International Journal of Innovative Computing, Information and Control
Volume **16**, Number **5**, October **2020**

ICIC International ©2020 ISSN 1349-4198
pp. 1611–1621

LOCAL SPATIAL INFORMATION WITH BAG-OF-VISUAL-WORDS MODEL VIA GRAPH-BASED REPRESENTATION FOR TEXTURE CLASSIFICATION

Srisupang Thewsuwan[1] and Keiichi Horio[2]

[1]Department of Computer Engineering
Srinakharinwirot University
Ongkharak, Nakhon Nayok 26120, Thailand
srisupang@g.swu.ac.th

[2]Graduate School of Life Science and Systems Engineering
Kyushu Institute of Technology
Wakamatsu, Kitakyushu 808-0796, Japan
horio@brain.kyutech.ac.jp

Received May 2020; revised August 2020

Abstract. *This paper proposes an enhanced feature descriptor for texture classification through graph-based representation. Searching the meaningful texture descriptor is a crucial process in pattern analysis and applications. Graph theory is a model-based approach that applies to texture analysis with outstanding results. Therefore, to develop feature descriptors that are robust against many variations images collected from random viewpoints, change in scale, and illumination remains a challenge for researchers. In this work, we propose an Automatically Local Spatial Pattern Mapping (LSPM^{Auto}) method based on the spatial-BoVW model that can extract local and global features information from the spatial arrangement of image pixels. The proposed approach is evaluated by using three different texture databases: Brodatz, UIUC, and Outex. The experimental results show that the proposed method can achieve highly discriminant descriptors superior to the other methods.*

Keywords: Spatial-BoVW, Graph theory, Feature descriptor, Texture analysis, LSPM^{Auto}

1. **Introduction.** Texture can be used to characterize the appearance of objects in the range of coarseness and smoothness caused by the uniformity of the image tone and color, representing an image appearance [1, 2]. In deformable objects with material properties, such as clothes, a capable texture descriptor is an essential part of clothing classification in robotic vision applications [3, 4]. The natural texture, such as fabric, has random and stochastic patterns, so it is still challenging to compose such patterns from their models. The capable techniques that can extract the texture pattern, for example, use first-order features [2], Fourier descriptors [5], Gabor filters [6], graph theory [7, 8], and complex network theory [9, 10]. Therefore, texture descriptor enhancement is focused on this research.

In recent years, image representation by graph theory has been introduced as a model-based approach for texture analysis [7, 11]. Backes et al. [11] have proposed an efficient complex network model using graph theory [7, 9, 10] to represent image pixels for texture analysis and classification with outstanding results. The coarseness and orientation of an image structure are described regarding the network's topological properties, which aids in discrimination. Although the complex network model was capable of texture analysis

DOI: 10.24507/ijicic.16.05.1611

and classification, the network's statistical properties discard fundamental properties such as a spatial arrangement. The related researchers have developed feature descriptors by combining the BoVW model and Complex Networks (BoVW-CN) [12, 13] for texture classification. They proposed a new methodology describing the relevant points of an image based on BoVW and a complex network model that can obtain good accuracy. In previous work [14], a texture extractor with a spatial-analysis approach was proposed using the complex network model. Specifically, Local Spatial Pattern Mapping (LSPM) was constructed to capture the pixel arrangement on an adopted complex network model. The results have been proven to be useful for improving performances in texture classification compared with the traditional complex network model and other methods. However, the method was lack of capability for extracting discriminative information, which was used to distinguish various pattern structures with invariant in an uncontrolled environment. Therefore, it is a challenge to achieve higher texture discriminative performance. Bag-of-Visual-Words (BoVW) model is widely used in the feature representation scheme and image classification [15, 16]. The method aims to collect local spatial pixels in the images and appearance of images as global features. To overcome the limitation of previous work, we combined the graph-based representation and the BoVW-based spatial feature into texture classification.

In this paper, we propose an enhanced feature descriptor for improving texture classification performance based on graph-based image representation. The spatial-BoVW model is employed to identify spatial information of weighted graphs for characterizing the textural pattern. We combine LSPM^{Auto} with the BoVW model as a feature extractor for texture classification. The LSPM^{Auto} extends the previous work, LSPM method, by using adaptive threshold values to construct the graph in texture images. The LSPM^{Auto} and BoVW aim to seek local and global spatial feature extractors, which are adaptability, and methodically for texture analysis. Three benchmark texture databases including – Brodatz [17], UIUC [18], and Outex [19] – are applied for the proposed system evaluation. The experimental results show that the proposed method achieves superior performance at the texture classification compared with the original approaches and conventional texture analysis methods.

To describe the proposed method, this paper is organized as follows. Section 2 introduces the proposed algorithm that can be separated into three parts: pixel-based representation by graph theory, multiscale region-level mapping, and feature descriptors. Section 3 describes the experiments and results, including detailed discussions. Finally, the conclusion presents in Section 4.

2. **Proposed Algorithm.** The overview of the proposed algorithm is illustrated by Figure 1. The system can be separated into three parts. The first part is the graph-based image representation process. The weight of edges is applied by graph theory in this step. This parameter represents pairwise connections between a node and neighbors, which can be used for representing a local textural structure. The edge graph can be constructed by using the Euclidean distance between a node and pairwise connections. Then, the set of weighted graph values passes through the next process, the multiscale region-level mapping. We can increase the texture pattern from pixel connectivity into three patterns based on their Euclidean distance. After that, the weight-row recorders are presented to represent the set of feature vectors. Finally, the last part of the proposed system is the feature descriptors. The region-neighbor sets are local spatial features based on graph-based representation. The feature descriptors consist of local and global features. The local feature is denoted by the LSPM^{Auto} method. On the other hand, the BoVW

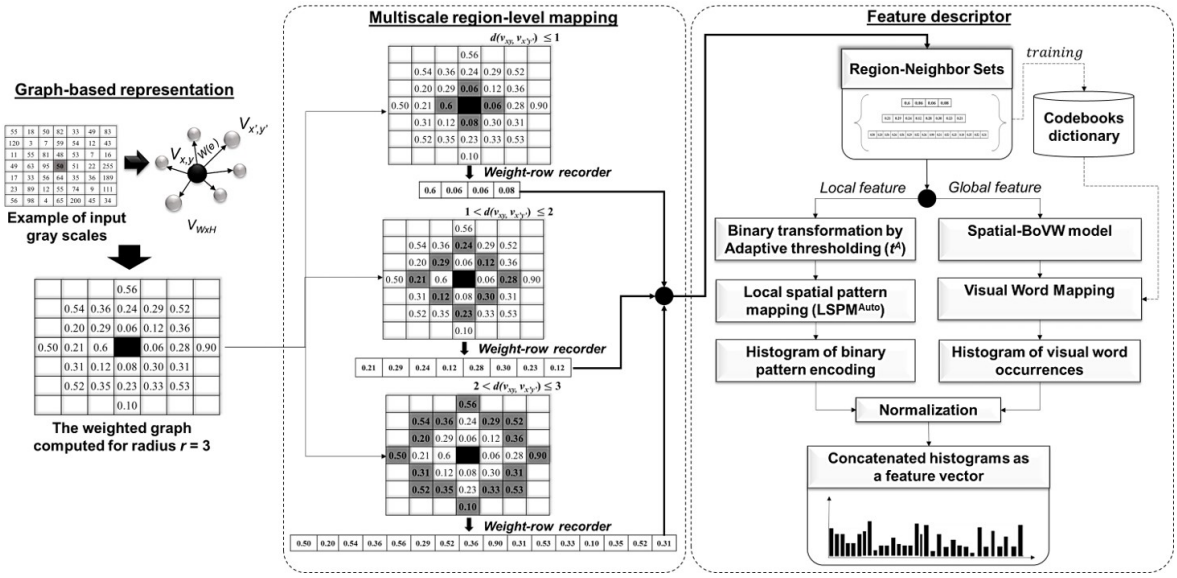

FIGURE 1. Overview of the proposed system

model is used for extracting global features. Then, the concatenated histograms are the final feature vector from the proposed system.

2.1. **Pixel-based representation by graph theory.** In this work, graph theory [9,10] is employed in modeling each pixel of an image as a graph. The image I can be defined with the resolution of *width* $(W) \times$ *height* (H) pixels. The graph $G = (V, E)$ generated from a given image, where $V = \{v_1, \ldots, v_{M\times N}\}$, is set of vertices, and set of edges E, where $E = \{e_1, \ldots, e_{M\times N}\}$. The edges of graph are created if the Euclidean distance between two pixels, $v(x, y) \in V$ and $v(x', y') \in V$ is less than or equal to radius r, then the two vertices are connected by an edge as defined by Equation (1). Regarding Equation (1), the vertex connects with its neighbors within the radial distance r. The maximum value of radius r is equal to 3, as recommended by Backes et al.'s work [11].

$$E = \left\{(x, y), (x', y') \in W \times H \,\middle|\, \sqrt{(x - x')^2 + (y - y')^2} \leq r \right\} \tag{1}$$

The weight of edge $W(e)$ connects with each edge $e = (v_{xy}, v_{x'y'}) \in E$ from the graph G. The weight of $W(e)$ is determined via connectivity of two pixels representing an edge e if the Euclidean distance less than or equal to radius r, their respective intensities, and maximum intensity value L in the image I, as proposed by [11,14].

$$W(e) = \begin{cases} \dfrac{(x - x')^2 + (y - y')^2 + r^2 \dfrac{|I(x, y) - I(x', y')|}{L}}{r^2 + r^2} & \text{if } dist(v_{xy}, v_{x'y'}) \leq r \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

2.2. **Multiscale region-level mapping.** The multiscale region-level mapping approach is applied to capturing the context of region-neighbors. Figure 1 shows an example of a vertex's spatial attributes regarding the range of Euclidean distance values. The process consists of three scales of fixed spatial-pixel arrangements. Then, the weighted graph values are collected by a weight-row recorder as an attribute mapping. We explain this

step using this equation

$$W_p^r(e) = \begin{cases} W_4^1(e) & \text{if } dist(v_{xy}, v_{x'y'}) \leq 1 \\ W_8^2(e) & \text{if } 1 < dist(v_{xy}, v_{x'y'}) \leq 2 \\ W_{16}^3(e) & \text{if } 2 < dist(v_{xy}, v_{x'y'}) \leq 3 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

From Equation (3), $W_p^r(e)$ is region-neighbor sets from the multiscale region-level mapping, where $(v_{xy}, v_{x'y'}) \in V$ is the set of vertices $\{v_1, v_2, \ldots, v_{M\times N}\}$ where r is the radius of each scale equal to 1, 2 and 3 and its neighbors p equal to 4, 8 and 16, respectively as regarding our previous work [14].

2.3. **Feature descriptors.** The feature descriptor of this work consists of the local and global feature extractors. The LSPM^{Auto} method is proposed to describe the local spatial feature as detailed in Algorithm 1. The spatial-BoVW model can be illustrated in Algorithm 2. This method is used to capture a global feature of pixels in images. We describe in detail in the following subsections.

2.3.1. *LSPM^{Auto} method.* This method is extended from the previous work, LSPM method [14], for encoding the spatial arrangement of image pixels. Regarding the complex network model properties, we had created the desired network that simulates pixel dynamic transformation in texture analysis by providing a set of thresholds. However, the network requires tuning the threshold parameters found through experiments. To overcome this limitation, an adaptive thresholding value is needed. Otsu's global threshold clustering method is employed to perform as automatic thresholding. This method seeks a threshold parameter that minimizes the intra-class variance [20]. It is defined by a weighted sum of variances between two vertices. In this binary pattern transformation process $WB_p^r(e)$, we apply an adaptive threshold t^A to the original set of edges E. It enables us to determine the attributes of spatial neighbors. Equation (4) explains the edge weight computation

Algorithm 1: LSPM^{Auto} method

Input: Region-Neighbor Sets: $W_p^r(e)$
Data: gray scale image I; $G = (V, E)$; radius of each scale r, $(r = 1, 2, 3)$ which corresponds to number of neighbors p, $(p = 4, 8, 16)$; adaptive threshold value t^A
Result: LSPM^{Auto} feature: H

```
begin
    for each of scale (r, p) where W(e) ∈ E do
        W_p^r(e) ← the region-neighbor sets by Equation (3)
        WB_p^r(e) ← the binary pattern transformation
        lspm_p^r(e) ← the uniformity of local binary pattern
        LSPM^Auto(r, p) ← the local spatial pattern mapping according to Equation (6)
        h_p^r ← the histogram of each scale
    end
    H ← {h_4^1, h_8^2, h_16^3}
end
return H
```

Algorithm 2: Spatial-BoVW model

Input: Region-Neighbor Sets: $W' = [W_4^1(e), W_8^2(e), W_{16}^3(e)]$
Data: training dataset β'; $G = (V, E)$; vocabulary of visual word VW; number of visual words K
Result: Spatial-BoVW feature: F

```
begin
    W' ← randomly choose the set of weighted graph from β'
    VW ← compute vocabulary of visual words from W' according to Equation (8)
    for each of number of vocabulary h(w_i), i = 1, 2, ..., K do
        h(w_i) ← represent the visual appearance using histogram
    end
    F ← set of feature vector from Equation (10), h(w)
end
return F
```

when it is less than or equal to the threshold t^A, is converted to 1, and otherwise to 0.

$$WB_p^r(e) = \begin{cases} 1 & \text{if } W_p^r(e) \leq t^A \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

The idea of the uniformity of a Local Binary Pattern (LBP) mapping [21,22] is adapted to this method. The method encodes the spatial arrangement of pixels for considering the uniformity local pattern, which we can define as follows:

$$lspm_p^r(e) = \sum_{i=1}^{p} WB_p^r(e) 2^{i-1} \tag{5}$$

where i is a neighbor of a vertex. The uniformity of $lspm$ can be defined as the following:

$$LSPM^{Auto}(r, p) = \begin{cases} lspm_p^r(e) & \text{if } U\left(lspm_p^r(e)\right) \leq 2 \\ p + 1 & \text{otherwise} \end{cases} \tag{6}$$

where U (bit-wise number transitions between 0 and 1) is used to determine whether a pattern is a uniform, which is defined as when U is at most two, for example, 00010000 (2 transitions) is uniform, whereas 01010100 (6 transitions) is non-uniform as justified by [21]. If $lspm(r, p)$ is such that $U \leq 2$, then the pattern is defined as a uniform, which can obtain the uniformity of a textured pattern or unique pattern each vertex. In practice, the mapping of LSPM^{Auto} is implemented with the reading value of template from a lookup table of 2^p elements, so there are $p + 2$ output bits. As a result, the set of feature descriptors for each radial distance is given by

$$H = \left\{h_4^1, h_8^2, h_{16}^3\right\} \tag{7}$$

2.3.2. *Spatial-BoVW model.* The BoVW model can be applied to image classification in the field of computer vision [12,15,16]. A BoVW is a vector that represents the frequency of a vocabulary. In this paper, the set weights of edges $W_p^r(e)$ in Equation (3) within training images (β') are used to create a codebook. For each training set, the k-mean clustering algorithm is approached to generate data vectors. The number of centers is denoted by K. The codebook generation is described as follows:

$$VW(K) = \underset{c}{\operatorname{argmin}} \sum_{k=1}^{K} \sum_{j=1}^{\sum p} \left\|W_j' - c_k\right\|^2 \tag{8}$$

where $k = 1, \ldots, K$ indexes clusters of Visual Words (VW). The data vectors are denoted by W' and obtained from Equation (3) as $[W_4^1(e), W_8^2(e), W_{16}^3(e)]$ from training dataset β'. The last step is representing visual appearance. We can calculate appearance by the following:

$$h(w_i) = \text{argmin}\, \|W - VW(K)\| \tag{9}$$

The set of the occurrence of VW properties based on spatial-BoVW model is shown below:

$$F = \{h(w_1), h(w_2), \ldots, h(w_K)\} \tag{10}$$

As the combined approach between the LSPM^{Auto} and spatial-BoVW method, the final feature vector computed as normalized concatenated histogram, is given by

$$\Phi = [H, F] \tag{11}$$

3. **Experiments and Results.** To evaluate the proposed method, Support Vector Machine (SVM) with a quadratic kernel is integrated as a discrimination function for texture classification. The implementation is accomplished using MATLAB 2016a, following default configuration for parameter adjustment. Multi-class classification used in this experiment consists of the one-versus-all strategy and 10-fold cross-validation. In this study, two experiments were conducted to compare the results among the LSPM^{Auto} method, spatial-BoVW model, and conventional methods in texture analysis.

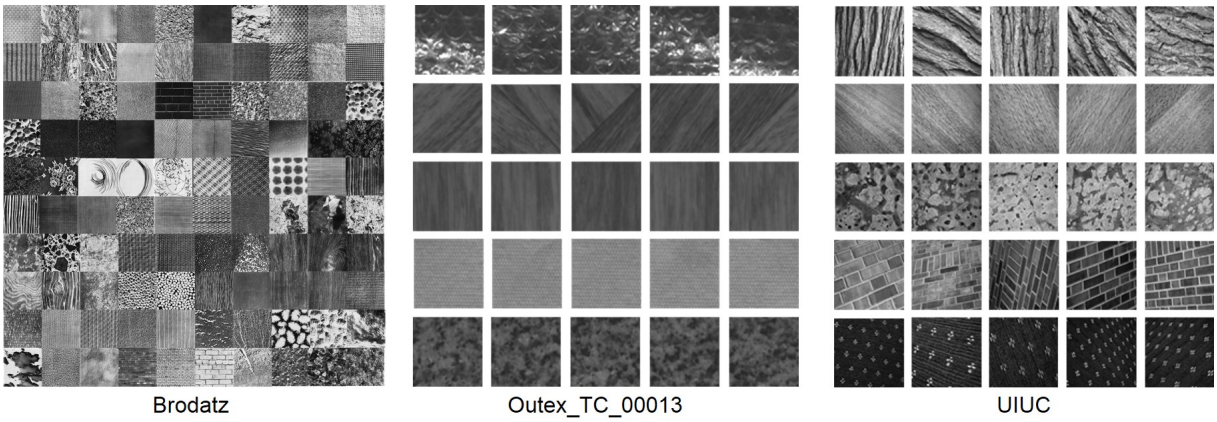

FIGURE 2. Example of texture images of the three standard texture databases

3.1. **Datasets.** The experiments were conducted using three texture databases for evaluation, as illustrated in Figure 2, which are explained as follows.

- Brodatz Texture Album dataset [17] is a benchmark texture database for evaluating texture analysis and its applications such as texture classification. The texture sample of each class is manipulated by separating partitive ten non-overlapping images. The experiments were considered total sample images in 111 classes, 10 grayscale images.
- The UIUC database is a complicated dataset for our study [18]. All samples are included by the significantly various image scales and viewpoints according to distortion appearances and deformable objects. Moreover, the illumination conditions are uncontrolled image appearances. The experiments were considered total texture images, including 25 classes, 40 grayscale images. The image size is 128×128 pixels.

- The Suite Outex_TC_00013, or simplify as Outex [19] database, contains surface textures and natural scenes. The texture images in 68 classes containing 20 images were considered in the experiments, each comprising 128×128 pixels.

3.2. **Parameter analysis.** In the experiments, the two main parameters of the proposed method are the set of radius and its neighbors (r, p) and number of visual words k. We have defined the set $(r, p) = \{(1, 4), (2, 8), (3, 16)\}$ as recommended by related work [11, 13, 14]. This parameter was fixed as our proposed method. Regarding our previous work [14], we believe that these parameters present satisfactory performance for applying in this paper. The multiscale region-level mapping is able to expand the local investigation of the texture pattern by separating the neighborhood, which a vertice can examine based on the range of the Euclidean distances. This fact would seem to affect the method's ability to discriminate. The number of visual words k has varied in a different vocabulary number composed by $k = \{10, 20, 30, \ldots, 100\}$. Table 1 presents the retrieval accuracy rate with a number of visual words ranging from 10 to 100 (#Vocab.) in the Brodatz, UIUC, and Outex databases. We notice that the best results were achieved when k is between 50 to 100. According to the experimental results, when $k > 100$ it gave results with a significantly decreased success rate. Also, the related work [12] has mentioned using k equal to 100 achieved the best result. Therefore, the best number of visual words selected in the experiment uses $k = 100$, as the accuracy results in Table 1.

Table 1. Accuracy results on different vocabulary sizes (#Vocab.) over Brodatz, UIUC and Outex databases. Bold values correspond to the best result for each database.

#Vocab.	**Database**		
	Brodatz	**UIUC**	**Outex**
10	75.86	46.14	59.42
20	85.39	67.61	75.37
30	87.95	68.29	80.79
40	88.73	74.65	80.29
50	91.13	77.01	83.10
60	90.86	76.39	83.10
70	91.44	79.61	82.83
80	90.81	**81.36**	83.82
90	92.09	80.81	84.53
100	**92.22**	80.36	**84.70**

3.3. **Results obtained by LSPM^{Auto} and spatial-BoVW model.** In our previous work [14], we proposed and discussed how spatial arrangement analysis by the LSPM method is a crucial extractor for texture characterization. Although the LSPM method exhibited an improvement in classification performance, it still required tuning a threshold parameter, the same as the traditional complex network model. We noted that the tuning threshold approach involves expensive processes of indeterminate length, and therefore the LSPM^{Auto} method is proposed in this paper. The experimental result of the LSPM^{Auto} method is shown in Table 3. In terms of accuracy, the result indicated a decrease in the success rate with 67.49% compared to our previous work [14], 81.80% on Brodatz database. Accordingly, we can note that this approach did not obtain enough descriptors to extract the local spatial relationship of neighbors, including to describe the coarseness of primitives in textures.

The spatial-BoVW method is employed to search among more texture information by collecting a set of weights of edges on radial distances based on k-means clustering. The repetitions of primitives in textures are collected as vocabulary. Thus, stochastic textures can be visualized by this method. Table 2 illustrates the classification results in each database using the Spatial-BoVW (S-BoVW) method and the combined approach (Combined) among the LSPM^{Auto} method and spatial-BoVW model. The table lists the number of features (#Features) corresponding to number of vocabulary sizes from the spatial-BoVW model (#S-BoVW) and total number of features from the combined approach (#Combined). The Combined can improve the accuracy rate compared with the proposed S-BOVW in the experimental results, significantly for all three databases. Furthermore, compared with the LSPM^{Auto} method, the S-BoVW is more successful at finding relevant informants on stochastic texture based on the experimental results.

Table 2. The classification results in each database using Spatial-BoVW (S-BoVW) method and combined approach (Combined). Bold values correspond to the best result for each approach over each database.

#Features		**Brodatz**		**UIUC**		**Outex**	
#S-BoVW	#Combined	S-BoVW	Combined	S-BoVW	Combined	S-BoVW	Combined
10	44	75.86	79.88	46.14	67.70	59.42	77.9
20	54	85.39	84.32	67.61	76.23	75.37	78.81
30	64	87.95	87.21	68.29	76.52	80.79	82.25
40	74	88.73	88.47	74.65	80.20	80.29	81.41
50	84	91.13	91.05	77.01	80.21	83.10	83.98
60	94	90.86	92.20	76.39	80.40	83.10	83.60
70	104	91.44	92.91	79.61	81.84	82.83	83.14
80	114	90.81	91.57	**81.36**	82.62	83.82	83.27
90	124	92.09	93.15	80.81	83.37	84.53	84.05
100	134	**92.22**	**93.51**	80.36	**83.40**	**84.70**	**85.44**

3.4. **Comparison of the proposed and traditional methods.** For more evaluation of the proposed system, the other methods are provided for comparison, as the following: Fourier descriptors [5], Graph-LBP [7], CNTD [11], and LSPM [14]. To each method, the used-parameters are the best result. This experiment also includes basic LBP and LBP^{riu2} [21, 22] operators. We use the concatenation of the histograms computed for $(p, r) = (8, 1), (16, 2), (24, 3)$ to characterize a texture pattern. In Gabor filters [6], the total of 40 filters is composed of 8 rotation filters and 5 scale filters by a frequency range between 1.2 to 1.4. For co-occurrence matrices [23], the experiments define the co-occurrence matrix comprising with $d = 1$ and 2 and angles = 0, 45, 90, and 135, in non-systemic version for each image. The energy and entropy descriptors are computed from each co-occurrence matrix to compose an image feature vector in the total number of descriptors equal to 24.

Table 3 represents comparison classification results between the proposed method with other methods based on the Brodatz, UIUC, and Outex databases. We chose the best result in Table 2 when k is equal to 100 feature dimensions for comparison. The Brodatz database includes many texture patterns, such as regular patterns and randomly scattered patterns. The experimental results show that when we applied the spatial-BoVW method together with LSPM^{Auto}, some randomly scattered patterns, that is, stochastic textures were extracted, so the success rate was improved with 93.51%. Although our proposed methods have achieved not the best results compared with CNTD 95.27%, this proposed

TABLE 3. The comparison accuracy results of the proposed methods and traditional texture analysis method in each database. Bold text corresponds to our proposed methods.

Method	**No. features**	**Success rate (%)**		
		Brodatz	**UIUC**	**Outex**
Fourier descriptors [5]	90	78.02	72.40	72.13
Gabor filters [6]	40	76.58	60.00	74.26
Graph-LBP [7]	68	90.92	84.51	78.34
CNTD [11]	108	95.27	–	86.76
BoVW-CN [13]	100	76.29	–	68.38
LSPM [14]	60	81.80	80.90	83.46
LBP^{riu2} [21]	54	82.34	76.60	77.28
Basic LBP [22]	256	82.52	51.30	71.47
Co-occurrence matrices [23]	24	73.69	69.77	48.97
Proposed method (LSPM^{Auto})	34	67.49	64.69	66.49
Proposed method (Spatial-BoVW model)	100	92.22	80.36	84.67
Proposed method (Combined-approach)	134	93.51	83.40	85.44

approach presents satisfactory results for an extensive range of image classes. The UIUC database is used for evaluating whether the spatial arrangement is an essential property since the image samples consisted of the various image viewpoints and scales. The LSPM^{Auto} method was used to analyze the spatial relationship and spatial arrangement of pixels. The results of the combined approach achieved higher accuracy than the LSPM method [14]. It indicates that the combined approach, LSPM^{Auto} and spatial-BoVW methods, had achieved good performances with 83.40% over the challenging UIUC database. This result shows that the proposed method is robust against many variations present in the UIUC database, such as containing images collected from random viewpoints, changes in scale, illumination, and image dimension. The Outex database contains surface texture and natural scenes under different illuminations. The proposed method results indicate an 85.44% performance, which is good compared with the CNTD and LSPM method.

Therefore, the experimental results confirmed that the proposed methods adequately described the spatial configuration of local and global texture features. These experiments assure us that the method has good invariance and is capable of texture classification using the combination between the LSPM^{Auto} and the spatial-BoVW model as an enhanced texture descriptor.

4. **Conclusions.** This paper proposed enhanced feature descriptors through graph-based representation for texture characterization. Regarding the previous work result, we extended and proposed the LSPM^{Auto} method as local feature extraction aids to texture discrimination. Spatial-BoVW model was proposed to build a vocabulary of visual words based on a spatially weighted graph with multiscale region-level properties. The combined approach proposed as a texture descriptor based on graph theory. We evaluated the proposed methods using three different texture databases: Brodatz, UIUC, and Outex. The results showed that the proposed method could enhance feature extractor to produce more accurate results that are superior and satisfactory compared with other methods and our previous work.

In future work, we aim to investigate how the proposed system could be applied to a clothing database containing images of non-rigidly deformable objects. Based on related work [3, 4, 7], a classification system focusing on the wrinkle characteristics of fabric can handle this problem, so, the proposed modification of our system could also be applied to such fabric characterization.

Acknowledgment. This paper disclosed receipt of the following financial support provided by Srinakharinwirot University under grant 028/2563. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved this work.

REFERENCES

[1] L. Urdang and S. Flexner, *The Random House Dictionary of the English Language*, Random House, New York, 1968.

[2] A. Materka and M. Strzelecki, Texture analysis methods – A review, *Tech. Rep.*, Institute of Electronic, Techinical University of Lodz, 1998.

[3] S. Thewsuwan and K. Horio, Preprocessing techniques based on LBP and Gabor filters for clothing classification, *Proc. of the International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, Phuket, Thailand, pp.1-6, 2016.

[4] K. Yamazaki and M. Inaba, Clothing classification using image features derived from clothing fabrics, wrinkles and cloth overlaps, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp.2710-2717, 2013.

[5] R. Azencott, J.-P. Wang and L. Younes, Texture classification using windowed Fourier filters, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.19, no.2, pp.148-153, 1997.

[6] M. Idrissa and M. Acheroy, Texture classification using Gabor filters, *Pattern Recognition Letters*, vol.23, no.9, pp.1095-1102, 2002.

[7] S. Thewsuwan and K. Horio, Texture-based features for clothing classification via graph-based representation, *Journal of Signal Processing*, vol.22, no.6, pp.299-305, 2018.

[8] J. J. de Mesquita Sá Junior, A. R. Backes and P. C. Cortez, Texture analysis and classification using shortest paths in graphs, *Pattern Recognition Letters*, vol.34, no.11, pp.1314-1319, 2013.

[9] L. F. Costa, F. A. Rodrigues, G. Travieso and P. R. V. Boas, Characterization of complex networks: A survey of measurements, *Tech. Rep. Condmat/0505185*, 2005.

[10] M. E. J. Newman, The structure and function of complex networks, *SIAM Review*, vol.45, no.2, pp.167-256, 2003.

[11] A. R. Backes, D. Casanova and O. M. Bruno, Texture analysis and classification: A complex network-based approach, *Information Sciences*, vol.219, pp.168-180, 2013.

[12] L. F. S. Scabini, W. N. Gonçalves and A. A. Castro Jr., Texture analysis by Bag-of-Visual-Words of complex networks, *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, pp.485-492, 2015.

[13] G. V. L. de Lima, P. T. M. Saito, F. M. Lopes and P. H. Bugatti, Classification of texture based on Bag-of-Visual-Words through complex networks, *Expert Systems with Applications*, vol.133, pp.215-224, 2019.

[14] S. Thewsuwan and K. Horio, Texture classification by local spatial pattern mapping based on complex network model, *International Journal of Innovative Computing, Information and Control*, vol.14, no.3, pp.1113-1121, 2018.

[15] K. S. Arun, V. K. Govindan and S. D. Madhu Kumar, Enhanced bag of visual words representations for content-based image retrieval: A comparative study, *Artif. Intell. Rev.*, vol.53, pp.1615-1653, 2020.

[16] O. A. Penatti, F. B. Silva, E. Valle, V. Gouet-Brunet and R. D. S. Torres, Visual word spatial arrangement for image retrieval and classification, *Pattern Recognition*, vol.47, no.2, pp.705-720, 2014.

[17] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, Dover Publications, New York, 1966.

[18] 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.

[19] T. Ojala, T. Maenpaa, M. Pietikainen, J. Viertola, J. Kyllonen and S. Huovinen, Outex – New framework for empirical evaluation of texture analysis algorithms, *Object Recognition Supported by User Interaction for Service Robots*, vol.1, pp.701-706, 2002.

[20] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Transactions on Systems, Man, and Cybernetics*, vol.9, no.1, pp.62-66, 1979.

[21] T. Ojala, M. Pietikainen and T. Maenpaa, 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.

[22] T. Ojala, M. Pietikainen and D. Harwood, A comparative study of texture measures with classification based on feature distributions, *Pattern Recognition*, vol.29, no.1, pp.51-59, 1996.

[23] R. M. Haralick, Statistical and structural approaches to texture, *Proc. of the IEEE*, pp.786-804, 1975.