

Research Article

A Novel Local Human Visual Perceptual Texture Description with Key Feature Selection for Texture Classification

Jianning Chi ¹, Xiaosheng Yu ¹, Yifei Zhang ², and Huan Wang²

¹Faculty of Robot Science and Engineering, Northeastern University, Shenyang 110819, China

²College of Information Science and Engineering, Northeastern University, Shenyang 110819, China

Correspondence should be addressed to Jianning Chi; chijianning@mail.neu.edu.cn and Xiaosheng Yu; yuxiaosheng7@163.com

Received 10 July 2018; Revised 30 November 2018; Accepted 8 December 2018; Published 4 February 2019

Academic Editor: Federica Caselli

Copyright © 2019 Jianning Chi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a novel local texture description method which defines six human visual perceptual characteristics and selects the minimal subset of relevant as well as nonredundant features based on principal component analysis (PCA). We assign six texture characteristics, which were originally defined by Tamura et al., with novel definition and local metrics so that these measurements reflect the human perception of each characteristic more precisely. Then, we propose a PCA-based feature selection method exploiting the structure of the principal components of the feature set to find a subset of the original feature vector, where the features reflect the most representative characteristics for the textures in the given image dataset. Experiments on different publicly available large datasets demonstrate that the proposed method provides superior performance of classification over most of the state-of-the-art feature description methods with respect to accuracy and efficiency.

1. Introduction

Texture feature description plays fundamental roles in many computer vision applications, especially general texture classification that can be widely used in material surface inspection [1], medical imaging [2–4], object recognition [5–7], scene recognition [8], and image retrieval [9, 10]. Because of its significance, a large number of texture description approaches have been proposed during the past decades [11–14]. However, it remains a challenging problem to extract highly representative and robust texture features to describe textural images.

Several classical texture descriptors have proved to be effective in reflecting the texture properties. For example, autoregression (AR) model based texture descriptor [15] and Markov random field (MRF) model based descriptor [16] described the texture using the parameters of the models that the pixel intensities were assumed to follow. Edge-based descriptors calculated the gradient magnitudes and directions of the edge pixels in the image and then measured the first-order or second-order statistics of edge distributions by some distance-dependent functions as the texture pattern [17]. Spatial frequency based descriptors have been proposed

based on the assumption that textural characteristics are directly related to the spatial frequencies of texture primitives [18]. The local binary pattern (LBP) [19] and its variants [20–24] exploited the direct comparison between the pixel intensity and its neighbours' intensities within a local area. These methods, the so-called low-level texture descriptors, described the texture by the intuitive features of the texture itself, including the occurrence and direction of the edges or corners, the shape of the textures, the occurrence frequency of the texture primitives, the variance of the pixel intensities, and other primitive level features. Since low-level descriptors typically lacked in giving significant attention to explore the relationship between the texture characteristics they described and human visual sense, image descriptors for high-level features that described how human beings observed the different texture regions have attracted more attention in recent years. However, the high-level texture descriptors still need to solve two problems:

- (1) how to bridge the high-level texture characteristics and the low-level statistical measurements;
- (2) how to reduce the unnecessary texture characteristics to avoid redundancy.

For the first problem, several methods have been proposed to define the high-level texture characteristics differently. One of the most classic high-level texture descriptor is the gray-level cooccurrence matrix (GLCM) texture descriptor proposed by Haralick et al. [25], where they defined 14 characteristics based on the gray-level cooccurrence matrix. In the 14 defined characteristics, the energy, entropy, contrast, homogeneity, and correlation were the most popular-used ones to describe a texture from others. Another famous high-level descriptor was proposed by Laws et al. in [26], where they calculated the uniformity, density, coarseness, roughness, regularity, linearity, directionality, frequency, and phase by assessing average gray level, edges, spots, ripples, and waves in texture [27]. Since then, a number of approaches were proposed to describe the high-level characteristics of the texture. Tomita and Tsuji defined the texture density, directionality, coarseness, randomness, linearity, and periodicity in [28]. M. Amadasun et al. [29] defined five basic characters of texture, namely, coarseness, contrast, busyness, complexity, and texture strength in terms of spatial changes in intensity. Rao and Lohse et al. [30] constructed a three-dimensional space for texture description, where the three orthogonal dimensions were (1) repetitive vs. nonrepetitive, (2) high-contrast and nondirectional vs. low-contrast and directional, and (3) granular, coarse, and low-complexity vs. nongranular, fine, and high-complexity. K. Fujii et al. [31] presented a set of texture characters (contrast, coarseness, and regularity) corresponding to perceptual properties of visual texture based on autocorrelation function (ACF), showing how human subjects seemed to be sensitive to the same parameters captured by a certain descriptor (ACF). H. Tamura et al. [32] presented that six textural characteristics had high correlation to human visual perception, which were coarseness, contrast, directionality, line-likeness, regularity, and roughness. They developed both computational and psychological measurements for these characteristics and made a comparison between these two measurements. The six Tamura's texture characteristics have proved to be most accurate and efficient in describing the texture according to the human visual perception so far. However, one drawback of the existing descriptors is the incorrect calculation of the high-level texture characteristics by using the low-level statistical measurements [33]. Moreover, most of the descriptors capture the texture characteristics globally for the given image, so they cannot precisely describe the variances of the texture characteristics in the textural images with multiple textures.

For the second problem, dimensionality reduction methods have been proposed as a usual solution, which reduced the number of random variables under consideration to obtain a set of "uncorrelated" principal variables [34]. Most of the dimensionality reduction techniques can be divided into feature extraction and feature selection [35]. Feature extraction approaches transformed data in a high-dimensional space to a lower-dimensional space. The transformations can be both linear (such as in principal component analysis (PCA) [36], linear discriminant analysis (LDA) [37], and maximum margin criterion (MMC) [38]) and nonlinear (such as kernel PCA, locally linear embedding (LLE) [39],

and isomap [40]). However, transformed features usually lacked in obvious semantic meaning, making it difficult to understand the image and analyse image components representing different high-level visual characteristics. Feature selection techniques [41–43] have proved to be more effective than feature extraction approaches for dimensionality reduction of image features because they pick a subset of the original features rather than finding a mapping that uses all of the original features. Several feature selection methods have been proposed and applied in different situations. Among them, the optimality properties of PCA have attracted research on PCA-based variable selection methods [44–47]. However, these methods have the disadvantage of either being too computationally expensive, or choosing a subset of features with redundant information.

To overcome the above problems coming from the definition and measurement of high-level texture characteristics and the feature selection, this paper proposes a novel high-level texture description with the selection scheme of representative texture characteristics. Figure 1 illustrates the processing of the proposed texture description method. We redefine six texture characteristics that were originally defined by Tamura et al. [32], including coarseness, contrast, directionality, line-likeness, regularity, and roughness, by measuring each of them with novel local metrics so that these measurements reflect the human perception of each characteristic more precisely. To achieve more effective performance of description, we propose a PCA-based feature selection method exploiting the structure of the principal components of the feature set to find a subset of the original feature vector, where the features reflect the most representative characteristics for the textures in the given image dataset. The experimental results demonstrate that the proposed method provides superior performance in classification accuracy and efficiency over the state-of-the-art methods on three different databases.

The rest of this paper is organized as follows. We present the details of the proposed method in Section 2. The experiments are implemented and the results are discussed in Section 3. Finally, we conclude our proposed work in Section 4.

2. Materials and Methods

The proposed texture descriptor consists of two parts: (1) novel local definitions and measurements of six texture characteristics that were named and defined by Tamura et al. [32] and (2) selection of the most representative features whose corresponding principal component analysis (PCA) coefficients exhibit the largest orthogonality to each other. The two parts are introduced in detail in this section.

2.1. Local Tamura's Texture Description

2.1.1. Tamura's Texture Description. In [32], Tamura et al. proposed six human visual perceptual texture features based on psychological experiments. The six texture features were named as coarseness, contrast, directionality, line-likeness, regularity, and roughness, of which the definitions are reviewed as follows.

TABLE 1: The mathematical definitions of six original Tamura's texture features: coarseness, contrast, directionality, line-likeness, regularity, and roughness.

Feature	Definition
coarseness	$F_{crs} = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n S_{best}(i, j)$
contrast	$F_{con} = \frac{\sigma}{(\alpha_4)^n}$
directionality	$F_{dir} = 1 - r \dot{n}_p \sum_p \sum_{\Phi \in w_p} (\Phi - \Phi_p)^2 \dot{H}_D(\Phi)$
line-likeness	$F_{lin} = \frac{\sum_{i=1}^m \sum_{j=1}^n P_{D,d}(i, j) \cos[(i-j)(2\pi/n)]}{\sum_i} = 1^m \sum_{j=1}^n P_{D,d}(i, j)$
regularity	$F_{reg} = 1 - r (\sigma_{crs} + \sigma_{con} + \sigma_{dir} + \sigma_{lin})$
roughness	$F_{rgh} = F_{crs} + F_{con}$

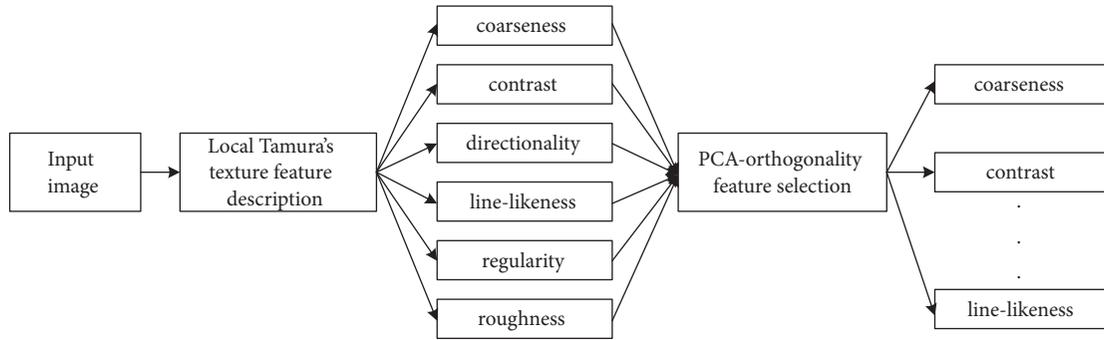


FIGURE 1: The process of our texture description method. Six texture characteristics: coarseness, contrast, directionality, line-likeness, regularity, and roughness are re-defined and extracted from the input image. Together with the PCA-orthogonality feature selection scheme, the most representative features are selected and used to describe the texture image.

Coarseness. Coarseness was defined as the measurement of the size of the primitive elements (texels) composing the texture. The computational procedure was defined as shown in Table 1, where m and n are the effective width and height of the image, and $S_{best}(i, j)$ denotes the neighbourhood size which generates the highest similarity of averaging intensity of the neighbourhood centred at (i, j) .

Contrast. In Tamura's work, they assumed the contrast difference between two texture patterns with different structures was influenced by the following two factors: (1) dynamic range of gray levels and (2) polarization of the distribution of black and white on the gray level histogram or ratio of black and white areas.

Given the two factors, the contrast was defined as F_{con} in Table 1, where σ represented the standard deviation, α_4 denoted the kurtosis of the intensity histogram as a measurement of polarization, and n was a positive number choosing from 8; 4; 2; 1; 1/2; 1/4; 1/8. This calculation of contrast could reduce the values for distributions with biased peaks while almost preserving those for polarized distributions.

Directionality. In Tamura's descriptor, directionality was measured using the histogram of directional angles of the oriented local edges, i.e., the frequency distribution of oriented local edges against their directional angles. The Tamura's directionality was calculated as the F_{dir} in Table 1, where H_D

was the direction histogram, n_p was the number of peaks and Φ_p was the p -th peak position in the histogram H_D , w_p was the range of p -th peak between valleys, and Φ was the quantized direction angle (cyclically in modulo 180°), while r was a normalizing factor related to the number of quantization levels of Φ .

Line-likeness. Tamura defined the word line-likeness as a characteristic of texture that was composed of lines. For this purpose, when the direction and the neighbouring edges' directions for a given edge were nearly equal, they regarded such a group of edge pixels as a line. Therefore, the measurement of F_{lin} in Table 1 was defined so that cooccurrences in the same direction were weighted by +1 and those in the perpendicular direction by -1. In the mathematical function, $P_{D,d}$ was the direction cooccurrence matrix whose element $P_{D,d}(i, j)$ was defined as the relative frequency with which neighbourhoods centred at two pixels separated by a distance d along the edge direction occur on the image, one with the quantized direction i and the other with the quantized direction j .

Regularity. Tamura assumed that if any feature of a texture varied over the whole image, the image is irregular. As shown in Table 1, they took the sum of the variation for each of the four defined features extracted from subimages of the image defined above as the measurement of regularity F_{reg} , where r

was the normalizing factor and each σ_{xxx} meant the standard deviation of F_{xxx} .

Roughness. According to the results of their psychological experiments on vision, Tamura emphasized the effects of coarseness and contrast on roughness and approximated a measurement of roughness as shown in Table 1.

From their own experiments [32] and the following research [48], Tamura's texture descriptor has proved to be successful in reflecting the human visual perception on textures and bridging the semantic gap between low-level features and high-level concepts. However, all the six Tamura's texture characteristics described the textural image globally, resulting in difficulties in discriminating the different textures in the image. Therefore, in the next part, Tamura's descriptor is improved to represent texture features locally, still following the definition and the understanding of the texture characteristics.

2.1.2. Local Tamura's Texture Description. Due to the discussion in Section 2.1.1, in this part we will focus on redefining six Tamura's texture features for local texture description. When extending descriptors from global to local, a common method is as follows: for each pixel, treat the neighbourhood centred at it as an image, then apply the texture descriptor to this neighbourhood, and assign the feature vector to the centre pixel. However, this simple "global to local" strategy is not applicable for the calculation of coarseness because it requires estimating the size of the image block that maximizes the intensity similarity in it and minimizing the intensity similarity to its adjacent blocks with the same estimated size, which cannot be calculated in a local neighbourhood. As a result, regularity and roughness cannot be calculated locally following the original definitions since they both rely on the measurement of coarseness. Therefore, we research the definition of each feature and calculate them by the low-level calculations that measure the relation among pixels in the local neighbourhood centred at each pixel.

Coarseness. As shown in Figures 2(a) and 2(g), larger texels composing the texture should result in higher coarseness value. In our work, coarseness is considered to measure the density of the texels in the neighbourhood: the more the texels are there in the local window, the smaller sizes the texels have, and the lower the coarseness should be. Therefore, the computation procedure is defined as follows.

Step 1. Calculate the edges in the given image I by applying the Laplacian of Gaussian (LoG) filter [49] to the image and finding the zero crossings, resulting in the edge map E .

Step 2. Connect the edges by the morphological transformation that sets 0-valued pixels to 1 if they have two nonzero neighbours that are not connected [50], so that we get the connected edge map E' :

$$\begin{aligned} E : \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} &\longrightarrow \\ E' : \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}. & \end{aligned} \quad (1)$$

After the edge connection, the edges surrounding the smooth regions as "texels" are mostly detected from the image.

Step 3. Calculate the "texels" image \hat{L} as the negative transformation of the connected edge image E' :

$$\hat{L} = 1 - E'. \quad (2)$$

Then the "texels" are labeled in the image \hat{L} .

Step 4. The coarseness is calculated as the ratio of the number of different texels to the number of the pixels in the neighbourhood:

$$F_{crs} = \frac{\text{diff}(\hat{L})}{M \times N'}, \quad (3)$$

where $\text{diff}(\hat{L})$ calculates the number of different "texels" in the neighbourhood of \hat{L} , while M and N are the size of the neighbourhood.

Contrast. Based on the discussion in Section 2.1 and the observation of the textures with different contrast, as shown in Figures 2(b) and 2(h), the following factors are considered to influence the contrast difference between different texture patterns:

- (1) the range of gray levels: the textures where the distances between the maximum and minimum gray level are large are more likely to have higher contrast;
- (2) the concentration of the distribution of the gray level histogram: the textures where the histogram distributes uniformly in a wide range will exhibit higher contrast;
- (3) the polarization of the distribution of the gray level histogram: a small difference of intensity is negligible if the intensities are biased to 255, while the same small difference matters if the intensities are biased to 0.

Therefore, the contrast of the image can be measured locally as

$$F_{con} = \frac{r_d}{p}, \quad (4)$$

where r_d corresponds to the dynamic range of the gray levels between the maximum intensity I_{max} and the minimum intensity I_{min} :

$$r_d = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}, \quad (5)$$

and p combines the skewness $\text{skew}(\bullet)$ and the kurtosis $\text{kurt}(\bullet)$ of the histogram of the pixel intensities in the neighbourhood I_l to measure the concentration and the polarization of the distribution of the intensities:

$$p = \frac{\text{kurt}(I_l)}{1 + \text{skew}(I_l)}. \quad (6)$$

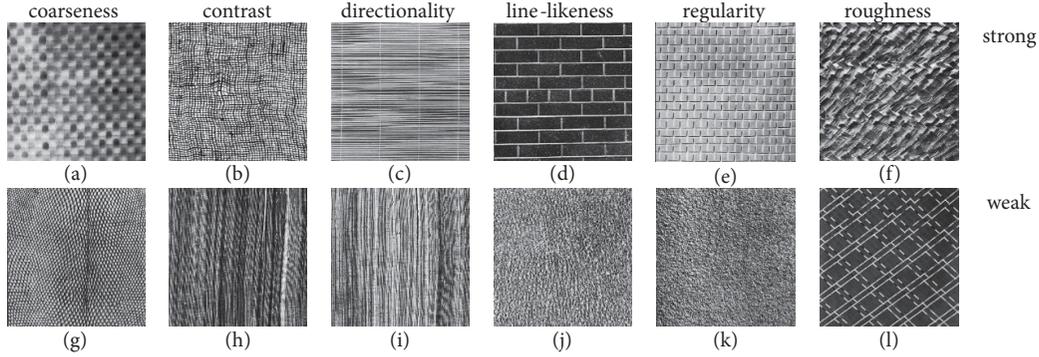


FIGURE 2: Example textures exhibiting the six texture characteristics strongly and weakly. Columns (from left to right): textures with high (upper row) and low (lower row) values with respect to coarseness, contrast, directionality, line-likeness, regularity, and roughness.

Directionality. The concepts of directionality and line-likeness in Tamura's texture descriptor are obscured because the directionality measures the edges with similar directions while the line-likeness measures the coincidence of the edge directions. It can partly explain the incorrectness of the Tamura's descriptor in describing directionality for some textures with random directions. Therefore, in this work, directionality is redefined as the measurement of the mean orientation of all the edges in the local neighbourhood. Moreover, we assume that the horizontal texture has high directionality value, as shown in Figure 2(c), while the vertical texture has low directionality value, as shown in Figure 2(i). Then the directionality of a texture pixel is calculated in its neighbourhood as

$$F_{dir} = \tan^{-1} \left(\frac{\sum_{x=1}^M \sum_{y=1}^N m(x, y) \sin d(x, y)}{\sum_{x=1}^M \sum_{y=1}^N m(x, y) \cos d(x, y)} \right), \quad (7)$$

where M and N are the size of the neighbourhood, and $m(x, y)$ and $d(x, y)$ represent the magnitude and direction of the edge pixel (x, y) , which are calculated by

$$\begin{aligned} m(x, y) &= \sqrt{\delta_x(x, y)^2 + \delta_y(x, y)^2}, \\ d(x, y) &= \tan^{-1} \frac{\delta_y(x, y)}{\delta_x(x, y)} + \frac{\pi}{2}, \end{aligned} \quad (8)$$

where $\delta_x(x, y)$ and $\delta_y(x, y)$ represent the pixel-wise horizontal and vertical derivatives using the Sobel operator [51] to the neighbourhood centred at (x, y) .

Line-likeness. The line-likeness measures how many of the edges in the neighbourhood have the same or similar directions. Figure 2(d) shows the texture with high line-likeness where most of the edges are in the similar direction, while Figure 2(j) shows the texture with low line-likeness because of the edges with various directions. The directions at the edge pixels $d(x, y)$ are quantized into n direction intervals

$[0, \pi/n), [2\pi/n, \pi/n) \dots [(n-1)\pi/n, \pi)$. As a result, the line-likeness of the pixel (x, y) is calculated as 1 minus the variance of the local edge orientations:

$$F_{lin}(x, y) = 1 - \frac{\sum_{i=x-\omega}^{x+\omega} \sum_{j=y-\omega}^{y+\omega} (D(i, j) - \mu_D)^2}{(2\omega + 1)^2}, \quad (9)$$

where $D(x, y)$ is the quantized direction of the edge direction $d(x, y)$, μ_D is the mean value of D in the local window, and ω is the radius of the window.

Regularity. The regularity of the texture measures the spatially repetitiveness of texels over a certain distance. For a regular texture shown in Figure 2(e), the texel edges repeat every several pixels, and the distance of this repetitiveness is relatively fixed in contrast to an irregular texture as shown in Figure 2(k). Therefore, this pattern can be calculated by the autocorrelation functions of the edge image as follows:

$$F_{reg}(x, y) = 1 - \frac{\max(C)}{\sum_{i=1}^N (C(i))}, \quad (10)$$

where C is the autocorrelation function of the neighbouring image centred at (x, y) :

$$C = \text{CORR}(I_l, I_l), \quad (11)$$

where $\text{CORR}(\cdot)$ represents the autocorrelation function, and I_l is the local window of the image I . Since the maximum of the autocorrelation function C will only appear when the displacement of the texture is 0, and the regular texture where there must exist another displacement of the repeating texels will have other local peak values in the autocorrelation function, the ratio of the maximum autocorrelation value to the sum of all the local peak autocorrelation values $\max(C) / \sum_{i=1}^N (C(i))$ will be smaller in regular texture than that in irregular one, leading to a larger measurement $F_{reg}(x, y)$.

Roughness. As shown in Figure 2(f), a rough texture means the surface of the texture is uneven and not smooth, and contains many edges, while the texture with quite smooth surface as shown in Figure 2(l) has the low value of roughness.

Therefore, the roughness can be approximated by the density of the edges in the neighbourhood:

$$F_{rgh} = \frac{\sum_{i=1}^M \sum_{j=1}^N E(i, j)}{M \times N}, \quad (12)$$

where E represents the edge map of the neighbourhood centred at (x, y) with the size $M \times N$. Similar to the calculation of coarseness, the edge is computed by the Laplacian of Gaussian (LoG) filter.

As discussed above, the six human visual perceptual texture features that were originally proposed in Tamura's work have been locally calculated with novel methods from low-level measurements. Therefore, the texture image can be described more precisely according to human visual perception and different types of textures can be better differentiated locally.

2.2. PCA-Orthogonality Key Texture Feature Selection. From the above discussion, the six novel-defined texture features can locally describe textural images. Using all these features in describing textural images, however, is sometimes redundant because these features are correlated to each other to some extent. In this part, a PCA-orthogonality key feature selection method is proposed to find the representative characteristics in differentiating textures in the image.

2.2.1. Principal Component Analysis (PCA) in Image Feature Reduction. Principal component analysis (PCA) is one of the most popular methods for dimensionality reduction. As described in [52], for a set of data with p dimensions, PCA aims to find a linear subspace of dimension q lower than p so that the data points mostly lie in this subspace, which is likely to maintain most of the variability of the original data. q orthogonal vectors that form a new coordinate system, called the "principal components," can specify the linear subspace. The principal components are orthogonal, linear transformations of the original data points, so there can be no more than p of them. However, the expectation is that only $q < p$ principal components are required to approximate the space spanned by the p original axes.

Mathematically, the transformation is defined [52] by a set of p -dimensional vectors of weights or loading $\mathbf{w}_{(i)} = (\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_p)_{(k)}$ that map each row vector $\mathbf{x}_{(i)}$ of X to a new vector of principal component scores $\mathbf{t}_{(i)} = (\mathbf{t}_1, \mathbf{t}_2, \dots, \mathbf{t}_k)_{(i)}$, given by

$$\mathbf{t}_{k(i)} = \mathbf{x}_{(i)} \cdot \mathbf{w}_{(k)}. \quad (13)$$

In such a way that the individual variables of \mathbf{t} considered over the data set successively inherit the maximum possible variance from \mathbf{x} , with each weights vector \mathbf{w} constrained to be a unit vector. Therefore the full principal components transformation of \mathbf{x} can be given as

$$\mathbf{T} = \mathbf{XW}, \quad (14)$$

where W is a $p \times p$ matrix whose columns are the eigenvectors of $\mathbf{x}^T \mathbf{x}$. Then the transformation $\mathbf{T} = \mathbf{XW}$ maps a data vector $\mathbf{x}_{(i)}$ from an original space of p variables to a new space of p

variables which are decorrelated over the dataset. However, not all the principal components need to be kept. Keeping only the first q principal components, produced by using only the first q loading vectors, gives the truncated transformation:

$$\mathbf{T}_q = \mathbf{XW}_q, \quad (15)$$

where the matrix \mathbf{T}_q now has n rows but only q columns. In other words, PCA learns a linear transformation $\mathbf{t} = \mathbf{W}^T \mathbf{x}$, $\mathbf{x} \in R^p$, $\mathbf{t} \in R^q$, where the columns of $p \times q$ matrix W form an orthogonal basis for the q features (the components of representation \mathbf{t}) that are decorrelated. By construction of all the transformed data matrices with only q columns, this score matrix maximizes the variance in the original data that has been preserved, while minimizing the total squared reconstruction error $\|\mathbf{TW}^T - \mathbf{T}_q \mathbf{W}_q^T\|_2^2$ or $\|\mathbf{X} - \mathbf{X}_q\|_2^2$.

Following this transformation scheme, PCA has been widely used to reduce the dimensionality of the image features selected by some certain image descriptor [44–47, 53, 54]. Given an image with size $M \times N$ and a local descriptor that extract p features for each pixel in the image, then the data matrix X can be considered as

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{np} \end{bmatrix}, \quad (16)$$

where $n = M \times N$ is the total number of the pixels in the image, and p is the dimensionality of the original feature space for each pixel. Then according to (15), only the first q principal components that generate the largest variances are preserved and the new feature matrix \mathbf{T}_q is

$$\mathbf{T}_q = \begin{bmatrix} y_{11} & \cdots & y_{1q} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nq} \end{bmatrix}, \quad (17)$$

where $\{y_{i1} \dots y_{iq}\}$ is the newly feature vector for each pixel in the image, with a reduced dimensionality of the image feature.

2.2.2. PCA-Orthogonality Key Feature Selection. Given the PCA transformation for image feature vector as described in (16) and (17), the "principal" features are the linear combination of the original features with the coefficients matrix:

$$[Y_1 Y_2 \dots Y_q] = [X_1 X_2 \dots X_p] \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1q} \\ k_{21} & k_{22} & \cdots & k_{2q} \\ \vdots & \vdots & \ddots & \vdots \\ k_{p1} & k_{p2} & \cdots & k_{pq} \end{bmatrix}, \quad (18)$$

where Y_1, Y_2, \dots, Y_q are the q principal components translated from the p original features X_1, X_2, \dots, X_p ($q \leq p$) by the coefficients matrix.

Assuming that there are n observations of the image with respect to the p features, then we have n sets of the original features with p -dimension. After PCA transformation, the n observations of each of the q features with the reduced dimensionality p of the original feature space as follows:

$$Y_i = \begin{bmatrix} k_{1i}X_{11} + k_{2i}X_{12} + \dots + k_{pi}X_{1p} \\ k_{1i}X_{21} + k_{2i}X_{22} + \dots + k_{pi}X_{2p} \\ \vdots \\ k_{1i}X_{n1} + k_{2i}X_{n2} + \dots + k_{pi}X_{np} \end{bmatrix} \quad (19)$$

$$= \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix} \begin{bmatrix} k_{1i} \\ k_{2i} \\ \vdots \\ k_{pi} \end{bmatrix} = X \begin{bmatrix} k_{1i} \\ k_{2i} \\ \vdots \\ k_{pi} \end{bmatrix},$$

where $i = 1, 2, \dots, q$. Since for every i , the observations of the original features X are the same, the column $[k_{1i} \ k_{2i} \ \dots \ k_{pi}]^T$ controls the difference between two principal components Y_i , while the absolute value of the k_{ri} , $r = 1, 2, \dots, p$ represents the weight of the r -th original feature to the i -th principal component.

Differently from the original PCA dimensionality reduction which directly uses Y_1, Y_2, \dots, Y_q as the features to describe the image, the PCA-based feature selection aims to find a subset X' of the original data vector X , so that with the corresponding coefficients k' , which are also the subset of the original coefficients k , the output components Y' are also the "principal" components. In another word, the orthogonality of the output Y' should remain close to that of the original principal components Y .

In our work, the orthogonality of the component matrix Y is defined by the cosine dissimilarity between each two components Y_i and Y_j in Y :

$$\text{Ortho}(Y) = \frac{\sum_{i=1}^q \sum_{j=1}^q \text{Diss}(Y_i, Y_j)}{q^2}, \quad (20)$$

Where q is the number of components in Y , and the cosine dissimilarity $\text{Diss}(\bullet)$ is defined as

$$\text{Diss}(Y_i, Y_j) = 1 - \frac{\sum_{p=1}^n Y_{ip}Y_{jp}}{\sqrt{\sum_{p=1}^n Y_{ip}^2} \sqrt{\sum_{p=1}^n Y_{jp}^2}}, \quad (21)$$

where n is the number of observations of the image data. However, it is difficult to calculate the dissimilarity of the Y_i and Y_j directly because n is always large as the total number of the image pixels and varies in images with different sizes. We instead use the weighting coefficients matrix k to measure the dissimilarity because according to (19), the columns $[k_{1i} \ k_{2i} \ \dots \ k_{pi}]^T$ and $[k_{1j} \ k_{2j} \ \dots \ k_{pj}]^T$ are able to control the difference between two principal components

Y_i and Y_j . Then the orthogonality of the component matrix Y can be calculated by the coefficient matrix k as

$$\text{Ortho}(Y) = \frac{\sum_{i=1}^q \sum_{j=1}^q \text{Diss}(k_i, k_j)}{q^2}, \quad (22)$$

where q is also the number of columns in k corresponding to the components in the Y , and

$$\text{Diss}(k_i, k_j) = 1 - \frac{\sum_{l=1}^p k_{li}k_{lj}}{\sqrt{\sum_{l=1}^p k_{li}^2} \sqrt{\sum_{l=1}^p k_{lj}^2}}, \quad (23)$$

where p is the dimensionality of the input features X . Since p is based on the descriptor used to represent the image, p is a constant number for all the images, which simplifies the calculation of the orthogonality. Therefore, the PCA-based feature selection can be mathematically described as

$$\{s\} = \arg_s \{ \text{Ortho}(k_{(s)}) \approx \text{Diss}(k_{(p)}) \}, \quad (24)$$

where s is the subset of the original features and the number of the dimensionality of the selected subset is l , while p represents the full set of the original features. The algorithm of the PCA-orthogonality feature selection (PCA-ORTH) is implemented as Algorithm 1:

- (1) implement the PCA to the data vector consisting of the features of the images in the dataset;
- (2) find all the possible subsets of the coefficients matrix k . Each subset is formed by selecting the rows from k , corresponding to the selected features;
- (3) calculate the dissimilarities between every two columns of all the possible subsets, and then compute the orthogonality of the components matrices represented by those subsets;
- (4) compare the orthogonality of the possible matrices, and select the key features represented by the subset of the coefficients matrix that derive the output components matrix with the orthogonality closest to the original one.

3. Results and Discussion

To evaluate the performance of the proposed PCA-orthogonality key local Tamura's texture description method (LT+PCA-ORTH), different texture description methods combined with different feature selection methods are compared to classify textures in the given testing texture databases.

4. Experimental Materials

4.1. *Image Database.* There are three image databases used in our experiments to evaluate different texture descriptors:

- (i) 32 texture images from the Brodatz texture database [55]; Figures 3(a), 3(b), and 3(c) show three example textures in the database;

- (1) Initialize $L = 1$, the number of selected features, $s_{select} = \{1, 2, \dots, p\}$, the selected features set, $ortho = 0$, the orthogonality of the principal components computed from the selected features, k is the principal component coefficients matrix with p rows corresponding to the original features.
- (2) Perform $p - 1$ times:
 - (i) Calculate the set s of possible combinations by choosing L features from the original p features and the number of combinations C_p^L .
 - (ii) Initialize $i = 1$, the order of the combination in the set s .
- (3) Perform C_p^L times:
 - (i) Calculate the orthogonality $ortho_{(i)}$ of the subset consisting of the i -th combination of rows in the principal component coefficients matrix k .
 - (ii) If $ortho_{(i)} > ortho$, update the orthogonality by $ortho = ortho_{(i)}$, update the selected features set by $s_{select} = s(i)$. Else, skip.
 - (iii) Update the order i by $i = i + 1$.
 - (iv) Update the number of selected features L by $L = L + 1$.
- (4) Return the selected features set s_{select} .

ALGORITHM 1: The algorithm for PCA-based feature selection.

- (ii) 25 texture images from the UIUCTex texture database [56]; Figures 3(d), 3(e), and 3(f) show three example textures in the database;
- (iii) 32 texture images from the CURET texture database [57]; Figures 3(g), 3(h), and 3(i) show three example textures in the database.

For Brodatz and UIUCTex texture database, since there are not enough samples of each texture (1 sample for each texture in Brodatz and 40 samples for each texture in UIUCTex), the training and testing samples for each texture are generated from the given texture itself. As Figure 4 shows, the training and testing samples for a given texture are generated as follows:

- (1) each texture with the size of 512×512 is cut into $8 \times 8 = 64$ patches with the size of 64×64 , shown as those in Figures 4(a) and 4(b); we choose 64×64 as the size of each texture patch because it is a good trade off: size larger than 128×128 would result in too few patches of the original image to generate enough texture samples, while size smaller than 32×32 would result in too shattered patches to generate the samples preserving enough original texture information;
- (2) the 64 patches are randomly divided into 8 groups, each of which contains 8 patches; in each group, the 8 patches are randomly combined (with repetition) into 8 rows and 8 columns and then merged into one image; this step is repeated 8 times in each group, resulting in 8 images with the same type of texture; Figures 4(c) and 4(d) show two example randomly generated texture images in two corresponding groups;
- (3) 8-fold cross validation is implemented with the generated $8 \times 8 = 64$ images: 8 generated images in 1 group are used as the testing image set for one texture, 56 generated images in the other 7 groups are used as the training image set; the process is repeated 8 times by

using every group of 8 images as testing set while the others as training set.

4.1.1. Comparator Texture Descriptor. The state-of-the-art texture description methods and feature selection methods are combined as the comparators to the proposed method in finding the most representative local texture features. The following texture description methods are used:

- (1) the local binary pattern (LBP) descriptor [19], which described the texture by computing the LBP code from comparing a pixel intensity with its neighbours;
- (2) the scale invariant feature transform (SIFT) descriptor [58], which combined a scale invariant region detector and a description based on the gradient distribution in the detected regions;
- (3) the Gabor transform based texture descriptor (Gabor) [59], which applied the Gabor coefficients of the image as the image description;
- (4) the Tamura's texture descriptor (TM) [32], which originally defined the coarseness, contrast, directionality, line-likeness, regularity, and roughness based on the computational and psychological measurements.

And the following state-of-the-art feature selection methods are combined with all the texture descriptors as the comparators to the proposed PCA-ORTH feature selection method:

- (1) the mean square prediction error minimization of PCA method (PCA-MSPEM) [46], which found the subset of the feature set that minimized the trace of a measure matrix of the covariance matrix of the input feature vector;
- (2) the approximation of principal features by using the absolute value of the coefficients of the principal components (PCA-APX) [52], which chose the variables corresponding to the highest coefficients of each of the first q principal components as the principal features from the given ones;

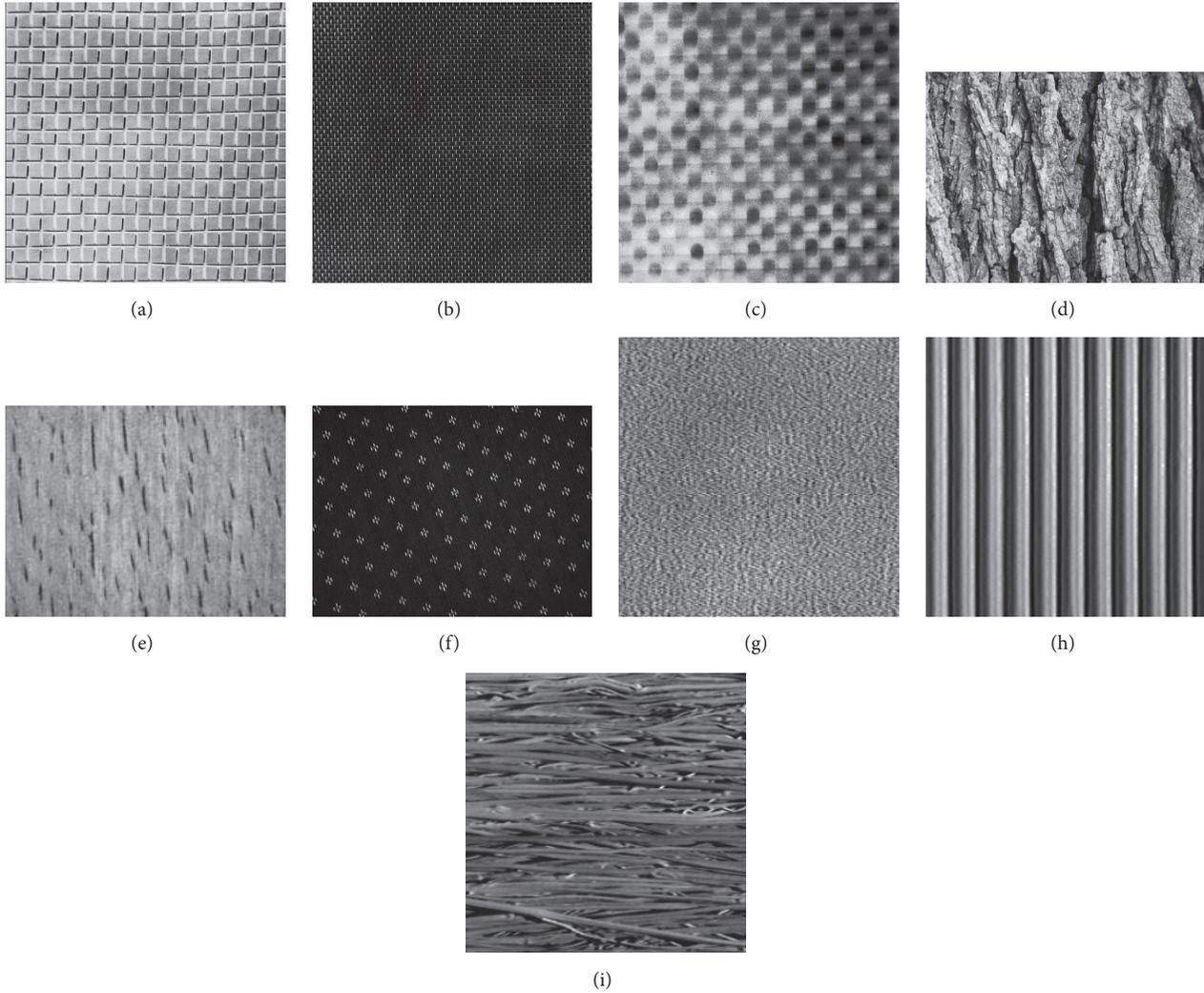


FIGURE 3: Examples of textures in Brodatz texture database, UIUCTex texture database, and CURET texture database. (a), (b), and (c) are the textures from Brodatz database; (d), (e), and (f) are the textures from UIUCTex database; (g), (h), and (i) are the textures from CURET database.

- (3) the principal feature analysis (PFA) [47], which clustered all the rows in the principal component coefficient matrix by K-Means algorithm and selected the feature corresponding to the row vector that was closest to the mean of each cluster as the principal feature.

4.1.2. Implementation of the Experiment. With the generated training and testing samples of each texture in each database and the given texture description methods, the experiment of classifying different textures in each database is implemented through MATLAB2017b on an Intel i7-7700 2.60GHz PC with 8G RAM as follows.

- (1) The local Tamura's features, consisting of the 6 specific characteristics, coarseness, contrast, directionality, line-likeness, regularity, and roughness, are calculated in the local neighbourhood. The selection of the neighbourhood size will be discussed in the following

sections. For each sample image, we first calculate the $(512 \times 512) \times 6 = 262144 \times 6$ features for all the $512 \times 512 = 262144$ pixels. Then the mean value of the feature values for all the pixels is considered as the certain feature value for the image; that is, 1×6 feature vector is calculated for each sample image. Similarly, other texture features are calculated by the comparing texture description methods.

- (2) The most discriminating features are selected from different texture features using different feature selection methods, and then for each sample image a $1 \times q$ ($q \leq 6$) vector of the most discriminating feature for each feature description method is generated.
- (3) The testing samples of textures are classified by the multiclass SVM classifier and different texture features. The classifiers are trained with the corresponding feature vectors of the training samples.

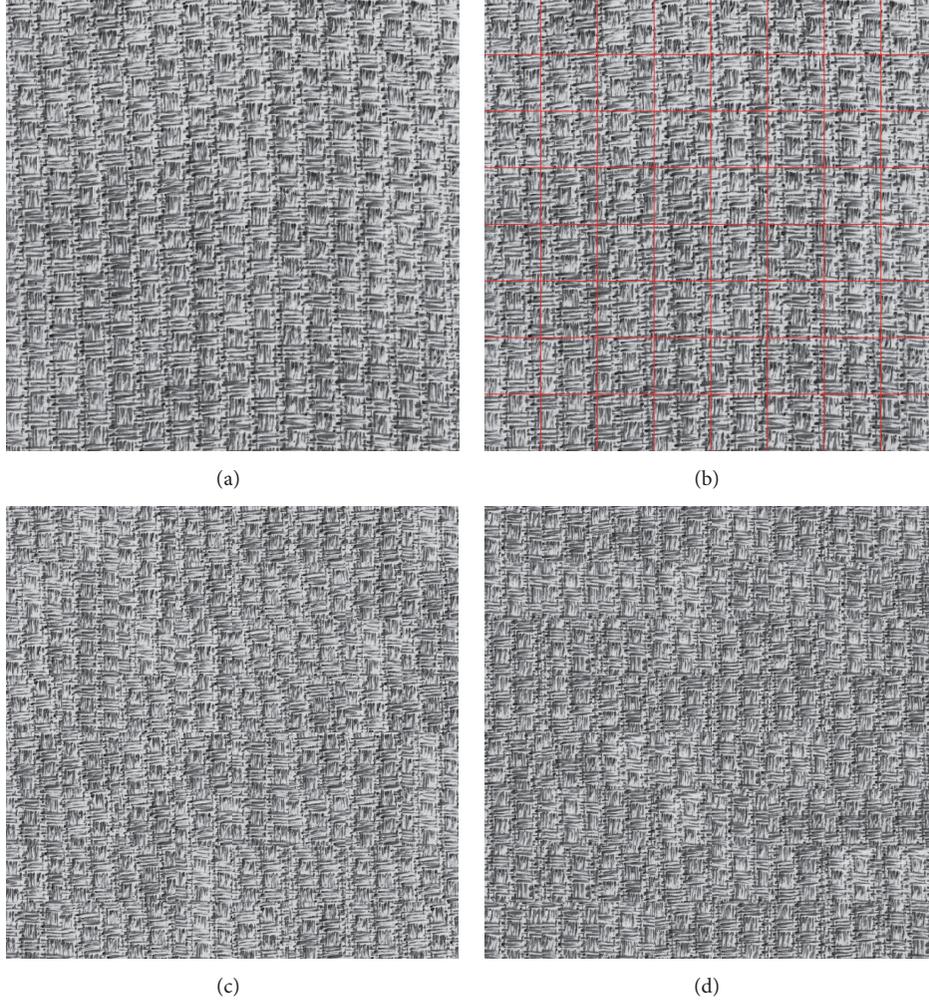


FIGURE 4: Example of generating multiple training and testing texture images from a given texture. The texture is cut into $8 \times 8 = 64$ patches with the size of 64×64 , and then the 64 patches are randomly divided into 8 groups, each of which contains 8 patches. In each group, the 8 patches are randomly combined (with repetition) into 8 rows and 8 columns and then merged into one image. This step is repeated 8 times in each group, resulting in 8 images with the same type of texture. (a) The sample texture. (b) 64 patches of the sample. (c) Regenerated texture by merging the 8 partitions in one group randomly. (d) Regenerated texture by merging the 8 partitions in another group randomly.

- (4) The classification accuracy of each classification is defined as

$$AR = \frac{TP + TN}{TP + FP + TN + FN}, \quad (25)$$

where the traditional true positive (TP), true negative (TN), false positive (FP), and false negative (FN) were described in [60].

- (5) The confusion matrix of each classification is defined as

$$CM_{i,j} = \frac{S_{i,j}}{S_i}, \quad \text{for } i = 1, 2, \dots, N., \quad j = 1, 2, \dots, N, \quad (26)$$

where $CM_{i,j}$ represents the element $(i; j)$ in the confusion matrix CM, $S_{i,j}$ is the number of samples

whose target is the i -th class that is classified as j -th class, and S_i denotes the total number of the samples belonging to i -th class. And the confusion rate is calculated as

$$R_c = 1 - \frac{\sum_{i=1}^N CM_{i,i}}{N}, \quad (27)$$

where N is the total number of the texture classes in each database.

- (6) The classification accuracy, the confusion matrix, and confusion rate of classifying textures in each database by all the texture description methods are compared to evaluate the effect of different descriptors in extracting the most representative features for the textures in a given database.

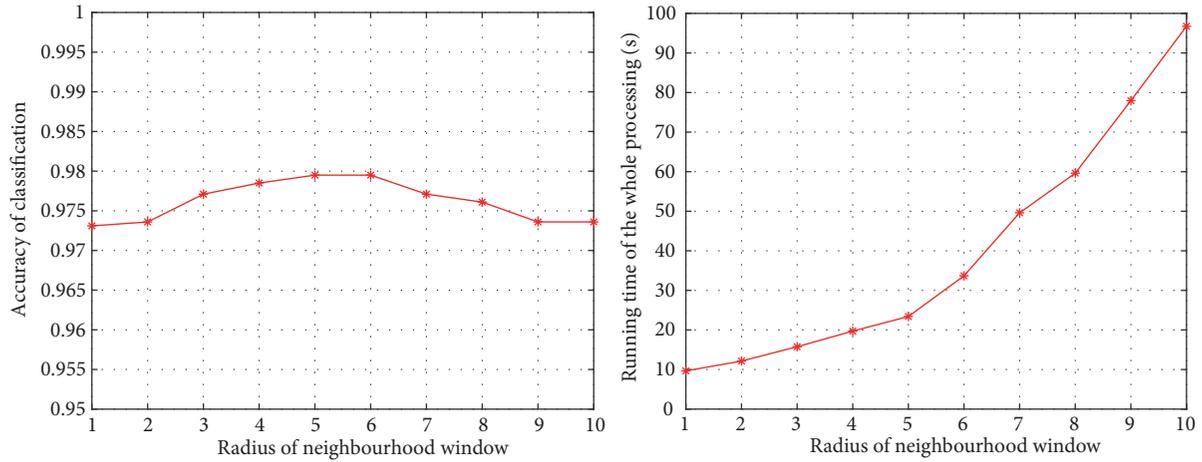


FIGURE 5: The relationship between the radius of the neighbourhood and the classification accuracy as well as the running time of the whole processing with the proposed method.

5. Experimental Results

In this section, we show and analyse the experimental results of classifying textures in different texture image databases separately.

5.1. Experiments on Brodatz Texture Database. Figure 5 shows the relationship between the radius of the neighbourhood and the classification accuracy as well as the running time of the whole processing of the proposed local Tamura’s texture description (LTM) and the PCA-orthogonality (PCA-ORTH) feature selection. It is considered that 5 is a good trade-off as the radius of the neighbourhood: it achieves the highest classification accuracy (97.95%) with the running time as short (23.40s) as possible.

Figure 6 shows the average accuracy of classifying 32 textures in Brodatz database by multi-SVM, using the texture features extracted by different methods. On one hand, the proposed “full-feature” local Tamura’s texture description (LTM) (without feature selection method) provides higher classification accuracy (97.41%) than the other “full-feature” descriptors (97.02%, 96.63%, 96.04%, and 95.80%). On the other hand, the proposed PCA-orthogonality key feature selection method (PCA-ORTH) can improve the performance of each texture description method more than the comparing feature selection methods. As a result, the proposed PCA-orthogonality local Tamura’s texture descriptor generates the highest average classification accuracy on this database (97.95%).

Figure 7 provides confusion matrices of the classification using different texture description methods. The original Tamura’s texture descriptor fails to identify the textures D1, D12, and D21. The LBP descriptor cannot distinguish the global arrangement of local structures between the textures D10 and D22. Even the “full-feature” local Tamura’s texture descriptor incorrectly considers the textures D9, D14, and D17. The PCA-orthogonality local Tamura’s texture descriptor correctly classifies most of the testing textures to their target classes, leading to the lowest confusion rate (0.97%).

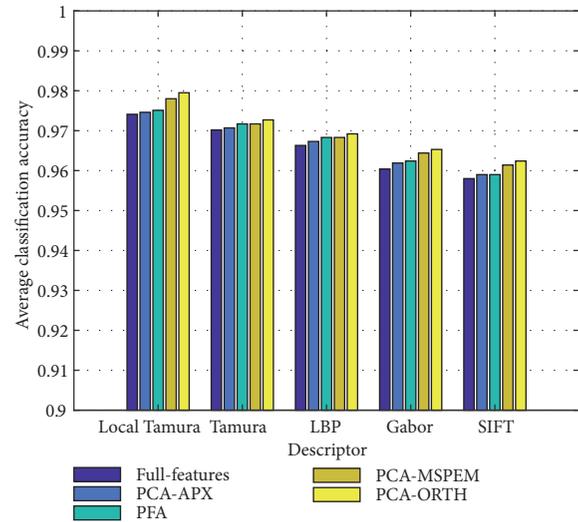


FIGURE 6: Average classification accuracy of textures in Brodatz database by different texture descriptors with different feature selection methods.

Furthermore, Table 2 illustrates the running time of feature extraction, feature selection, and classification using the selected features for a general evaluation of the efficiency of each texture descriptor. The LBP and Gabor provide shorter running time in feature extraction, but the classification time is much longer because of the high feature dimensionality, even with the feature selection step. The SIFT provides high running time in both the feature extraction and the classification step. In contrast, the two Tamura’s texture descriptors cost medium running time in feature extraction step but the shortest classification time since the feature vectors only have 6 dimensions. For the feature selection part, both PCA-MSPERM and PFA methods need long time to select features. The approximation method takes the shortest running time in selecting q key features since it needs no computation other than checking the highest absolute value of the coefficients.

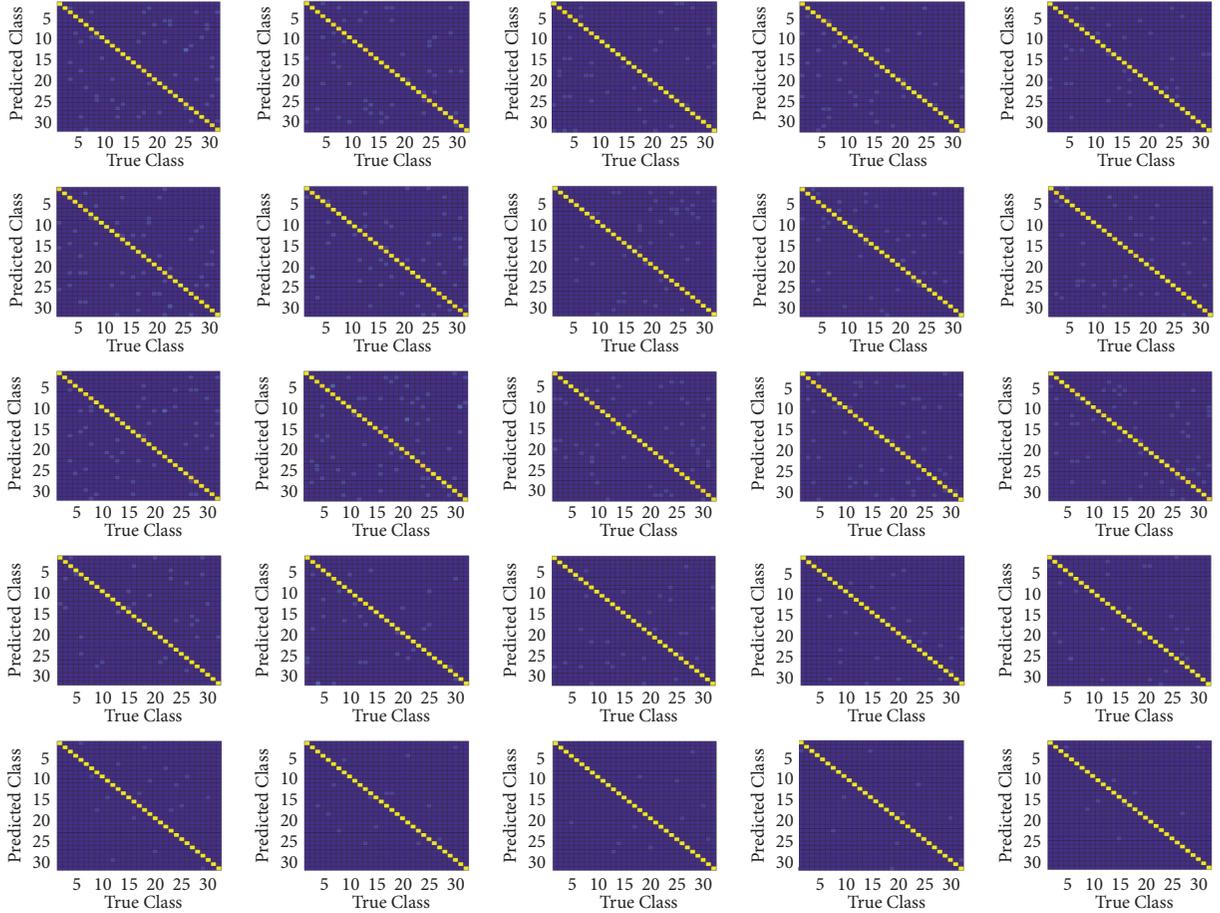


FIGURE 7: The confusion matrices of classifying textures in Brodatz database via applying different texture descriptors with different feature selector. Rows from top to bottom: confusion matrix of using LBP descriptor, Gabor descriptor, SIFT descriptor, Tamura descriptor, and the proposed local Tamura descriptor. Columns from left to right: confusion matrix of using full set of features, PCA-APX feature selection, PFA feature selection, PCA-MSPeM feature selection, and the proposed PCA-ORTH feature selection. The yellow colour represents high classification rate of the certain class of texture, while the blue colour represents low classification rate of the certain class of texture. The more yellow the elements on diagonal line are, the more blue the elements on diagonal line are, and the better the texture descriptor performs.

The proposed PCA-ORTH feature selection method achieves the running time close to the approximation method because it only repeats the calculation of cosine distance between vectors in the possible subsets of the coefficient matrix k , of which the size is not quite large for the chosen image descriptor. For example, only $C_6^2 \times C_2^2 + C_6^3 \times C_3^2 + C_6^4 \times C_4^2 + C_6^5 \times C_5^2 + C_6^6 \times C_6^2 = 240$ cosine distances need to be calculated in total for each of the two Tamura's descriptors. Although the size of the coefficient matrix could be larger in the case of other 3 descriptors, the running time of PCA-ORTH is still comparable to the approximation method considering that the calculating ability of the computer is stronger than human in seeking the variables corresponding to the highest coefficients of each of the first q principal components. Over all, the proposed method relates the low-level statistical measurements to the high-level visual characteristics locally, applies the cosine distance as the simpler dissimilarity measurement of principal components, and finds the feature vectors that are most dissimilar to each other. Therefore, the most representative features are computed with

less computational complexity than other methods and able to represent the most representative features of the original ones, leading to higher classification accuracy than the "full-feature" classification and much shorter running time in both feature selection and classification.

5.1.1. Experiments on UIUCTex Texture Database. Figure 8 illustrates how the classification accuracy and the running time of the proposed LTM+PCA-ORTH method change with different selections of the neighbourhood size. Similar as the results for the Brodatz texture database, 5 is a good trade-off as the radius of the neighbourhood: it achieves the highest classification accuracy (98.75%) with short running time (23.32 s).

Figure 9 shows the classification accuracy of applying different texture descriptors combined with different feature selection methods to the texture images in UIUCTex database. Figure 10 shows the confusion matrices of the classification using different texture description methods. Table 3 lists the running time of each step in the whole

TABLE 2: The running time of feature extraction, feature selection, classification, and total processing of classifying textures in Brodatz database with different texture descriptors and different feature selection methods.

Method	Feature extraction(s)	Feature selection(s)	Classification(s)	Total(s)
LTM	2.12	-	65.63	67.75
LTM+PCA-MSPEM	2.12	2.74	54.65	59.51
LTM+PFA	2.12	0.10	21.63	23.85
LTM+PCA-APX	2.12	0.12	20.48	22.72
LTM+PCA-ORTH	2.12	0.18	21.10	23.40
TM	3.26	-	65.63	68.89
TM+PCA-MSPEM	3.26	2.74	54.65	60.65
TM+PFA	3.26	0.10	21.23	24.59
TM+PCA-APX	3.26	0.12	20.48	23.68
TM+PCA-ORTH	3.26	0.18	21.10	24.54
LBP	0.15	-	82.27	82.42
LBP+PCA-MSPEM	0.15	5.49	68.51	74.15
LBP+PFA	0.15	0.30	27.11	27.56
LBP+PCA-APX	0.15	0.32	25.67	26.14
LBP+PCA-ORTH	0.15	0.38	26.45	26.98
SIFT	1.19	-	73.95	75.14
SIFT+PCA-MSPEM	1.19	3.81	61.58	66.58
SIFT+PFA	1.19	0.14	24.37	25.70
SIFT+PCA-APX	1.19	0.17	23.08	24.44
SIFT+PCA-ORTH	1.19	0.25	23.77	25.21
Gabor	0.32	-	75.60	75.92
Gabor+PCA-MSPEM	0.32	4.22	62.94	67.48
Gabor+PFA	0.32	0.12	24.88	25.32
Gabor+PCA-APX	0.32	0.14	23.54	24.00
Gabor+PCA-ORTH	0.32	0.24	24.28	24.84

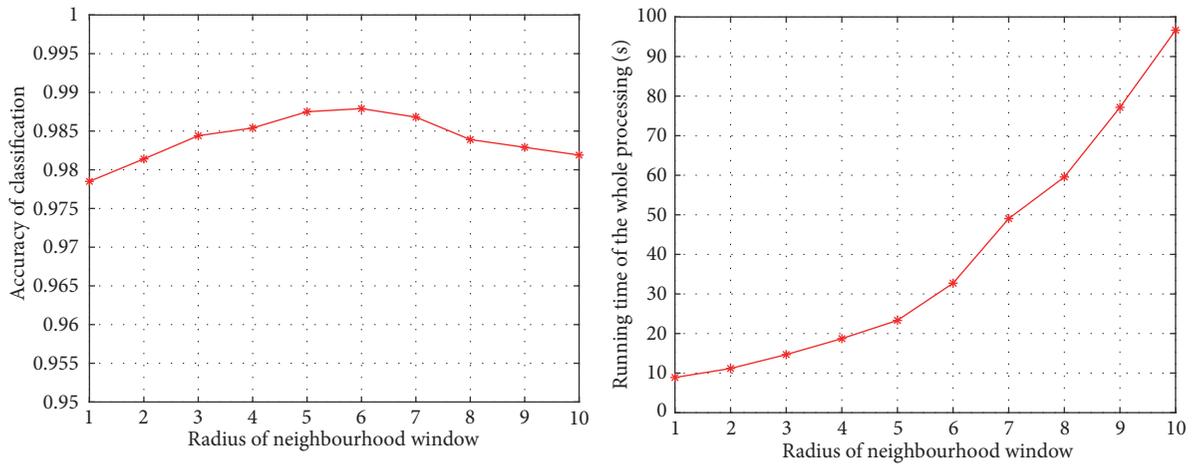


FIGURE 8: The relationship between the radius of the neighbourhood and the classification accuracy as well as the running time of the whole processing with the proposed method.

classification processing. Similar as the results on the Brodatz database, the proposed local Tamura's texture descriptor combined with the proposed PCA-ORTH feature selection method provides better classification performance than the comparators, with respect to the average classification accuracy (98.75%), confusion rate (1%) and the running time (23.32 s).

5.1.2. *Experiments on CURET Texture Database.* Figure 11 provides the classification accuracy of using the proposed feature description and selection method with different neighbourhood radius and the corresponding running time of the whole procedure. It can be considered that 5 is a good choice as neighbourhood radius, leading to highest classification rate (98.83%) in quite short running time (23.40 s).

TABLE 3: The running time of feature extraction, feature selection, classification, and total processing of classifying textures in UIUCTex database with different texture descriptors and different feature selection methods.

Method	Feature extraction(s)	Feature selection(s)	Classification(s)	Total(s)
LTM	2.08	-	65.61	67.69
LTM+PCA-MSPEM	2.08	2.72	54.62	59.42
LTM+PFA	2.08	0.09	21.16	23.33
LTM+PCA-APX	2.08	0.11	20.44	22.63
LTM+PCA-ORTH	2.08	0.17	21.07	23.32
TM	3.19	-	65.61	68.80
TM+PCA-MSPEM	3.19	2.72	54.62	60.53
TM+PFA	3.19	0.09	21.16	24.44
TM+PCA-APX	3.19	0.11	20.44	23.74
TM+PCA-ORTH	3.19	0.17	21.07	24.43
LBP	0.14	-	82.24	82.38
LBP+PCA-MSPEM	0.14	5.47	68.48	74.09
LBP+PFA	0.14	0.28	27.08	27.50
LBP+PCA-APX	0.14	0.30	25.63	26.07
LBP+PCA-ORTH	0.14	0.36	26.41	26.91
SIFT	1.16	-	73.91	75.07
SIFT+PCA-MSPEM	1.16	3.75	61.54	66.45
SIFT+PFA	1.16	0.11	24.32	25.59
SIFT+PCA-APX	1.16	0.13	23.02	24.31
SIFT+PCA-ORTH	1.16	0.22	23.72	25.10
Gabor	0.32	-	75.60	75.92
Gabor+PCA-MSPEM	0.32	4.22	62.94	67.48
Gabor+PFA	0.32	0.12	24.88	25.32
Gabor+PCA-APX	0.32	0.14	23.54	24.00
Gabor+PCA-ORTH	0.32	0.24	24.28	24.84

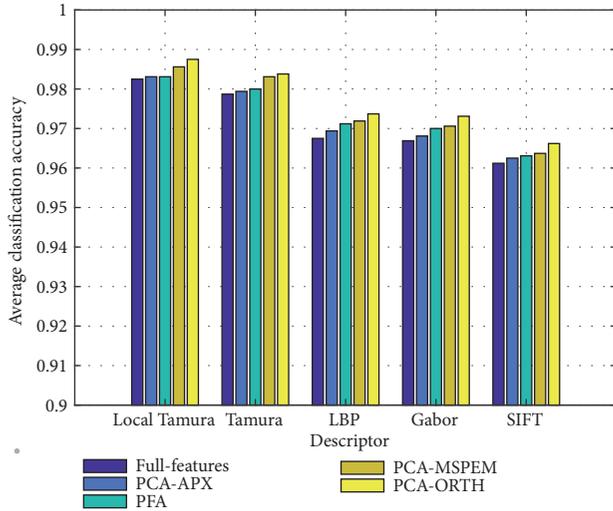


FIGURE 9: Average classification accuracy of textures in UIUCTex database by different texture descriptors with different feature selection methods.

Figure 12 shows the results of describing and classifying the texture images in the CURET database using the features extracted by different description methods. Figure 13 shows the confusion matrices of the classification using different

texture description methods. Table 4 illustrates the running time of different classification using different texture descriptors and feature selection methods. It can be considered that the proposed description algorithm, that is, combining the proposed local Tamura's descriptor and the proposed PCA-ORTH feature selection, leads to higher classification accuracy (98.83%), lower confusion rate (0.58%) and shorter running time (23.40 s).

6. Discussion

We will discuss the experimental results on the above three database with respect to the aspects of texture description and feature selection since the proposed method consists of these two parts.

The classification of textures in different datasets by the "full-feature" version of different descriptors demonstrates their performance in reflecting the differences between different textures. The LBP descriptor firstly calculates the LBP code, which reflects the intensity distribution in the local neighbourhood, but it cannot distinguish the textures with similar structure but different intensity contrast because it cannot measure the exact difference between the neighbouring pixel and the centre pixel. The SIFT descriptor needs to detect the key points in the image first, and therefore it cannot describe the textures where the edges are quite

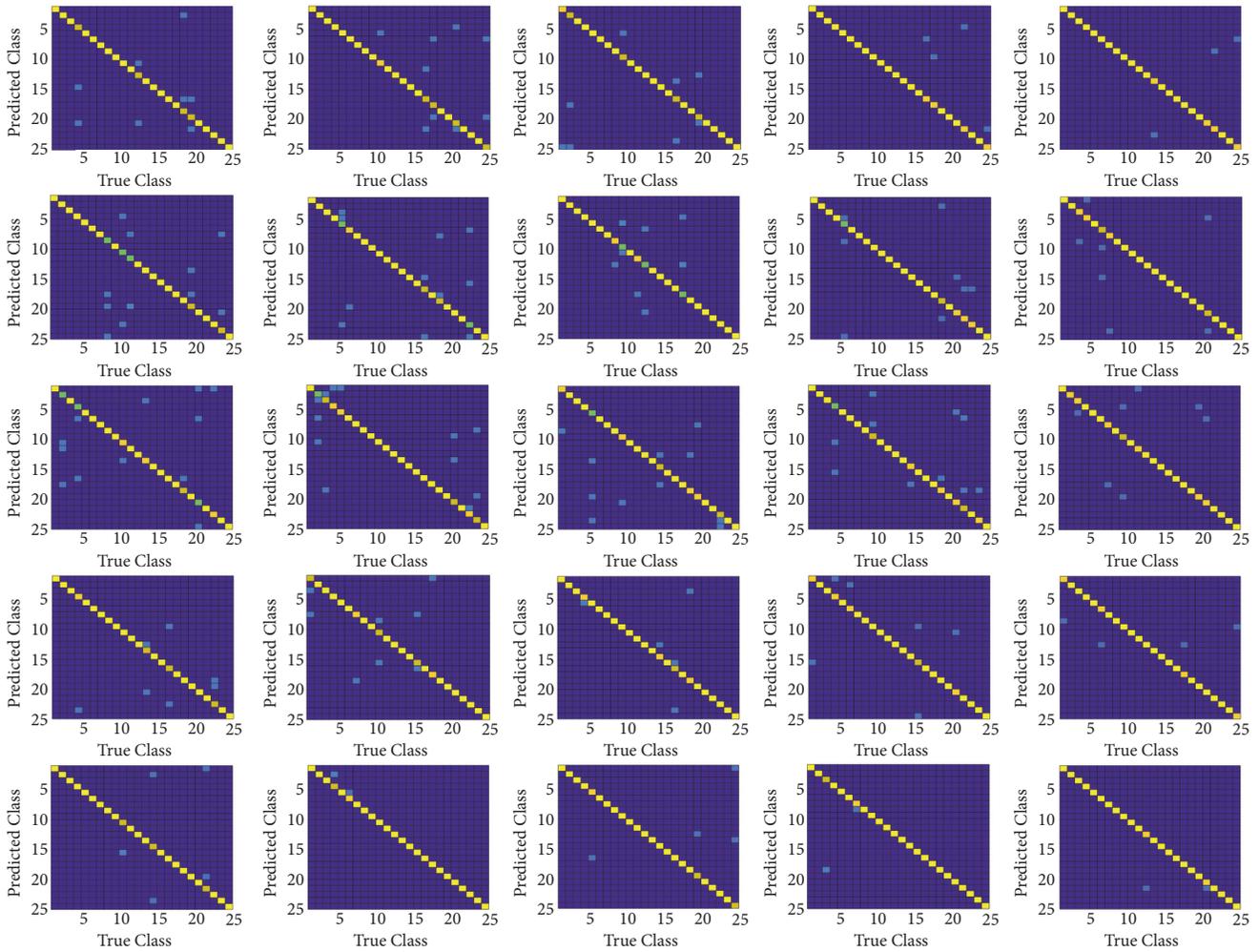


FIGURE 10: The confusion matrices of classifying textures in UIUCTex database via applying different texture descriptors with different feature selector. Rows from top to bottom: confusion matrix of using LBP descriptor, Gabor descriptor, SIFT descriptor, Tamura descriptor, and the proposed local Tamura descriptor. Columns from left to right: confusion matrix of using full set of features, PCA-APX feature selection, PCA-MSPM feature selection, and the proposed PCA-ORTH feature selection. The yellow colour represents high classification rate of the certain class of texture, while the blue colour represents low classification rate of the certain class of texture. The more yellow the elements on diagonal line are, the more blue the elements on diagonal line are, and the better the texture descriptor performs.

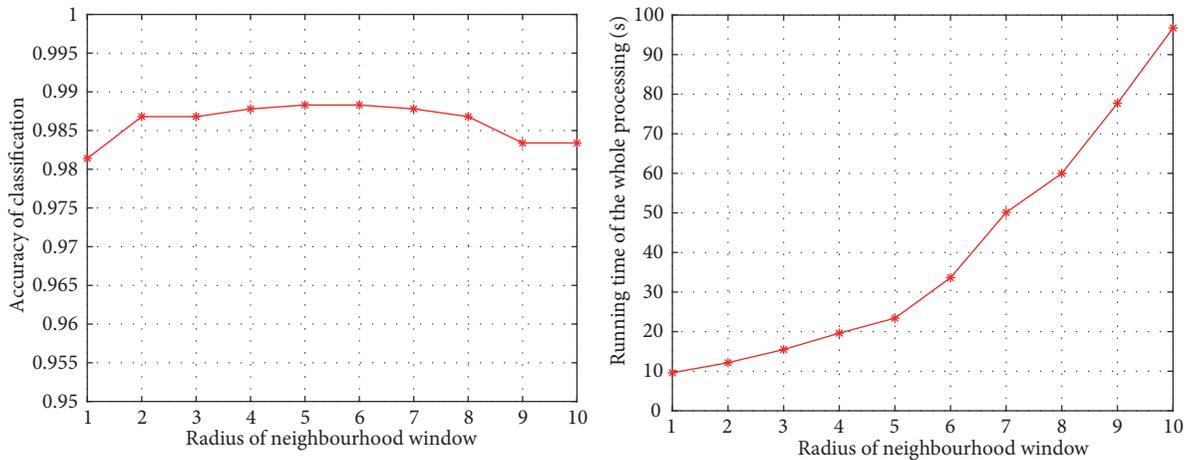


FIGURE 11: The relationship between the radius of the neighbourhood and the classification accuracy as well as the running time of the whole processing with our proposed method on CURET texture database.

TABLE 4: The running time of feature extraction, feature selection, classification, and total processing of classifying textures in CURET database with different texture descriptors and different feature selection methods.

Method	Feature extraction(s)	Feature selection(s)	Classification(s)	Total(s)
LTM	2.66	-	65.63	67.75
LTM+PCA-MSPEM	2.66	2.74	54.65	59.51
LTM+PFA	2.66	0.10	21.63	23.85
LTM+PCA-APX	2.66	0.12	20.48	22.72
LTM+PCA-ORTH	2.66	0.18	21.10	23.40
TM	4.17	-	65.63	68.89
TM+PCA-MSPEM	4.17	2.74	54.65	60.65
TM+PFA	4.17	0.10	21.63	24.59
TM+PCA-APX	4.17	0.12	20.48	23.86
TM+PCA-ORTH	4.17	0.18	21.10	24.54
LBP	0.19	-	87.48	87.67
LBP+PCA-MSPEM	0.19	6.03	70.15	76.37
LBP+PFA	0.19	0.38	29.26	29.83
LBP+PCA-APX	0.19	0.40	27.88	28.47
LBP+PCA-ORTH	0.19	0.48	25.57	26.24
SIFT	1.20	-	74.15	75.35
SIFT+PCA-MSPEM	1.20	3.82	61.62	66.64
SIFT+PFA	1.20	0.16	24.39	25.75
SIFT+PCA-APX	1.20	0.19	23.10	24.49
SIFT+PCA-ORTH	1.20	0.27	23.79	25.26
Gabor	0.34	-	75.62	75.96
Gabor+PCA-MSPEM	0.34	4.23	62.98	67.55
Gabor+PFA	0.34	0.16	24.88	25.38
Gabor+PCA-APX	0.34	0.18	23.58	24.10
Gabor+PCA-ORTH	0.34	0.29	24.31	24.94

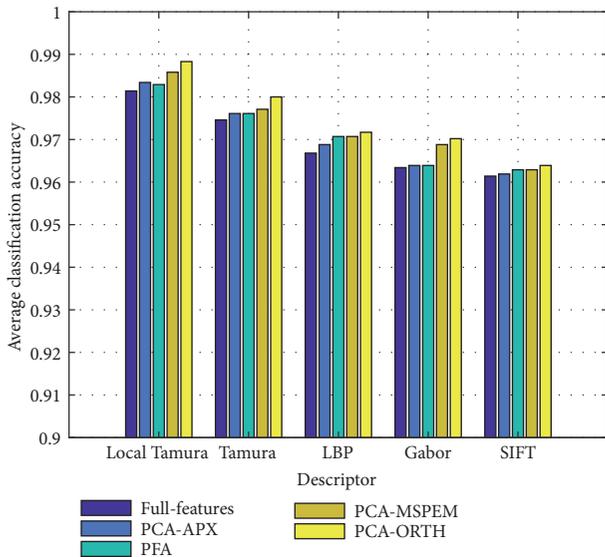


FIGURE 12: Average classification accuracy of textures in CURET database by different texture descriptors with different feature selection methods.

smooth or the patterns changed gradually and it is difficult to find the key points. The Gabor descriptor makes use of the Gabor wavelet coefficients as the texture features, so the

performance depends on the Gabor transform of the given texture image. The Gabor-based description may fail when the scale or the size of the Gabor window is not selected suitably. The LBP, SIFT, and Gabor descriptors focus on the local features too much, lacking in giving significant attention to explore the relationship between the texture characteristics they described and human visual sense globally. Therefore, they cannot differentiate those textures with tiny differences in visual system but similar edges or intensity distribution. Moreover, there is much redundant information in LBP, SIFT, and Gabor-based description, which increases the computational complexity, leading to longer running time than the Tamura (TM) and local Tamura (LTM) descriptors. The Tamura descriptor, as a high-level descriptor, pays more attention on the textural differences with respect to human visual perception, leading to higher classification accuracy. However, when applying the original Tamura's definitions for the six characteristics directly to local neighbourhood, it lacks in measuring the local contrast, rotation, illuminance, and scale changes, resulting in difficulty in distinguishing the textures with similar appearances but different local details. The proposed local Tamura description defines the six Tamura's characteristics using the local information in each neighbourhood directly. It correctly bridge the "semantic gap" between the low-level measurements and the high-level visual perceptual characteristics, leading to accurate

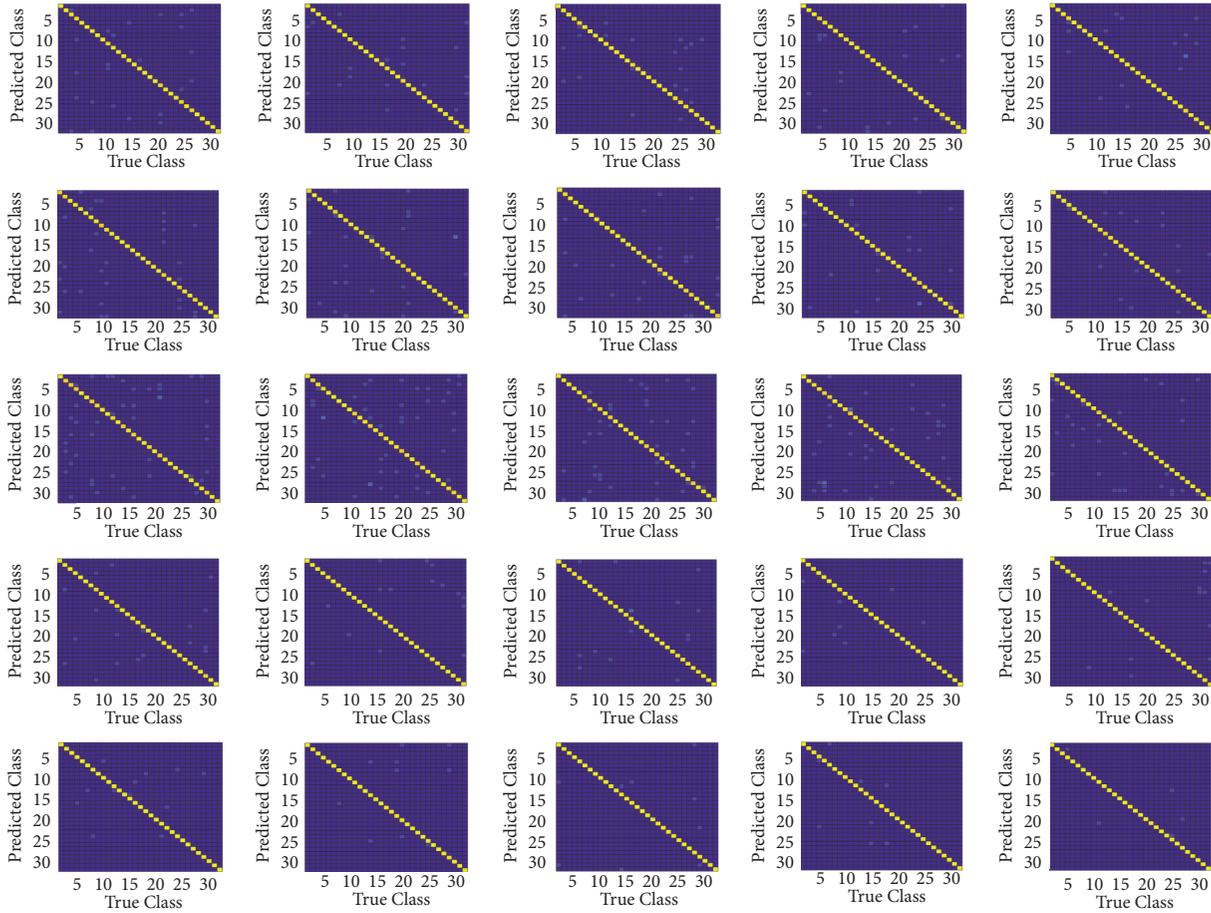


FIGURE 13: The confusion matrices of classifying textures in CURET database via applying different texture descriptors with different feature selector. Rows from top to bottom: confusion matrix of using LBP descriptor, Gabor descriptor, SIFT descriptor, Tamura descriptor, and the proposed local Tamura descriptor. Columns from left to right: confusion matrix of using full set of features, PCA-APX feature selection, PFA feature selection, PCA MSPEM feature selection, and the proposed PCA-ORTH feature selection. The yellow colour represents high classification rate of the certain class of texture, while the blue colour represents low classification rate of the certain class of texture. The more yellow the elements on diagonal line are, the more blue the elements on diagonal line are, and the better the texture descriptor performs.

description of local textural region with respect to human visual perception and accurate classification of both intraclass and interclass textures.

In the feature selection part, the PCA-MSPERM method preserves 4 or 5 of the 6 input features so the redundancy is still possible to exist. And the computational complexity of finding the representative features is high because the mean square prediction error is calculated for all possible combinations of the q selected features. The approximation based on the magnitudes of coefficients of the principal components is a very intuitive and computationally feasible method because it requires no computation other than checking the highest absolute value of the coefficients. But the number of the selected features is the same as that of the principal components, losing some representative features; therefore the performance of classifying textures with these features varies a lot over the whole database. The PFA method achieves good performance in selecting the features that are largely spread in the lower dimensional space and good

representation of the original data. However, it requires the prior parameter K to be set so it is not adaptive in selecting the features. Moreover, the clustering method is high in computational cost when the number of coefficients is large. The proposed PCA-ORTH feature selection method applies the cosine distance as the simpler dissimilarity measurement and finds the feature vectors that are most dissimilar to each other. Therefore, the most representative features are computed with less computational complexity than other methods and able to select the most representative features of the original ones, leading to better classification performance and shorter running time in both feature selection and classification.

The experimental results on the above three database proves that the proposed local Tamura's texture descriptor can correctly bridge the "semantic gap" between the low-level measurements and the high-level visual perceptual characteristics. Moreover, the proposed PCA-ORTH feature selection method can find the key features that reflect the essential

differences between textures and avoid the redundancy in description, reducing the confusion of intraclass textures and the computational complexity of the following classification.

7. Conclusions

In this paper, we proposed a local Tamura's texture description and a PCA-ORTH feature selection scheme. Texture images can be locally described by the high-level characteristics that fit the human visual perception, including coarseness, contrast, directionality, line-likeness, regularity, and roughness. Then a PCA-ORTH feature selection scheme was applied on a given database to find the most representative features that can differentiate the textures with low redundancy. Experiments were implemented to compare the performance of using the proposed description methods and the others in classifying the textures in different databases. The experimental results demonstrate that the local Tamura's texture features can describe different textures in the database accurately and the proposed feature selection method can effectively find the features that are essential in differentiating the textures. As a result, the differences between different textures can be measured more precisely with respect to these key texture characteristics and different textures can be classified more accurately.

Data Availability

All of the data are available. The links are listed as follows: Brodatz texture database: <http://multibandtexture.recherche.usherbrooke.ca/original.brodatz.html> and UIUCTex texture database: http://slazebni.cs.illinois.edu/research/uiuc_texture_dataset.zip CURET texture database: <http://www.cs.columbia.edu/CAVE/software/curet/>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grants 61701101, U1713216, 61803077, and 61603080, the National Key Robot Project Grant 2017YFB1300900, and the Fundamental Research Funds for the Central Universities (N172603001 and N172604004). The authors would like to thank Ying Wang and Wei Zhou (Northeastern University, China) for their help and discussions.

References

- [1] X. Hong, G. Zhao, M. Pietikainen, and X. Chen, "Combining LBP difference and feature correlation for texture description," *IEEE Transactions on Image Processing*, vol. 23, no. 6, pp. 2557–2568, 2014.
- [2] F. Florea, E. Barbu, A. Rogozan, A. Bensrhair, and V. Buzuloiu, "Medical image categorization using a texture based symbolic description," in *Proceedings of the 2006 IEEE International Conference on Image Processing, ICIP 2006*, pp. 1489–1492, USA, October 2006.
- [3] Y. Lv, J. Liu, and W. Zeng, "A method of cattle follicle ultrasound images detection based on HOG + improved LBP + SVM," in *Proceedings of the 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 2176–2181, Wuhan, China, May 2018.
- [4] W. Tan, Y. Yuan, A. Chen, L. Mao, Y. Ke, and X. Lv, "An approach for pulmonary vascular extraction from chest CT images," *Journal of Healthcare Engineering*, vol. 2019, 11 pages, 2019.
- [5] D. T. Nguyen, Z. Zong, P. Ogunbona, and W. Li, "Object detection using non-redundant local binary patterns," in *Proceedings of the 2010 17th IEEE International Conference on Image Processing, ICIP 2010*, pp. 4609–4612, Hong Kong, September 2010.
- [6] Z. Wang, B. Fan, G. Wang, and F. Wu, "Exploring Local and Overall Ordinal Information for Robust Feature Description," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 11, pp. 2198–2211, 2016.
- [7] L. Wan, L. Zheng, H. Huo, and T. Fang, "Affine invariant description and large-margin dimensionality reduction for target detection in optical remote sensing images," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 7, pp. 1116–1120, 2017.
- [8] T. Song and H. Li, "WaveLBP based hierarchical features for image classification," *Pattern Recognition Letters*, vol. 34, no. 12, pp. 1323–1328, 2013.
- [9] F. Liu, B. Wang, and F. Zeng, "Trademark image retrieval using hierarchical region feature description," in *Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP)*, pp. 3620–3624, Beijing, China, September 2017.
- [10] S. R. Dubey, S. K. Singh, and R. K. Singh, "Local neighbourhood-based robust colour occurrence descriptor for colour image retrieval," *IET Image Processing*, vol. 9, no. 7, pp. 578–586, 2015.
- [11] H. Lategahn, S. Gross, T. Stehle, and T. Aach, "Texture classification by modeling joint distributions of local patterns with Gaussian mixtures," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1548–1557, 2010.
- [12] J. Chen, S. Shan, C. He et al., "WLD: a robust local image descriptor," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1705–1720, 2010.
- [13] M. Crosier and L. D. Griffin, "Using basic image features for texture classification," *International Journal of Computer Vision*, vol. 88, no. 3, pp. 447–460, 2010.
- [14] H. Ji, X. Yang, H. Ling, and Y. Xu, "Wavelet domain multifractal analysis for static and dynamic texture classification," *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 286–299, 2013.
- [15] K. Deguchi and I. Morishita, "Texture characterization and texture-based image partitioning using two-dimensional linear estimation techniques," *IEEE Transactions on Computers*, vol. 27, no. 8, pp. 739–745, 1978.
- [16] G. R. Cross and A. K. Jain, "Markov random field texture models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 5, no. 1, pp. 25–39, 1983.
- [17] F. Tomita and S. Tsuji, *Computer Analysis of Visual Textures*, Kluwer Academic Publishers, 1990.
- [18] Z.-C. Huang, P. P. K. Chan, W. W. Y. Ng, and D. S. Yeung, "Content-based image retrieval using color moment and gabor

- texture feature,” in *Proceedings of the 2010 International Conference on Machine Learning and Cybernetics, ICMLC 2010*, pp. 719–724, China, July 2010.
- [19] T. Ojala, M. Pietikäinen, and T. Mäenpää, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.
- [20] S. R. Dubey, S. K. Singh, and R. K. Singh, “Multichannel decoded local binary patterns for content-based image retrieval,” *IEEE Transactions on Image Processing*, vol. 25, no. 9, pp. 4018–4032, 2016.
- [21] X. Zhao, Y. Lin, and J. Heikkilä, “Dynamic texture recognition using volume local binary count patterns with an application to 2D face spoofing detection,” *IEEE Transactions on Multimedia*, vol. 20, no. 3, pp. 552–566, 2018.
- [22] S. Jia, B. Deng, J. Zhu, X. Jia, and Q. Li, “Local Binary Pattern-Based Hyperspectral Image Classification with Superpixel Guidance,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 2, pp. 749–759, 2018.
- [23] Q. Wang, B. Li, X. Chen, J. Luo, and Y. Hou, “Random Sampling Local Binary Pattern Encoding Based on Gaussian Distribution,” *IEEE Signal Processing Letters*, vol. 24, no. 9, pp. 1358–1362, 2017.
- [24] H. Fan, P. C. Cosman, Y. Hou, and B. Li, “High-Speed Railway Fastener Detection Based on a Line Local Binary Pattern,” *IEEE Signal Processing Letters*, vol. 25, no. 6, pp. 788–792, 2018.
- [25] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural features for image classification,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, no. 6, pp. 610–621, 1973.
- [26] K. I. Laws, “Rapid Texture Identification,” *Processing SPIE 0238, Image Processing Integration For Missile Guidance*, vol. 238, pp. 376–381, 1980.
- [27] K. I. Laws, “Texture energy measures,” in *Proceedings of the Image Understanding Workshop*, pp. 47–51, University of Southern California Los Angeles Image Processing Institution, 1979.
- [28] R. N. Sutton and E. L. Hall, “Texture measures for automatic classification of pulmonary disease,” *IEEE Transactions on Computers*, vol. 21, no. 7, pp. 667–676, 1972.
- [29] M. Amadasun and R. King, “Textural features corresponding to textural properties,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, no. 5, pp. 1264–1274, 1989.
- [30] A. R. Rao and G. L. Lohse, “Towards a texture naming system: Identifying relevant dimensions of texture,” in *Proceedings of the 4th Annual IEEE Conference on Visualization, VIS 1993*, pp. 220–227, USA, October 1993.
- [31] K. Fujii and T. Arikawa, “Urban object reconstruction using airborne laser elevation image and aerial image,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 10, pp. 2234–2240, 2002.
- [32] H. Tamura, S. Mori, and T. Yamawaki, “Textural features corresponding to visual perception,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 8, no. 6, pp. 460–473, 1978.
- [33] L. Feng and B. Bhanu, “Semantic concept co-occurrence patterns for image annotation and retrieval,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 4, pp. 785–799, 2016.
- [34] S. T. Roweis and L. K. Saul, “Nonlinear dimensionality reduction by locally linear embedding,” *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [35] P. Pudil and J. Novovicova, “Novel methods for feature subset selection with respect to problem knowledge,” in *Feature Extraction, Construction and Selection*, pp. 101–116, IEEE Educational Activities Department, 1998.
- [36] I. T. Jolliffe, *Principal Component Analysis*, Wiley Online Library, New York, NY, USA, 2nd edition, 2002.
- [37] A. M. Martinez and A. C. Kak, “PCA versus LDA,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 228–233, 2001.
- [38] H. Li, T. Jiang, and K. Zhang, “Efficient and robust feature extraction by maximum margin criterion,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 17, no. 1, pp. 157–165, 2006.
- [39] J. B. Tenenbaum, V. de Silva, and J. C. Langford, “A global geometric framework for nonlinear dimensionality reduction,” *Science*, vol. 290, no. 5500, pp. 2319–2323, 2000.
- [40] P. J. G. Lisboa and A. R. Mehri-Dehnavi, “Sensitivity methods for variable selection using the MLP,” in *Proceedings of the 1996 1st International Workshop on Neural Networks for Identification, Control, Robotics, and Signal/Image Processing, NICROSP’96*, pp. 330–338, August 1996.
- [41] J. Gui, Z. Sun, S. Ji, D. Tao, and T. Tan, “Feature selection based on structured sparsity: a comprehensive study,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 7, pp. 1490–1507, 2017.
- [42] B. Xue, M. Zhang, W. N. Browne, and X. Yao, “A Survey on Evolutionary Computation Approaches to Feature Selection,” *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 4, pp. 606–626, 2016.
- [43] X. Zhu, X. Li, S. Zhang, C. Ju, and X. Wu, “Robust joint graph sparse coding for unsupervised spectral feature selection,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 6, pp. 1263–1275, 2017.
- [44] Y. Ke and R. Sukthankar, “PCA-SIFT: a more distinctive representation for local image descriptors,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR ’04)*, vol. 2, pp. 506–513, July 2004.
- [45] S. B. Alam, R. Nakano, S. Kobashi, and N. Kamiura, “Feature selection of manifold learning using principal component analysis in brain MR image,” in *Proceedings of the 4th International Conference on Informatics, Electronics and Vision, ICIEV 2015*, pp. 1–5, Fukuoka, Japan, June 2015.
- [46] C. D. Sigg and J. M. Buhmann, “Expectation-maximization for sparse and non-negative PCA,” in *Proceedings of the 25th International Conference on Machine Learning*, pp. 960–967, ACM, New York, NY, USA, July 2008.
- [47] I. Cohen, Q. Tian, X. Zhou, and T. S. Huang, *Feature Selection and Dimensionality Reduction Using Principal Feature Analysis*, University of Illinois Image Formation and Processing group, 2015.
- [48] T. Majtner and D. Svoboda, “Extension of tamura texture features for 3D fluorescence microscopy,” in *Proceedings of the 2nd Joint 3DIM/3DPVT Conference: 3D Imaging, Modeling, Processing, Visualization and Transmission, 3DIMPVT 2012*, pp. 301–307, Switzerland, October 2012.
- [49] H. Xu, C. Lu, R. Berendt, N. Jha, and M. Mandal, “Automatic Nuclei Detection Based on Generalized Laplacian of Gaussian Filters,” *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 3, pp. 826–837, 2017.
- [50] J. Yezep and S. Ko, “Improved license plate localisation algorithm based on morphological operations,” *IET Intelligent Transport Systems*, vol. 12, no. 6, pp. 542–549, 2018.

- [51] Y. Zhang, X. Han, H. Zhang, and L. Zhao, "Edge detection algorithm of image fusion based on improved Sobel operator," in *IEEE 3rd Information Technology and Mechatronics Engineering Conference (ITOEC)*, pp. 457–461, Chongqing, China, 2017.
- [52] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, no. 4, pp. 433–459, 2010.
- [53] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: a review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 1798–1828, 2013.
- [54] Y. Luo, C.-M. Wu, and Y. Zhang, "Facial expression recognition based on fusion feature of PCA and LBP with SVM," *Optik - International Journal for Light and Electron Optics*, vol. 124, no. 17, pp. 2767–2770, 2013.
- [55] P. Brodatz, *A Photographic Album for Arts and Design*, Dover Publishing Company, Toronto, Canada, 1966.
- [56] S. Lazebnik, C. Schmid, and J. Ponce, "A sparse texture representation using local affine regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 8, pp. 1265–1278, 2005.
- [57] K. J. Dana, B. Van Ginneken, S. K. Nayar, and J. J. Koenderink, "Reflectance and texture of real-world surfaces," *ACM Transactions on Graphics*, vol. 18, pp. 1–34, 1999.
- [58] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [59] M. Abdulrahman, T. R. Gwadabe, F. J. Abdu, and A. Eleyan, "Gabor wavelet transform based facial expression recognition using PCA and LBP," in *Proceedings of the 2014 22nd Signal Processing and Communications Applications Conference, SIU 2014*, pp. 2265–2268, Turkey, April 2014.
- [60] D. G. Altman and J. M. Bland, "Diagnostic tests 1: sensitivity and specificity," *British Medical Journal*, vol. 308, no. 6943, p. 1552, 1994.

