawFooT database experiments, considering different values of Q.

	Experiment #									
Q	1	2	3	4	5	6	7	8	9	Average
8	98.34	90.85	67.75	96.88	96.09	95.77	48.99	35.39	95.07	80.57
16	98.72	90.35	69.71	97.51	95.53	95.00	47.26	33.56	95.83	80.39
32	98.88	86.41	68.62	97.71	92.44	93.53	42.91	30.33	95.87	78.52
64	98.93	83.49	67.33	97.46	90.89	92.69	39.73	27.80	95.5	77.09
128	98.94	81.74	66.29	97.39	90.63	92.48	39.02	27.12	95.62	76.58
256	98.93	82.40	64.28	97.50	90.55	92.59	38.23	26.05	95.62	76.24

Table 7 Classification accuracies (%) obtained by the ECLMP descriptor on the KTH-TIPS-2b database experiments, considering different values of Q.

	Experiment #			
Q	1	2	3	Average
8	73.86	71.59	71.59	72.35
16	76.70	71.02	73.86	73.86
32	75.57	72.73	74.43	74.24
64	76.70	71.59	75.00	74.43
128	75.57	73.30	75.00	74.62
256	75.00	71.02	73.30	73.11

Table 8 Classification accuracies (%) obtained by the ECLMP descriptor on the USPtex database, considering different values of Q.

Q	Classification accuracy
8	87.52
16	91.80
32	90.66
64	93.72
128	92.50
256	93.02

For the experiments in the RawFooT database (Table 6), the best Q value varies between 8 and 32, depending on the variation in the lighting condition (experiments 2 to 9). When there are no changes in the illumination scene (experiment 1), $Q = 128$ yields the highest classification accuracy. Considering the average accuracy, $Q = 8$ presented the best performance, followed by $Q = 16$.

In the KTH-TIPS-2b experiments (Table 7), the Q value that leads to the best accuracies varies between 64 and 128, depending on the experiment. For the USPtex database, as shown in (Table 8), the highest classification accuracy was obtained when $Q = 64$.

Importantly, an improvement of only 4.33, 2.27, and 6.20 pp was achieved when the best Q value was used, compared with the worse (within the investigated range). Considering the results presented for all the databases, and since increasing Q leads to increased processing time, we chose to adopt $Q = 16$.

Acknowledgments

The authors would like to thank the Sao Paulo Research Foundation (FAPESP), grant #2015/20812–5. The authors have no relevant financial interests and no other potential conflicts of interest to disclose.

References

1. A. Drimbarean and P. Whelan, "Experiments in colour texture analysis," *Pattern Recognit. Lett.* **22**(10), 1161–1167 (2001).
2. D. Iakovidis, D. Maroulis, and S. Karkanis, "A comparative study of color-texture image features," in *12th Int. Workshop on Systems, Signals and Image Processing*, pp. 22–24 (2005).
3. T. Maenpaa and M. Pietikainen, "Classification with color and texture: jointly or separately?" *Pattern Recognit.* **37**(8), 1629–1640 (2004).
4. C. Palm, "Color texture classification by integrative co-occurrence matrices," *Pattern Recognit.* **37**(5), 965–976 (2004).
5. M. A. Akhloufi, X. Maldague, and W. B. Larbi, "A new color-texture approach for industrial products inspection," *J. Multimedia* **3**, 44–50 (2008).
6. A. R. Backes, D. Casanova, and O. M. Bruno, "Color texture analysis based on fractal descriptors," *Pattern Recognit.* **45**(5), 1984–1992 (2012).
7. F. Sandid and A. Douik, "Robust color texture descriptor for material recognition," *Pattern Recognit. Lett.* **80**, 15–23 (2016).
8. S. Banerji, A. Sinha, and C. Liu, "New image descriptors based on color, texture, shape, and wavelets for object and scene image classification," *Neurocomputing* **117**, 173–185 (2013).
9. Z. Shao et al., "Improved color texture descriptors for remote sensing image retrieval," *J. Appl. Remote Sens.* **8**, 083584 (2014).
10. C. Cusano, P. Napoletano, and R. Schettini, "Combining local binary patterns and local color contrast for texture classification under varying illumination," *J. Opt. Soc. Am. A* **31**, 1453–1461 (2014).
11. S. H. Lee et al., "Local color vector binary patterns from multichannel face images for face recognition," *IEEE Trans. Image Process.* **21**(4), 2347–2353 (2012).
12. F. Bianconi et al., "Improved opponent colour local binary patterns for colour texture classification," *Lect. Notes Comput. Sci.* **10213**, 272–281 (2017).
13. D. E. Ilea and P. F. Whelan, "Image segmentation based on the integration of colour-texture descriptors-a review," *Pattern Recognit.* **44**(10–11), 2479–2501 (2011).
14. F. Bianconi et al., "Theoretical and experimental comparison of different approaches for color texture classification," *J. Electron. Imaging* **20**(4), 043006 (2011).
15. F. Bianconi et al., "Robust color texture features based on ranklets and discrete Fourier transform," *J. Electron. Imaging* **18**(4), 043012 (2009).
16. T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002).
17. A. Jain and G. Healey, "A multiscale representation including opponent color features for texture recognition," *IEEE Trans. Image Process.* **7**, 124–128 (1998).
18. A. Ledoux, O. Losson, and L. Macaire, "Color local binary patterns: compact descriptors for texture classification," *J. Electron. Imaging* **25**(6), 061404 (2016).
19. C. Cusano, P. Napoletano, and R. Schettini, "Evaluating color texture descriptors under large variations of controlled lighting conditions," *J. Opt. Soc. Am. A* **33**, 17–30 (2016).
20. T. T. Negri et al., "A robust descriptor for color texture classification under varying illumination," in *Proc. of the 12th Int. Joint Conf. on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017)*, Vol. **4**, pp. 378–388, INSTICC, SciTePress (2017).
21. C. T. Ferraz, O. Pereira, Jr., and A. Gonzaga, "Feature description based on center-symmetric local mapped patterns," in *Proc. of the 29th Annual ACM Symposium on Applied Computing (SAC '14)*, pp. 39–44, ACM (2014).
22. L. Liu et al., "Extended local binary patterns for face recognition," *Inf. Sci.* **358**(Suppl. C), 56–72 (2016).
23. L. Liu et al., "Extended local binary patterns for texture classification," *Image Vision Comput.* **30**(2), 86–99 (2012).
24. B. Caputo, E. Hayman, and P. Mallikarjuna, "Class-specific material categorisation," in *Tenth IEEE Int. Conf. on Computer Vision (ICCV 2005)*, Vol. **2**, pp. 1597–1604 (2005).
25. F. Bianconi and A. Fernández, "An appendix to 'texture databases–a comprehensive survey'," *Pattern Recognit. Lett.* **45**(Suppl. C), 33–38 (2014).
26. R. T. Vieira et al., "Human epithelial type 2 cell classification using a multiresolution texture descriptor," in *Proc. of the XIII Workshop of Computer Vision (WVC 2017)* (2017).
27. A. Hanbury, U. Kandaswamy, and D. A. Adjeroh, "Illumination-invariant morphological texture classification," in *Proc. of the 7th Int. Symp. Mathematical Morphology*, pp. 377–386 (2005).
28. A. Oliva and A. Torralba, "Modeling the shape of the scene: a holistic representation of the spatial envelope," *Int. J. Comput. Vision* **42**(3), 145–175 (2001).
29. G. Paschos, "Fast color texture recognition using chromaticity moments," *Pattern Recognit. Lett.* **21**(9), 837–841 (2000).
30. O. L. Junior et al., "Trainable classifier-fusion schemes: an application to pedestrian detection," in *12th Int. IEEE Conf. on Intelligent Transportation Systems (ITSC '09)*, pp. 1–6 (2009).

31. A. Porebski, N. Vandenbroucke, and L. Macaire, "Supervised texture classification: color space or texture feature selection?" *Pattern Anal. Appl.* **16**(1), 1–18 (2013).
32. R. Bello-Cerezo et al., "Experimental comparison of color spaces for material classification," *J. Electron. Imaging* **25**(6), 061406 (2016).

Tamiris Trevisan Negri is an assistant professor at the Federal Institute of Education, Science and Technology of Sao Paulo, Brazil. She received her MSc degree in applied and computational mathematics in 2012 from the Sao Paulo State University. Currently, she is a PhD candidate in the Laboratory of Computer Vision at the University of Sao Paulo. Her major research interests include texture analysis, pattern recognition, and wavelets.

Fang Zhou received her PhD in computer science from the University of Helsinki, Finland. Currently, she is a postdoctoral researcher at the Center for Data Analytics and Biomedical Informatics at Temple University. Her research interests include graph mining, pattern recognition, and social media mining.

Zoran Obradovic is an academician at the Academia Europaea (the Academy of Europe) and a Foreign Academician at the Serbian Academy of Sciences and Arts. He is a L.H. Carnell professor of data analytics at Temple University with appointments at computer and information sciences, statistical science, and the Center for Data Analytics and Biomedical Informatics. He has published more than 350 articles and is cited more than 20,000 times (H-index 53).

Adilson Gonzaga is an associate professor at University of Sao Paulo, Brazil. He is the coordinator of the Computer Vision Laboratory at the Department of Electrical and Computer Engineering, since 1991. He is the CIO of the Information Technology Center of University of Sao Paulo, Sao Carlos campus. He is a member of IEEE, ACM, and SBC (Brazilian Computing Society) and is currently researching in the field of computer vision, image processing, and pattern recognition.