Two decades of local binary patterns: A survey

9

Matti Pietikäinen and **Guoying Zhao**

Center for Machine Vision Research, Department of Computer Science and Engineering,
University of Oulu, Finland

9.1 INTRODUCTION

Texture is an important characteristic of many types of images. It can be seen in images ranging from multispectral remotely sensed data to microscopic images. Texture can play a key role in a wide variety of applications of computer vision and image analysis. Therefore, the analysis of textures has been a topic of intensive research since the 1960s. Most of the proposed methods have not been, however, capable to perform well enough for real-world textures.

In recent years, very discriminative and computationally efficient local texture descriptors have been developed, such as local binary patterns (LBPs), which have led to significant progress in applying texture methods to different problems and applications. The focus of research has broadened from 2-D textures to 3-D textures and spatiotemporal (dynamic) textures. Due to this progress, the division between texture descriptors and more generic image or video descriptors has been disappearing.

The LBP operator can be seen as a unifying approach to the traditionally divergent statistical and structural models of texture analysis. Perhaps the most important property of the LBP operator is its robustness to monotonic grayscale changes caused, for example, by illumination variations. Another important property is its computational simplicity, which makes it possible to analyze images in challenging real-time settings.

The original LBP was invented two decades ago, published first at ICPR 1994 and then in *Pattern Recognition* [1], but at that time, one could not imagine what a great success it would be today. In the 1990s, LBP did not receive interest in the scientific community, because it was regarded as an ad hoc method. However, the promising power of LBP was already evident, because it performed much better in texture classification and segmentation tasks than the state of the art at that time [2]. Due to its computational simplicity and good performance, LBP was also successfully used in some applications such as industrial inspection. The theoretical foundations of LBP became much more clear after our research on multidimensional distributions of signed gray-level differences, carried out jointly with Prof. Erkki Oja and Dr. Kimmo Valkealahti [3].

To a large extent, the scientific community accepted LBP after its generalized version was published in *IEEE PAMI* journal [4], and further it was shown to be highly successful in face recognition, published first at ECCV 2004 and then in *IEEE PAMI* [5]. In 2014, the ECCV paper was awarded with the prestigious Koenderink Prize for Fundamental Contributions in Computer Vision. Different types of applications of LBP to motion analysis have been proposed after the spatiotemporal LBP was introduced, also in *IEEE PAMI* [6]. All these papers are highly cited, reflecting the increasing popularity of LBP. Due to this success of LBP, a book describing the basic methods and surveying different variants and applications was published [7]. Another survey on LBP and its applications in face analysis appeared in Ref. [8]. An edited book on some LBP variants and applications was published in 2013 [9].

After these surveys, which covered the progress until 2010, the interest on LBP has been growing further. LBP is no longer just a simple texture operator, but it forms the foundation for a new direction of research dealing with local binary image and video descriptors. Many different variants of LBP have been proposed to improve its robustness and to increase its discriminative power and applicability to different types of problems. A large number of new papers have been published in leading journals and conferences. Due to its discriminative power and computational simplicity, the LBP operator has become highly popular in various applications, including facial image analysis, biometrics, medical image analysis, motion and activity analysis, and content-based retrieval from image and video databases. Figure 9.1 depicts the number of citations in Web of Science and Google Scholar to the landmark LBP paper published in 2002 [4], showing a clear increase especially after LBP was successfully adopted for face recognition in 2006 and its spatiotemporal version was proposed in 2007.

Based on all these, the publication of this survey is very timely, covering the progress until early 2014. It complements and extends our preliminary survey presented in Sections 2.9 and 3.5 of Ref. [7], by focusing on the most important recent developments, new types of variants, and future challenges.
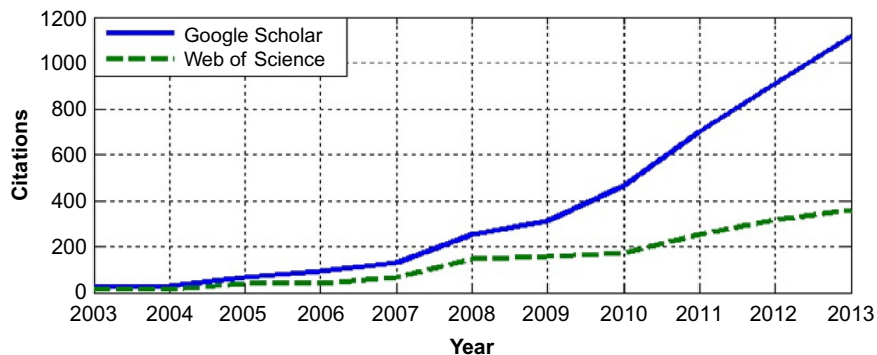
FIGURE 9.1

Annual citations to the PAMI 2002 paper on LBP.

9.2 AN OVERVIEW OF BASIC LBP OPERATORS

The LBP is based on the assumption that texture has locally two complementary aspects, a pattern and its strength. The pixels of an image are labeled by thresholding the neighborhood of each pixel, and the result is considered as a binary number. The distribution of the LBP labels computed over a region is then used for texture description.

The original LBP operator shown in Figure 9.2 works in a 3×3 neighborhood, using the center value as a threshold [1]. The thresholded values are multiplied with weights of the corresponding pixels, and by summing up the result an LBP code is obtained. The contrast measure C is obtained by subtracting the average gray level of the pixels below the threshold from that of those above (or equal to) the center pixel. If all neighbors of the center pixel have the same value (1 or 0), the value of C is set to 0.

The distributions of LBP codes or 2-D distributions of LBP and C are used as features in texture recognition. The LBP operator was extended to use neighborhoods of different sizes in Ref. [4]. Using a circular neighborhood and bilinearly interpolating values at noninteger pixel coordinates allow any radius and number of pixels in the neighborhood. The multiscale LBP (MLBP) is obtained by concatenating histograms produced by operators at different radii. The grayscale variance of the local neighborhood can be used as the complementary contrast measure.

Ojala et al. [4] found that some of the LBP patterns occur more frequently than others, representing, for example, edges, curves, line ends, flat areas, and spots. Based on this observation, the so-called uniform patterns were defined to reduce the number of patterns. Uniform patterns have been widely used and were necessary, for example, to reduce the length of the feature vector in face description.
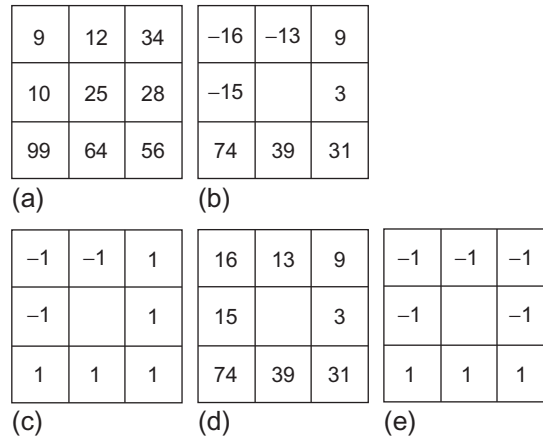
In their recent paper, Guo et al. [10] proposed a completed modeling of the LBP operator and developed an associated completed LBP (CLBP) scheme for texture classification. The local differences are decomposed into two complementary components: the signs (like LBP) and the magnitudes (Figure 9.3). The magnitude component provides an effective alternative for the complementary contrast measure of LBP. In addition to this texture-related information, they also include information about image intensity in their representation.

	Example			Thresholded			Weights	
6	5	2	1	0	0	1	2	4
7	6	1	1		0	128		8
9	8	7	1	1	1	64	32	16

Pattern = 11110001   **LBP =** 1 + 16 + 32 + 64 + 128 = 241
C = (6+7+9+8+7)/5 − (5+2+1)/3 = 4.7

FIGURE 9.2

The original LBP.

9	12	34
10	25	28
99	64	56

(a)

−16	−13	9
−15		3
74	39	31

(b)

−1	−1	1
−1		1
1	1	1

(c)

16	13	9
15		3
74	39	31

(d)

−1	−1	−1
−1		−1
1	1	1

(e)

FIGURE 9.3

CLBP: (a) 3 × 3 sample block, (b) the local differences, (c) the sign, (d) magnitude components, and (e) thresholded binary value of (d) against the average magnitude in whole image region.

Dynamic texture descriptors provide a very effective new tool for motion analysis. Zhao and Pietikäinen [6] extended the original LBP to a spatiotemporal representation for dynamic texture analysis. For this purpose, the so-called volume local binary pattern (VLBP) operator was proposed, in which dynamic texture is considered as a set of volumes in the (X, Y, T) space where X and Y denote the spatial coordinates and T denotes the frame index (time). The neighborhood of each pixel is thus defined in three-dimensional space. The VLBP combines motion and appearance together to describe dynamic textures. To make VLBP computationally simple and easy to extend, an operator based on co-occurrences of LBPs on three orthogonal planes (LBP-TOPs) considers three orthogonal planes – XY, XT, and YT – and concatenates LBP co-occurrence statistics in these three directions as shown in Figure 9.4. The circular neighborhoods are generalized to elliptical sampling to fit to the space-time statistics.

The first application problems to which spatiotemporal LBP was adopted were facial expression recognition, face and gender recognition, lip-reading, and action recognition [7].

Facial images are composed local texture patterns, and based on this observation, Ahonen et al. [5] introduced a method for combining local and global information for face (object) description. The face image is divided into several regions from which the LBP feature distributions are extracted and concatenated into an enhanced feature vector to be used as a face descriptor.

This approach has emerged as one of the key paradigms for facial image analysis and has also been adopted to other application domains. The length of the feature vector is an important design parameter in this description. Ahonen et al. used uniform patterns to reduce the dimensionality. Figure 9.5 illustrates the use of LBP for face description [5].
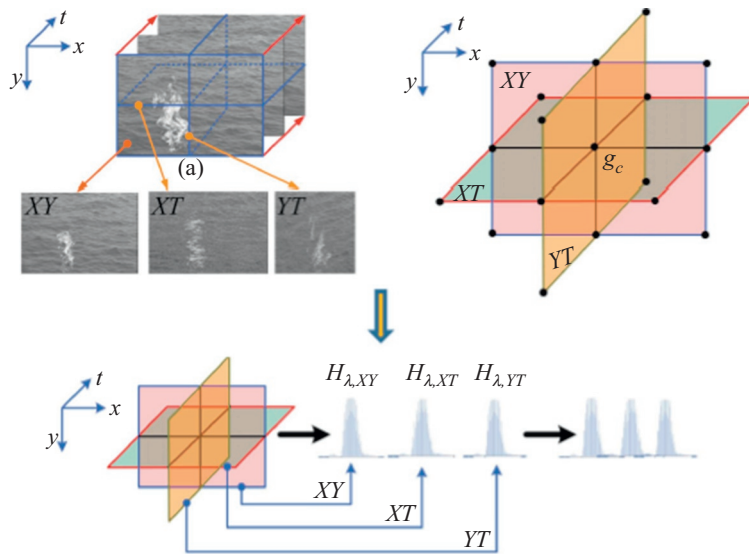
FIGURE 9.4

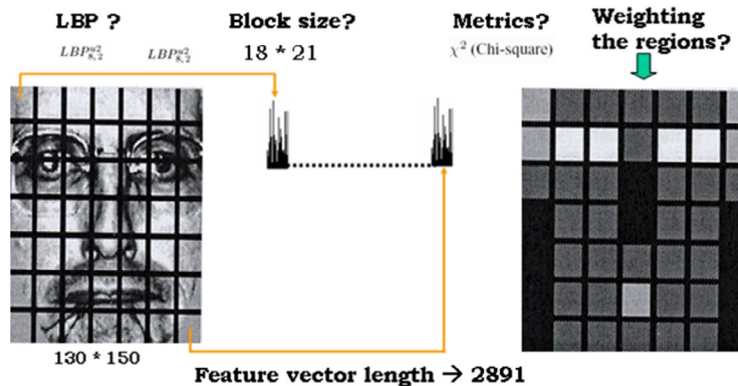LBP-TOP representation for spatiotemporal description.

FIGURE 9.5

Face description using LBP.

9.3 LBP VARIANTS IN THE SPATIAL DOMAIN
9.3.1 DIFFERENT TYPES OF VARIANTS

Due to its flexibility, the LBP method can be easily modified to make it suitable for the needs of different types of problems. Several extensions and modifications of LBP have been proposed with an aim to increase its robustness and discriminative power. In this section, different variants are divided into such categories that describe

their roles in feature extraction. Some of the variants could belong to more than one category, but only the most obvious categories are considered here. The choice of a proper method for a given application depends on many factors, such as the discriminative power, computational efficiency, robustness to lighting and other variations, and the imaging system used.

9.3.1.1 Preprocessing

In many applications, it is useful to preprocess the input image prior to LBP feature extraction. Especially, multiscale Gabor filtering and edge detection have been used for this purpose. Gabor filtering has been widely used before LBP computation in face recognition. A motivation for this is that methods based on Gabor filtering and LBP provide complementary information: LBP captures small and fine details, while Gabor filters encode appearance information over a broader range of scales. For example, Zhang et al. [11] proposed the extraction of LBP features from images obtained by filtering a facial image with 40 Gabor filters of different scales and orientations. The extracted features are called local Gabor binary patterns (LGBPs). A downside of the method is the high dimensionality of the LGBP representation.

Tan and Triggs [12] developed a very effective preprocessing chain for compensating illumination variations in face images. It is composed of gamma-correction, difference of Gaussian filtering, masking (optional), and equalization of variation. This approach has been very successful in LBP-based face recognition under varying illumination conditions. When using it for the basic LBP, the last step can be omitted due to LBP's invariance to monotonic grayscale changes.

In many recent studies, edge detection has been used prior to LBP computation to enhance the gradient information. Gradient images are more insensitive to lighting variations than the original images. Perhaps the first one was Yao and Chen [13] proposing local edge patterns (LEPs) to be used with color features for color texture retrieval. In LEPs, the Sobel edge detection and thresholding are used to find strong edges, and then LBP-like computation is used to derive the LEPs.

Li et al. [14] presented an approach based on capturing the intrinsic structural information of face appearances with multiscale heat kernel matrices. Heat kernels perform well in characterizing the topological structural information of face appearance. Histograms of LBPs computed for nonoverlapping blocks are then used for face description.

Also, other types of preprocessing have been applied with LBP, including wavelets and momentograms, as described in Sections 9.3.1.4 and 9.3.1.9.

9.3.1.2 Neighborhood topology and sampling

One important factor that makes the LBP approach so flexible to different types of problems is that the topology of the neighborhood from which the LBP features are computed can be different, depending on the needs of the given application.

The extraction of LBP features is usually done in a circular or square neighborhood. A circular neighborhood is important especially for rotation-invariant operators. However, in some applications, such as face recognition, rotation invariance is not required, but anisotropic information may be important. To exploit this,

Liao and Chung [15] proposed an elliptical binary pattern for face recognition. Nanni et al. [16] investigated the use of different neighborhood topologies (circle, ellipse, parabola, hyperbola, and Archimedean spiral) and encodings in their research on LBP variants for medical image analysis. An operator using quinary encoding in an elliptic neighborhood (EQP) provided the best performance. He et al. [17] used rotation-invariant LBP with elliptic sampling in four directions together with circular sampling to get anisotropic and isotropic information. A multistructure LBP (Ms-LBP) was achieved by applying this operator at different layers of an image pyramid. A downside of the method is that quite large image samples are needed for extracted macrostructures. Wolf et al. [18] considered different ways of using bit strings to encode the similarities between patches of pixels, which could capture complementary information to pixel-based descriptors. They proposed a three-patch LBP (TPLBP) and four-patch LBP (FPLBP). For each pixel in TPLBP, for example, a $w \times w$ patch centered at the pixel and S additional patches distributed uniformly in a ring of radius r around it are considered. Then, the values for pairs of patches located on the circle at a specified distance apart are compared with those of the central patch. The value of a single bit is set according to which of the two patches is more similar to the central patch. The code produced will have S bits per pixel. In FPLBP, two rings centered on the pixel were used instead of one ring in TPLBP.

Orjuela et al. [19] presented geometric local textural patterns (GLTPs), which are based on exploring intensity changes on oriented neighborhoods. An oriented neighborhood describes a particular geometry composed of points on circles with different radii around the center pixel. A digital representation of the points on the oriented neighborhood defines a GLTP code. The simple case called geometric local binary pattern (GLBP) is based on Boolean comparisons.

Wang et al. [20] proposed a sampling structure based on combining pixel and patch to mimic the retinal sampling grid (pixel to patch [PTP]). Also, a neighboring intensity relationship (NIR) operator was proposed to complement LBP texture information by exploring grayscale properties between neighborhoods. Two rotation-invariant descriptors were also proposed: the local intensity relationship pattern (LNIRP) based on the NIR operator and LNIRP_PTP using the PTP sampling structure. The method is computationally simple, has small feature dimensionality, and is training-free. Ylioinas et al. [21] introduced a dense sampling approach through the form of upsampling to extract more stable and discriminative texture patterns in local regions. Experiments on face recognition, texture classification, and age group estimation problems on various challenging benchmark databases demonstrate the efficiency of the proposed scheme.

9.3.1.3 Thresholding and encoding

A drawback of the LBP method, as well as of all local descriptors that apply vector quantization, is that they are not robust in the sense that a small change in the input image would always cause a small change in the output. LBP may not work properly for noisy images or on flat image areas of constant gray level. This is due to the thresholding scheme of the operator.

Instead of using the value of the center pixel for thresholding in the local neighborhood, other techniques have also been considered. Hafiane et al. [22] proposed median binary pattern operator by thresholding the local pixel values, including the center pixel, against the median within the neighborhood. The so-called improved LBP, on the other hand, compares the values of the neighboring pixels against the mean gray level of the local neighborhood [23].

Tan and Triggs [12] proposed a three-level operator called local ternary patterns (LTPs) (using one threshold T), for example, to deal with problems on near constant image areas. In ternary encoding, the difference between the center pixel and a neighboring pixel is encoded by three values $(1, 0, \text{and } -1)$ according to a threshold T. The ternary pattern is divided into two binary patterns taking into account its positive and negative components. The histograms from these components computed over a region are then concatenated. Figure 9.6 depicts an example of splitting a ternary code into positive and negative codes.

In order to make the LBP more robust against negligible changes in pixel values, the thresholding scheme of the operator was modified in Ref. [24] by adding a small constant a in local difference computations. The bigger the value of $|a|$, the bigger the allowed changes in pixel values without affecting the thresholding results. In order to retain the discriminative power of the LBP operator, a relatively small value should be used. An advantage of this modified LBP compared with the three-valued LBPs described above is that the feature vector length remains the same as in the ordinary LBP. Heikkilä et al. [25] proposed center-symmetric LBP (CS-LBP) to reduce the feature vector length of the original LBP. The center-symmetric pairs of pixels in the neighborhood are compared, instead of comparing all neighbors with the center pixel. Therefore, the CS-LBP produces only 16 patterns instead of 256 for the basic LBP. An interest region descriptor based on CS-LBP compared favorably to the SIFT operator in image matching and categorization experiments, especially for images with lighting variations. A similar thresholding approach as in Ref. [24] was used to improve the robustness of the CS-LBP descriptor.
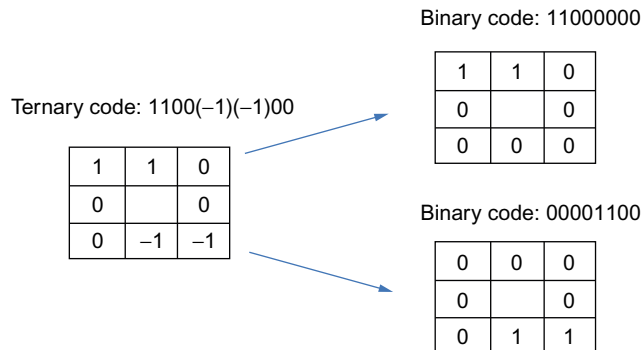
FIGURE 9.6

Local ternary pattern (LTP) operator.

Nanni et al. [16] studied the effects of different encodings of the local grayscale differences, using binary (B), ternary (T), and quinary (Q) encodings. In binary coding, the difference between a neighboring pixel and the center pixel is encoded by two values (0 and 1) as in LBP; in ternary encoding, it is encoded by three values as in LTP; and in quinary encoding, it is encoded by five values (-2, -1, 0, 1, and 2) according to two thresholds (T_1 and T_2). A quinary code can be split into four binary LBP codes.

A downside of the methods using one or two thresholds is that the methods are not strictly invariant to local monotonic gray-level changes as the original LBP. The feature vector lengths of these operators are also longer.

A soft three-valued LBP using fuzzy membership functions was proposed by Ahonen and Pietikäinen [26] to improve the robustness. In soft LBP, one pixel typically contributes to more than one bin in the histogram. An extensive study on generalized fuzzy LBPs was carried out by Katsigiannis et al. [27]. A disadvantage of the fuzzy methods is their increased computational cost.

Liao et al. [28] noticed that adding the small threshold (T) in LTP is not invariant under scaling of intensity values. The intensity scale-invariant property of a local comparison operator is very important, for example, in background modeling, because lighting variations, either global or local, often cause sudden changes of grayscale intensities of neighboring pixels simultaneously, which would approximately be a scale transform with a constant factor. Therefore, a scale-invariant local ternary pattern operator was developed for dealing with the grayscale intensity changes in complex background. They also proposed a pattern kernel density estimation technique to effectively model the probability distribution of local patterns in the pixel process. Ylioinas et al. [29] adopted a related method for image appearance description. Compared to the standard LBP histogram statistics where one labeled pixel always contributes to one bin of the histogram, the proposed method exploits a kernel-like similarity function to determine weighted votes contributing several possible pattern types in the statistic. As a result, the method yields a more reliable estimate of the underlying LBP distribution of the given image, providing improved performance especially for small-sized samples.

Trefny and Matas [30] proposed two encoding schemes, which are complementary to the standard LBPs and also invariant to monotonic intensity transformations. The binary value transition coded LBP is composed of neighbor pixel comparisons in a clockwise direction for all pixels except the central pixel (CI), encoding the relation between neighboring pixels. Direction coded LBP is related to the CS-LBP operator but uses also center pixel information for encoding. Intensity variation along each of the four basic directions is coded into two bits. The first bit encodes whether the center pixel is an extreme point and the second bit encodes whether the difference of border pixels compared to the center pixel grows or falls.

Liu et al. [31] extracted two different and complementary types of features, pixel intensities and differences, from local patches. The intensity-based features consider the intensities of the CI and its neighbors (NI), while for the difference-based feature, two components are computed: the radial difference (RD) and the angular difference

(AD). Two intensity-based descriptors CI-LBP and NI-LBP and two difference-based descriptors RD-LBP and AD-LBP were proposed.

Inspired by LBP, higher-order local derivative patterns (LDPs) were proposed by Zhang et al. [32], with applications in face recognition. The basic LBP represents the first-order circular derivative pattern of images, a micropattern generated by the concatenation of the binary gradient directions as was shown in Ref. [33]. The higher-order derivative patterns extracted by LDP will provide more detailed information but may also be more sensitive to noise than in LBP. The length of the feature vector is also four times the length of LBP, and the issue of rotation invariance was not addressed. To address these problems, Guo et al. [34] proposed local directional derivative patterns (LDDPs), a special case of which is the LBP. Similar to LBP, rotation invariance can be easily defined also for LDDP. Best recognition accuracy can be obtained by combining first- and second-order LDDPs.

Hussain and Triggs [35] tackled the problem of earlier LBP variants based on hand-specific codings, which limits them to small spatial support and coarse gray-level comparisons. They proposed local quantized patterns (LQPs), using vector quantization to code larger or deeper patterns. Precomputed lookup tables are used to make coding very fast. Their approach outperformed HOG, LBP, and LTP in challenging object detection and texture classification problems. Later, Huang et al. [36] introduced CLQP to include also magnitude (like in CLBP) and orientation information. A better and much faster way for initializing the vector quantization was also proposed.

Jiang et al. [37] proposed gradient LBPs for human detection. After LBP is calculated for each pixel, a 56-bin histogram is built with the co-occurrence of width value and angle value of LBPs. Width is the number of value 1s in the binary code of this pixel. Eight direction codes with 0-7 are defined in the direction of eight NIs. Angle value is the direction code of the middle pixel in 1 area of its binary code. Magnitude of gradient is used as weight to this 56-bin histogram.

9.3.1.4 Multiscale analysis

From a signal processing point of view, the sparse sampling used by the MLBP operator [4] may not result in an adequate representation of the signal, resulting in aliasing effects. Due to this, some low-pass filtering would be needed to make the operator more robust. From the statistical point of view, however, even sparse sampling is acceptable provided that the number of samples is large enough. The sparse sampling is commonly used, for example, with the methods based on grayscale difference or co-occurrence statistics. Mäenpää and Pietikäinen [38] proposed to use Gaussian low-pass filters for collecting texture information from a larger area than the original single pixel. The filters and sampling position were designed to cope with the neighborhood while minimizing the redundant information. With this approach, the radii of the LBP operators used in the multiscale version grow exponentially.

Another extension of MLBP operator is the multiscale block local binary pattern (MB-LBP) [39], which has gained popularity especially in facial image analysis. The key idea of MB-LBP is to compare average pixel values within small blocks instead

of comparing pixel values. Instead of the fixed uniform pattern mapping, MB-LBP was proposed to be used with a mapping that is dynamically learned from training data. In this mapping, the N most often occurring MB-LBP patterns receive labels $0.N - 1$, and all the remaining patterns share a single label. The number of labels, and consequently the length of the MB-LBP histogram, is a parameter the user can set.

A straightforward way for multiscale analysis is to utilize a pyramid of the input image computed at different resolutions and then concatenate LBP distributions computed from different levels of the pyramid. In their research on contextual analysis of textured scene images, Turtinen and Pietikäinen [40] combined this kind of idea with the original MLBP approach: image patches at three different scales were resized to the same size, and then LBP features were computed using LBPs with three different radii.

He et al. [41] developed a pyramid-based Ms-LBP for texture classification. It is obtained by executing the LBP on different layers of image pyramid, allowing to extract both micro- and macrostructures from textures. Five templates are used for creating the pyramid. The first one is a 2-D Gaussian function used to smooth the image. Other four anisotropic filters are used to create anisotropic subimages of the pyramid in four directions. Qian et al. [42] also extended LBP to pyramid transform domain. Different ways of sampling in pyramid construction are considered, including no sampling, partial sampling, and spatial pyramid sampling, of which the last one leads to the smallest computational cost.

Song and Li [43] combined wavelets and LBP. They built up the image description using a hierarchical framework based on low-dimensional wave LBP features, which not only extracts multiscale-oriented features and local image patterns but also captures multilevel (pixel, patch, and image) features. A very competitive performance was demonstrated in experiments.

Liu et al. [44] proposed a computationally simple approach to multiscale analysis with their binary rotation-invariant and noise-tolerant texture descriptor. Points are sampled in a circular neighborhood, but the number of bins in a single scale LBP histogram is kept constant and small by averaging over several contiguous pixels in the circle before binarization. This allows to encode a large number of scales and also reduces the effects of noise. Both sign and magnitude components, like in CLBP, are considered.

9.3.1.5 Handling rotation and scale variations

LBPs have been used for rotation-invariant texture recognition since the late 1990s [45]. In the most widely used version proposed in Ref. [4], the neighboring n binary bits around a pixel are clockwise-rotated n times. A maximal number of the most significant bits are used to express this pixel.

Guo et al. [46] developed an adaptive LBP (ALBP) by incorporating the directional statistical information for rotation-invariant classification. The directional statistical features, specifically the mean and standard deviation of the local absolute difference, are extracted and used to improve the LBP classification efficiency. In addition, the least square estimation is used to adaptively minimize the local

difference for more stable directional statistical features. Garcia-Olalla et al. [47] extended the ALBP method to an approach called adaptive local binary pattern with oriented standard deviation, which adds an oriented standard deviation term to the LBP operator instead of using this information in the matching. Excellent results were obtained when assessing boar sperm vitality.

In Ref. [48], LBP variance (LBPV) was proposed as a rotation-invariant descriptor. For LBPV, there are three stages:

1. using the local contrast (grayscale variance) as an adaptive weight to adjust the contribution of the LBP code in histogram calculation,
2. learning the principal directions: the extracted LBPV features are used to estimate the principal orientations, and then the features are aligned to the principal orientations, and
3. determining the nondominant patterns and thus by reducing them, feature dimension reduction was achieved.

Zhang et al. [49] proposed monogenic LBP, which integrates the traditional rotation-invariant LBP operator with two other rotation-invariant measures: the local phase and local surface type computed by the first- and second-order Riesz transforms, respectively. The local phase corresponds to a qualitative measure of local structure (step, peak, etc.), whereas the monogenic curvature tensor extracts local surface type information. Zhao et al. [50] extracted grayscale difference information but totally abandoned structural information by proposing a completed local binary count operator. Their results suggest that microstructure is not always needed for rotation-invariant classification.

Zhao et al. [51] proposed an approach to compute rotation-invariant features from histograms of local, noninvariant patterns. They applied this approach to both static and dynamic LBP descriptors. For static textures, they presented local binary pattern histogram Fourier (LBP-HF) features. LBP-HF is computed from discrete Fourier transforms of LBP histograms. The approach can also be generalized to embed any uniform features into this framework, and combining supplementary information, for example, sign and magnitude components of LBP, together can improve the description ability. Moreover, two variants of rotation-invariant LBP-HF descriptors were proposed for LBP-TOP, which is not rotation-invariant. Experiments showed that LBP-HF and its extensions perform very well in rotation-invariant texture classification. They are also robust with respect to changes in viewpoint, outperforming recent methods proposed for view-invariant recognition of dynamic textures.

Li et al. [52] considered scale- and rotation-invariant LBP. A circular neighboring set of a pixel is defined as a self-adaptive texton. The optimal scale of each pixel is adaptively selected based on the scale space derived by the Laplacian of Gaussian and used to determine the radius of the scale-adaptive texton. Different pixels have different optimal scales, resulting in scale invariance. The subuniform patterns of each uniform pattern are defined to improve the discrimination, and the circular shift LBP histogram is computed to obtain rotation invariance.

9.3.1.6 *Considering co-occurrences*

Recent studies show that encoding co-occurrences of LBPs can significantly improve the performance. Co-occurrences can be considered in different ways, within the same LBP operator, between adjacent operators, or at region level.

Qi et al. [53] introduced a pairwise rotation-invariant co-occurrence LBP (PRI-CoLBP), which incorporates two types of context: spatial co-occurrence and orientation co-occurrence. The method aims to preserve the relative angle between the orientations of individual features. The relative angle provides information about local curvature. For each co-occurrence pattern, the gradient magnitude of the two points is used to weight the copattern. Excellent results using a three-scale PRI-CoLBP are reported for many different datasets. The length of the feature vector (590 for the single scale operator and 6×590 when extracting six co-occurrence patterns in two angles and three scales) may limit the applicability of this and also many other co-occurrence methods in certain applications, such as face recognition. Qi et al. [54] also proposed a rotation-invariant multiscale joint encoding of LBP (MSJ-LBP) operator. In the original MLBP, each scale is encoded into histograms separately, and then the histograms are concatenated. This ignores the correlation between different scales. MSJ-LBP encodes jointly the LBPs of two scales around the center point to capture their correlation. For each point at two chosen scales, its joint MS-LBP pattern (one of the 590 patterns) and its gradient magnitude are computed. The gradient magnitude is used to weight the given joint pattern.

Nosaka et al. [55] proposed a co-occurrence among adjacent LBP (CoALBP) operator. The co-occurrence of adjacent LBPs is defined as an index of how often their combination occurs in the whole image (or region). Co-occurrence is measured with an autocorrelation matrix generated from multiple LBPs. Later, they extended it to rotation-invariant co-occurrence among adjacent LBP operator, which is enabled by introducing the concept of rotation equivalence class to CoALBP [56]. Louis and Plataniotis [57] considered co-occurrences of rotation-invariant LBPs (CoLBPs) at region level. Basic rotation-invariant LBPs are computed for all possible scales in the examined scanning window. Then, instead of computing histograms, multiple instances of rotation-invariant LBPs are selected using the sequential forward selection algorithm.

9.3.1.7 *Handling color*

LBP has also been widely applied to color images. To describe color and texture jointly, opponent color LBP (OCLBP) was defined in Ref. [58]. In OCLBP, the operator is used on each color channel independently and then for pairs of color channels so that the center pixel is taken from one channel and the neighboring pixels from the other. Opposing pairs, such as R-G and G-R, are highly redundant, so either of them can be used in the analysis. In total, six histograms (out of nine) are utilized (R, G, B, R-G, R-B, and G-B), making the descriptor six times longer than the monochrome LBP histogram. The OCLBP descriptor fares well in comparison with other color texture descriptors. It has been later used successfully, for example, for face recognition. However, the authors of this paper did not recommend joint color

and texture description as in their experiments: "all joint color texture descriptors and all methods of combining color and texture on a higher level are outperformed by either color or gray-scale texture alone."

A popular way is to apply the ordinary LBP to different color channels separately. Instead of the original R, G, and B channels, other more discriminative and invariant color features derived from them can be used for LBP feature extraction as well. Along this line, Zhu et al. [59] proposed multiscale color LBPs for visual object class recognition. Six operators were defined by applying MLBP on different types of channels and then concatenating the results. From these, the hue LBP (computed from the hue channel of the HSV color space), opponent LBP (computed over all three channels of the opponent color space), and an opponent LBP (computed over two channels of the normalized opponent space) provided good performance in experiments. Later, Zhu et al. [60] extended their approach to image region description using orthogonal combination of LBPs (OC-LBPs) and six new local descriptors based on OC-LBP enhanced with color information. OC-LBP is obtained by combining the histograms of $P/4$ different 4-orthogonal neighbor operators. The dimension of the descriptor is $4 \times P$, which is linear with the number of neighboring pixels in comparison to 2P for the original LBP. Color OC-LBP descriptor is obtained by concatenating OC-LBP histograms computed over selected color channels. Experiments on image matching, object recognition, and scene classification were used to show the effectiveness of the approach.

Qi et al. [61] presented an approach to encode cross channel texture correlation for color texture classification. The texture correlation between different RGB channels is first empirically studied using LBP as texture descriptor and Shannon's information entropy as correlation measurement. For color texture description, pairwise color channels are jointly encoded to obtain cross channel LBPs (CCLBPs). A multiscale CCLBP is also developed. Excellent results were reported.

Banerji et al. [62] present new image descriptors based on color, texture, shape, and wavelets for object and scene image classification. A three-dimensional LBP descriptor is proposed for encoding both color information and texture information and H-descriptor to integrate the 3-D LBP and the HOG of its wavelet transform, to encode color, texture, shape, and local information. The H-descriptor is comparatively assessed on seven well-known color spaces. A new H-fusion is also presented by fusing the principal component analysis (PCA) features of the H-descriptors in the seven color spaces. Experimental results on the Caltech 256 object categories, the UIUC sports event, and the MIT scene datasets demonstrate excellent performance.

9.3.1.8 Handling noise
Sensitivity to noise in images is one of the key problems of the original LBP. Kylberg and Sintorn [63] evaluated the noise robustness of eight LBP variants on five different datasets. None of the descriptors was generally more noise robust for all datasets and noise levels. However, LTP was often among the best performing descriptors, and improved LBP [23] often performed slightly better than the original LBP. Median LBP [22] and local quinary pattern [16] performed worst.

Chen et al. [64] proposed a simple robust version of LBP (i.e., RLBP) by changing the coding of bits of LBP, which could otherwise be changed by noise. Experimental results on texture datasets demonstrated that the RLBP outperforms many widely used descriptors and other variants of LBP, especially when noise is added in the images. Experimental results in face recognition also provided very promising results. Ren et al. [65] introduced a noise-resistant LBP aiming to preserve the local image structures in the presence of noise. A small pixel difference is vulnerable to noise, and thus it is first encoded as uncertain, and then its value is determined based on the other bits of LBP code to form a code of local uniform pattern. They also proposed extended noise-resistant LBP to capture line patterns. Experiments with added Gaussian and uniform noise on various datasets demonstrate the efficiency of their approach. Zhao et al. [66] proposed a completed robust local binary pattern, in which each center pixel is replaced with the average of the neighborhood, as in improved LBP. To make the operator more robust and stable, a weighted local gray level is also introduced to replace the value of the center pixel. The reported results for the UIUC database are clearly worse than those of Chen et al. [64].

9.3.1.9 Combining local and global information
LBP reflects the correlation among pixels within a local area (e.g., 3×3 area for $LBP_{8,1}$), which mainly represents the local information. Recently, there have been many works combining more global or more local information with LBP, for getting a more discriminative description from different feature levels.

Liao and Chung [67] extracted first dominant patterns, the so-called advanced LBPs from images, and labeled their locations with "1" and "0" otherwise. The gray-level aura matrix was used to extract the spatial information from each binary image. Guo et al. [68] investigated rotation-invariant image description with a linear model-based descriptor named MiC, which is suited to modeling microscopic configuration of images. To explore multichannel discriminative information on both the microscopic configuration and local structures, the feature extraction process is formulated as an unsupervised framework. It consists of (1) the configuration model to encode image microscopic configuration and (2) local patterns to describe local structural information. In this way, images are represented by a novel feature: local configuration pattern. Khellah [69] proposed a method that computes global rotation-invariant features from estimated dominant neighborhood structure and combines them with local LBP features, providing very good performance also in experiments with additive Gaussian noise.

Covariance matrices (CovMs) capture correlation among elementary features of pixels over an image region. Ordinary LBP features cannot be used as elementary features, since they are not numerical variables in Euclidean spaces. To address this problem, Hong et al. [70] developed a powerful descriptor, named COV-LBP. Firstly, a variant of LBPs in Euclidean spaces, named the LBP difference (LBPD) feature, was proposed. LBPD reflects how far one LBP lies from the LBP mean of a given image region. Secondly, applying LBPD together with some other features provided a bank of discriminative features for CovMs, providing very good performance in

experiments. Papagostas et al. [71] introduced moment-based LBPs. Their approach consists of two steps: the momentogram construction and the application of LBP on it. As a result, an enhanced LBP histogram is obtained, which is invariant under common geometric transformations (translation, rotation, and scaling), enclosing local and global information.

9.3.1.10 Complementary descriptors

A current trend in the development of new effective local image and video descriptors is to combine the strengths of complementary descriptors. From the beginning, the LBP operator was designed as a complementary measure of local image contrast. An interesting alternative for putting the local contrast into the one-dimensional LBP histogram was proposed by Guo et al. [48] in their LBPV method. As presented earlier, Guo et al. [10] also introduced the CLBP operator, in which the magnitude component is used as an improved contrast measure. Magnitude LBP and LBPV both contain supplementary information to LBP. They were embedded to the histogram Fourier framework [51] and concatenated to LBP-HF features as complementary descriptors to improve the description power for dealing with rotation variations.

In addition to applying LBP to Gabor-filtered face images, the joint use of LBP and Gabor and LBP and local phase quantization (LPQ) has provided excellent results in face recognition [12,72]. The HOG-LBP, combining LBP with the histogram of oriented gradients operator [73], has performed very well in human detection with partial occlusion handling [74]. Combining ideas from Haar and LBP features have given excellent results in accurate and illumination-invariant face detection [75]. A CS-LBP method for combining the strengths of SIFT and LBP in interest region description has also been developed [25].

The Weber law descriptor (WLD) is based on the fact that human perception of a pattern depends not only on the change of a stimulus (such as sound and lighting) but also on the original intensity of the stimulus [76]. WLD consists of two components: differential excitation and orientation. The differential excitation component is a function of the ratio between two terms: one is the relative intensity differences of a current pixel against its neighbors and the other is the intensity of the current pixel. The orientation component is the gradient orientation of the current pixel. For a given image, the two components are used to construct a concatenated WLD histogram. The joint use of LBP and the excitation component of WLD descriptor, together with the histogram of optic flow in dynamic texture segmentation, was considered by Chen et al. [77]. This indicates that the excitation component could be useful in replacing the contrast measure of LBP in other problems. The method of Liu et al. [78] also uses differential excitation of WLD together with LBP. The former is improved by using Laplacian of Gaussian.

Jun et al. [79] proposed local gradient patterns and binary histograms of oriented gradients and a hybrid feature that combines several local transform features using AdaBoost. Excellent results were reported in face and human detection experiments. In Nguyen et al. [80], contour templates representing object shape are used to derive a set of ("edge-like") key points at which local appearance features are extracted.

Both spatial information and orientation information are used in their computation. At each keypoint, nonredundant local binary pattern (NR-LBP) is computed. An object descriptor is formed by concatenating NR-LBP features from all keypoints. Very good performance is obtained for MIT and INRIA datasets. A downside of the method is that computation is quite time-consuming. The paper by Ma et al. [81] demonstrates that orientation information is critical in human detection. They proposed an oriented local binary pattern (OLBP) feature, which integrates pixel intensity difference information with texture orientation information to capture a salient object feature. Also, a set of edge orientation histograms and OLBP-based intrablock and interblock features is proposed to describe cell-level and block-level information. Experiments on INRIA and Caltech datasets demonstrate that the approach has a competitive performance and higher speed than existing detectors.

9.3.2 FEATURE SELECTION AND LEARNING
It has been shown by many studies that the dimensionality of the LBP distribution can be effectively decreased by reducing the number of neighboring pixels or by selecting a subset of bins available. In many cases, a properly chosen subset of LBP patterns can perform better than the whole set of patterns.

9.3.2.1 Rule-based selection
Already, the early studies on LBP indicated that some problems considering only four neighbors of the center pixel (i.e., 16 bins) can provide almost as good results as eight neighbors (256 bins). Mäenpää et al. [82] showed that a major part of the discriminative power lies in a small properly selected subset of patterns. In addition to the uniform patterns, they also considered a method based on beam search in which, starting from one, the size of the pattern set is iteratively increased up to a specified dimension D, and the best B pattern sets produced so far are always considered.

Smith and Windeatt [83] used the fast correlation-based filtering (FCBF) algorithm to select the most discriminative LBP patterns. FCBF operates by repeatedly choosing the feature that is most correlated with a given class (e.g., person identity in the case of face recognition), excluding those features already chosen or rejected and rejecting any features that are more correlated with it than with the class. As a measure of correlation, the information-theoretic concept of symmetric uncertainty is used. When applied to the LBP features, FCBF reduced their number from 107,000 to 120.

Liao et al. [84] introduced dominant local binary patterns that make use of the most frequently occurred patterns of LBP to improve the recognition accuracy compared to the original uniform patterns. The method has also rotation-invariant characteristics. To obtain discriminative patterns, Guo et al. [85] presented a learning model that is formulated into a three-layered model. It estimates the optimal pattern subset of interest by simultaneously considering the robustness, discriminative power, and representation capability of features. This model is generalized and can be integrated with existing LBP variants such as conventional LBP, rotation-invariant

patterns, local patterns with anisotropic structure, CLBP, and LTP to derive new image features for texture classification.

9.3.2.2 Boosting

Boosting has become a very popular approach for feature selection. It has been widely adopted for LBP feature selection in various tasks, for example, 3-D face recognition and face detection. AdaBoost is commonly used for selecting optimal LBP settings (such as the size and the location of local regions and the number of neighboring pixels) or for selecting the most discriminative bins of an LBP histogram. For instance, Zhang et al. [86] used AdaBoost learning for selecting an optimal set for local regions and their weights for face recognition. Since then, many related approaches have been used at region level for LBP-based face analysis. Shan and Gritti [87], on the other hand, used AdaBoost for learning discriminative LBP histogram bins, with an application to facial expression recognition.

Heng et al. [88] presented a "shrink boost" method for selecting features from multiple LBP histograms. Motivation for it was that feature selection from sparse and high-dimension features using conventional greedy-based boosting leads to poor generalization. The shrink boost method solves the sparse regularization problem with two iterative steps. A "boosting" step first uses weighted training samples to learn a full high-dimensional classifier on all features, avoiding overfitting to few features, and improves generalization. Then, a "shrinkage" step shrinks least discriminative classifier dimension to zero to remove the redundant features. In well-known (INRIA human detection, DaimlerChrysler pedestrian detection, and bird detection) object detection problems, they used "shrink boost" to select sparse features from concatenated LBP histograms of multiple quantization and image channels to learn classifier of additive lookup tables, obtaining improved generalization even under limited training samples.

9.3.2.3 Subspace learning

Another approach for deriving compact and discriminative LBP-based feature vectors consists of applying subspace methods for learning and projecting the LBP features from the original high-dimensional space into a lower-dimensional space. For instance, Chan et al. [72] used linear discriminant analysis (LDA) to project high-dimensional MLBP features into a discriminant space, yielding very good results. To deal with the small sample size problem of LDA, Shan et al. [89] constructed ensemble of piecewise Fisher discriminant analysis classifiers, each of which is designed based on one segment of the high-dimensional histogram of LGBP features. Their approach was shown to be more effective than applying LDA to high-dimensional holistic feature vectors.

Tan and Triggs [12] combined Gabor wavelets and LBP features and projected them to PCA space. Then, the kernel discriminative common vectors are applied to extract discriminant nonlinear compact features for face recognition. Zhao et al. [90] applied Laplacian PCA (LPCA) for LBP feature selection and pointed out the superiority of LPCA over PCA and KPCA for feature selection. Hussain and Triggs [91] exploited the complementarity of three sets of features, namely, HOG, LBP, and

LTP, and adopted partial least squares dimensionality reduction for selecting the most discriminative features, yielding fast and efficient visual object detector.

Nanni et al. [92] explored the use of random subspace, known to work well with noise and correlated features, to train features based also on nonuniform patterns. The approach fuses classifiers (SVM) trained considering the uniform patterns and random subspace classifiers trained considering only the nonuniform patterns.

9.3.2.4 Other methods
Ren et al. [93] found that most existing approaches rely on a predefined LBP structure to extract features and that those structures can be generalized as the patterns constructed from the binarized pixel differences in a local neighborhood. Instead of using a predefined structure, they learn binarized pixel-difference patterns (BPPs), casting the BPP structure discovery as a feature selection problem, which is solved via incremental minimal-redundancy-maximal-relevance algorithms. The method outperformed existing methods (e.g., CENTRIST [94]) in two scene recognition problems.

From the observation that LBP is equivalent to the application of a fixed binary decision tree, Maturana et al. [95] proposed a new method for learning discriminative LBP-like patterns from training data using decision tree induction algorithms. For each local image region, a binary decision tree is constructed from training data, thus obtaining an adaptive tree whose main branches are specially tuned to encode discriminative patterns in each region. Among the drawbacks of the proposed decision tree LBP is the high cost of constructing and storage of the decision trees especially when large pixel neighborhoods are used.

9.3.3 OTHER METHODS INSPIRED BY LBP
LBP has also inspired the development of new effective local image descriptors related to LBP.

The LPQ descriptor is based on quantizing the Fourier transform phase in local neighborhoods [96]. The phase can be shown to be a blur-invariant property under certain commonly fulfilled conditions. In texture analysis, histograms of LPQ labels computed within local regions are used as a texture descriptor. Generation of the labels and their histograms is similar to the LBP method. Extensions of LPQ to multiple scales [72], spatiotemporal domain [97], and color images have also been developed. The LPQ descriptor has recently received considerable interest in blur-invariant face recognition [72,96].

Lategahn et al. [98] developed a framework that filters a texture region by a set of filters and subsequently estimates the joint probability density functions by Gaussian mixture models (GMMs). Using the oriented difference filters of the LBP method [33], they showed that this method avoids the quantization errors of LBP, obtaining better results than with the basic LBP. Additional performance improvement of the GMM-based density estimator was obtained when the elementary filters were replaced by wavelet frame transform filter banks.

Vu and Caplier [99] proposed a feature descriptor called pattern of oriented edge magnitudes (POEMs) for face recognition and image matching. First, image gradient is computed, and then a histogram of orientations is accumulated and assigned to the CI of a cell (a spatial region around the current pixel). The gradient magnitude is used to weight the contribution of each pixel. Finally, the accumulated magnitudes are encoded using the LBP operator. Descriptors based on these principles performed very well in face recognition and image matching experiments.

Sharma et al. [100] employed local higher-order statistics (LHSs) of local nonbinarized patterns for image description. The LHS requires neither any user-specified quantization of the space of pixel patterns nor any heuristics for discarding low occupancy volumes of the space. Experiments with texture and face databases, with an SVM classifier, demonstrate very good performance. Zhang et al. [101] proposed local energy pattern (LeP) for texture classification using self-adaptive quantization thresholds. The method generates local feature vectors obtained by rectifying the responses of the 2-D Gaussian-like second derivative filters and then utilizes N-nary coding quantization instead of binary one. LEP was also extended to spatiotemporal analysis.

In Ref. [102], texture images are first decomposed by the shearlet transform, followed by construction of local energy features. These are then quantized and encoded to be rotation-invariant. The energy histograms accumulated over all decomposition levels reflect the different energy distributions. The method extracts more directional features like orientations and is robust with respect to noise. Experiments show very promising performance, especially with additive Gaussian noise.

Inspired by LBP, Maani et al. [103] introduced a method in which local frequency components are computed by applying 1-D Fourier transform on a neighboring function defined on a circle of radius R at each pixel. They observed that the low-frequency components are the major constituents of the circular functions and can effectively represent textures. Three sets of features were extracted from the low-frequency components, two based on the phase and one based on the magnitude. The method has advantages of a very good performance, relatively small number of features, and robustness to noise.

Wu and Rehg [94] proposed CENTRIST, a holistic *cen*sus *tr*ansform h*ist*ogram-based visual descriptor for recognizing places and scene categories. LBP-like census transform [104] histograms are used to encode structural properties within an image and suppress detailed textural information. Constraints between neighboring pixels are utilized to capture the structural characteristic within a small image patch. In larger scales, spatial hierarchy of CENTRIST is used to catch rough geometric information. A downside of the method is that it is not rotation-invariant.

Crosier and Griffin [105] proposed basic image features (BIFs) for texture classification. BIF engineers, like LBP, a dataset-independent dictionary of local features over which textures are represented statistically. Zeroth-, first-, and second-order Gaussian derivative filters are used for local description. Inspired by methods like LBP that produce binary codes, Kannala and Rahtu [106] proposed binarized

statistical image features (BSIFs). BSIF computes a binary code for each pixel by linearly projecting local image patches onto a subspace, whose basis vectors are learned from natural images via independent component analysis.

LBP has also given inspiration to recent interest on binary local feature descriptors, including BRIEF [107], ORB, and BRISK. The binary descriptors provide a comparable matching performance with the widely used interest region descriptors such as SIFT and SURF but have very fast extraction times and very low memory requirements needed, for example, in emerging applications using mobile devices with limited computational capabilities. Comparative evaluations of these descriptors can be found in Refs. [108,109].

9.4 SPATIOTEMPORAL AND OTHER DOMAINS
9.4.1 VARIANTS OF SPATIOTEMPORAL LBP

The high success of spatiotemporal LBP methods in various computer vision problems and applications has led to many other teams investigating the approach, and several extensions and modifications of spatiotemporal LBP have been proposed to increase its robustness and discriminative power.

Zhao and Pietikäinen [110] extended LBP-TOP to multiscale spatiotemporal space, with an application to facial expression recognition. AdaBoost was used to learn the principal appearance and motion, for selecting the most important expression-related features for all the classes or between every pair of expressions. Rotation-invariant variants of LBP-TOP based on histogram Fourier features were proposed by Zhao et al. [51].

LTP-TOP was developed by Nanni et al. [111]. The encoding function was modified for considering both the ternary patterns and the three orthogonal planes. Their experiments on a 10-class Weizmann dataset obtained very good results. WLD has also been extended to the spatiotemporal domain in the same way as LBP-TOP, yielding WLD-TOP for supplementing LBP-TOP in dynamic texture segmentation [77]. Mattivi and Shao [112] proposed Extended Gradient LBP-TOP for action recognition. Two modifications were made on the basis of LBP-TOP. Firstly, the computation of LBP was extended to nine slices, three for each axis. Therefore, on the *XY* dimension, there is the original *XY* plane (centered in the middle of the cuboid) plus two other *XY* planes located at 1/4 and 3/4 of the cuboid's length. The same is done for *XT* and *YT* dimensions. Secondly, computation of LBP operator on gradient images was introduced. The gradient image contains information about the rapidity of pixel intensity changes along a specific direction, has large magnitude values at edges, and can further increment LBP operator's performances, since LBP encodes local primitives such as curved edges, spots, and flat areas. For each cuboid, the brightness gradient is calculated along *x*, *y*, and *t* directions, and the resulting three cuboids containing specific gradient information are summed in absolute values. Before computing the image gradients, the cuboid is slightly smoothed with a Gaussian filter

in order to reduce noise. The extended LBP-TOP is then performed on the gradient cuboid. Experiments on a KTH human action dataset showed the effectiveness of the method.

For dealing with 2-D face recognition, Lei et al. [113] proposed effective LBP operator on three orthogonal planes of Gabor volume (E-GV-LBP). Firstly, the Gabor face images are formulated as a third-order Gabor volume. Then, LBP operator is applied on three orthogonal planes of Gabor volume, respectively, named GV-LBP-TOP in short. In this way, the neighboring changes both in spatial space and during different types of Gabor faces can be encoded. Moreover, in order to reduce the computational complexity, an effective GVLBP (E-GV-LBP) descriptor was developed that describes the neighboring changes according to the central point in spatial, scale, and orientation domains simultaneously for face representation.

Visual information from captured video is important for speaker identification under noisy conditions. Combination of LBP dynamic texture and EdgeMap structural features was proposed to take both motion and appearance into account [114], providing the description ability for spatiotemporal development in speech. Spatiotemporal dynamic texture features of LBPs extracted from localized mouth regions are used for describing motion information in utterances, which can capture the spatial and temporal transition characteristics. Structural edge map features are extracted from the image frames for representing appearance characteristics. Combination of dynamic texture and structural features takes both motion and appearance together into account, providing the description ability for spatiotemporal development in speech. In the experiments on BANCA and XM2VTS databases, the proposed method obtained promising recognition results compared to the other features.

Goswami et al. [115] proposed a novel approach to ordinal contrast measurement called local ordinal contrast patterns (LOCPs). Instead of computing the ordinal contrast with respect to any fixed value such as that at the center pixel or the average intensity value, it computes the pairwise ordinal contrasts for the chain of pixels representing the circular neighborhoods starting from the center pixel. Then, it was extended for dynamic texture analysis by extracting the LOCP in three orthonormal planes to generate LOCP-TOP. Together with LDA, its performance of mouth-region biometrics in the XM2VTS database received good results.

Spatial representation of LPQ was extended to a dynamic texture descriptor called the volume local phase quantization by Päivärinta et al. [97]. The local Fourier transform is computed by 1-D convolutions for each dimension in a 3-D volume. The data achieved are compressed to a smaller dimension before a scalar quantization procedure. Finally, a histogram of all code words from dynamic texture is formed.

Huang et al. [116] proposed to use spatiotemporal monogenic binary patterns to describe the appearance and motion information of the dynamic sequences. Firstly, they used monogenic signals analysis to extract the magnitude, the real picture, and the imaginary picture of the orientation of each frame, since the magnitude can provide much appearance information and the orientation can provide complementary information. Secondly, the phase-quadrant encoding method and the local

bit exclusive operator are utilized to encode the real and imaginary pictures from orientation in three orthogonal planes, and the LBP operator is used to capture the texture and motion information from the magnitude through three orthogonal planes. Finally, both the concatenation method and multiple kernel learning method are exploited to handle the feature fusion. The experimental results on the extended Cohn-Kanade and Oulu-CASIA facial expression databases demonstrated a state-of-the-art performance.

Ruiz-Hernandez and Pietikäinen [117] proposed a method for encoding LBPs using a reparametrization (RP) of the second-local order Gaussian jet. The information provided by RP generates robust and reliable histograms and is thus suitable for different facial analysis tasks. The proposed method has two main processes: the RP process that is used to compute needed parameters in a video sequence and the encoding process that combines the textural information provided by the LBP and the robustness of the RP. They showed that this approach can be used for recognizing facial microexpressions from videos, obtaining competitive performance on two different datasets.

Video texture synthesis is the process of providing a continuous and infinitely varying stream of frames. Guo et al. [118] proposed a variant of LBP-TOP called multiframe LBP-TOP to find the most appropriate matching pairs of frames for video texture synthesis. To achieve seamless synthesis results, a diffeomorphic growth model was applied to matching frames identified. The proposed approach has potential in other applications, for example, motion interpolation for videos and dynamic texture retrieval.

Inspired by spatiotemporal LBP, the Haar-like features were extended to represent the dynamic characteristic of facial expression [119]. The dynamic Haar-like features were built by two steps: (1) thousands of Haar-like features are extracted in each frame, and (2) the same Haar-like features in the consecutive frames are combined as dynamic features, as Figure 9.7 shows.

With the increasing amount of surveillance data, segmentation of moving objects in the compressed domain is receiving growing attention. Yang et al. [120] proposed a method in which the motion vectors are first accumulated and filtered to have
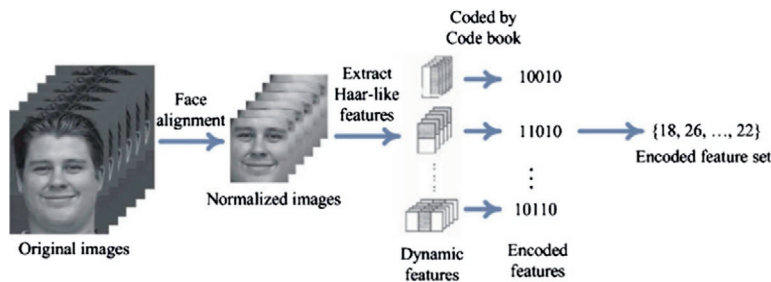
FIGURE 9.7

Dynamic Haar-like features.

reliable motion information. Then, spatiotemporal LBP features of motion vectors are extracted to obtain coarse and initial object regions in the H.264 compression domain. Final region refinement is done according to the distribution of DCT coefficients.

A local spatiotemporal directional descriptor was proposed for speaker identification by analyzing mouth movements [121]. The movements of mouth regions are described using LBPs and intensity contrast from six directions in three orthogonal planes. In addition, besides sign features, magnitude information encoded as weight for the bins with the same sign value was developed to improve the discriminative ability. Moreover, decorrelation is exploited to remove the redundancy of features. Experimental results on the challenging XM2VTS database showed the effectiveness.

9.4.2 LBP IN OTHER DOMAINS
9.4.2.1 Analysis of depth and 4-D images
Research on LBP variants for depth (range) images has focused on applications in face analysis. In most of the early papers, the depth values were first interpolated onto a regular grid, and then a more or less ordinary 2-D LBP approach was applied to the intensity (depth) values [122].

Huynh et al. [123] improved these by considering orientation differences in LBP computations, obtaining eight $LBP_{8,1}$-based orientations of the depth differences in a local 3×3 neighborhood. Eight depth difference histograms are obtained for the given region and concatenated to form an oriented histogram of depth differences. Sandbach et al. [124] extended the LBP methodology to 3-D face data by utilizing surface orientation information to detect facial action units (AUs). Local normal binary patterns employ the normals of the triangular polygons that form the 3-D face mesh to encode the shape of the mesh at each point. This is related to the gradient of a 2-D image, providing a richer source of local shape than the intensity alone. Bayramoglu et al. [125] also used surface orientation information in their method for facial AU detection. To get a short feature vector without degrading the performance, they used the CS-LBP approach with 3-D surface orientation information in computing the CS-3-D LBP operator. In Ref. [126], with the help of extracted keypoints, the 3-D face is divided into a mesh. LBP is computed using eight neighboring vertices of a current vertex. The depth and normal information of each vertex are extracted separately and encoded by LBP. The depth information describes the concavo-convex attribute, and the normal information indicates the face curvature variety.

Due to the progress of depth sensor technology, the analysis of 4-D images (i.e., depth images in motion) is an emerging research topic, for example, in facial expression recognition. Fang et al. [127] adopted the spatiotemporal LBP-TOP operator for 4-D facial expression recognition. A robust approach for registering consecutive frames of 4-D data is first applied, and the resulting "geometry" images are considered as frames in 2-D videos. LBP-TOP descriptors are computed on the difference between a frame and the first frame of the sequence, using the idea of the so-called flow image representing the deformation vectors in a subsequent frame w.r.t.

the first frame. Promising results are obtained for the BU-4DFE facial expression database. Reale et al. [128] used LBP-TOP as a comparative method in their paper on the 4-D spatiotemporal "Nebula" feature. The LBP-TOP approach, applied to 2-D and depth images, was in many ways similar to that of [127], but they did not use their alignment method and their tests on a flow image. The results for the Nebula feature were the best, but also, LBP-TOP performed reasonably well considering its simple way of implementation in this study.

9.4.2.2 Analysis of 3-D volume images

Extension of LBP to 3-D volume images is challenging. A circle in 2-D translates to a sphere in 3-D, and equidistant sampling on a sphere is not so trivial. The notion of ordering is also lost in 3-D due to the dimensionality [129].

Fehr and Burkhardt [130] extended the original LBP from 2-D images to 3-D volume data, achieving a full rotation invariance. For each LBP computation, the correlation between the values of all points on the neighborhood sphere with radius R and the weight factor that is a volume representation in an arbitrary but fixed order binomial factor is performed in the spherical harmonic domain. Rotation invariance is obtained from the computation of the minimum over all angles. Another method for rotationally invariant 3-D LBP using spectral harmonic decomposition was proposed by Banerjee et al. [129]. Unlike the original approach, the invariance is constructed implicitly, without considering all possible combinations of the pattern. Experiments were carried out with phantom data and clinical liver CTA data.

Morgado et al. [131] proposed an approach able to closely replicate in 3-D and without any approximation both uniformity and rotation invariance concepts originally proposed for the 2-D setting. Feature selection is done using correlation coefficients to quantify the relevance of each individual feature, and the SVM is used as the learning machine. The experiments demonstrate that the proposed method is able to enhance the diagnostic system and that the texture of the FDG-PET scans contains distinctive information about the presence of both Alzheimer's disease and mild cognitive impairment. Burner et al. [132] considered sign, contrast, and intensity information in computing texture bags at multiple scales for 3-D volumes. Instead of using histograms to match entire images, they search for regions that have similar local appearance. This is necessary to cope with the variability encountered in medical imaging data and the comparably subtle effects of disease. Based on the similarity of the texture structure of local regions, images are being ranked.

9.4.2.3 LBP in 1-D signal analysis

Analysis of 1-D signals is an emerging and potentially very important application area for LBP. For example, speech systems such as hearing aids require fast and inexpensive signal processing. Chatlani and Soraghan [133] proposed a straightforward simplification of the ordinary LBP to 1-D signals, with a preliminary application to simple signal segmentation and voice activity detection (VAD) to estimate periods of

speech and nonspeech. Later, Zhu et al. [134] used LBP-based VAD in HMM-based speech recognition. Speech is first denoised by adaptive empirical model decomposition and processed with LBP-based VAD, in which 1-D LBP is used to find start and end points of voiced speech segment so that it is distinguished from noise, unvoiced, or mute segments. In another application example, 1-D LBP is applied for bone texture classification by Houam et al. [135]. Global texture information is characterized by image projections, and local information is extracted from these using 1-D LBP.

Another interesting direction for 1-D signal analysis is to first derive a 2-D representation of the signal and then apply the spatial domain LBP to this representation. Lazic and Aarabi [136] successfully applied 2-D LBP for detecting spoken terms from visual spectrogram representation derived from the audio signal. A similar approach was used by Costa et al. [137] for classifying music genre from visual spectrograms derived from the audio signal. Esfahanian et al. [138], on the other hand, borrowing principles of facial image analysis [5], divided the visual spectrograms first into nonoverlapping regions and then used concatenated histograms to classify dolphin calls.

9.5 FUTURE CHALLENGES

Due to the considerable effort on spatial domain variants, future research should focus more on new challenges.

A great majority of the research on dense texture descriptors has been based on the assumption that there are no significant view or scale variations in the scene or objects to be analyzed. In practice, these variations are very common and may include self-occlusion when an object is imaged from different views. An obvious but impractical solution would be to train the recognition system with samples viewed from a large number of different positions and then derive a set of representative models for each class [7]. Moore and Bowden [139] showed that a larger-sized operator (Gabor LBP or MLBP) works better than small-sized operators in multiview facial expression recognition. It would very be useful to develop new approaches that are designed for handling effectively problems of view and scale variation.

For a large number of applications, an ability to analyze small sample sizes at high speed is vital, including face analysis, interest region description, segmentation, background subtraction, and tracking. This means that a compact region description with a short feature vector is needed. Many of the proposed descriptors would fail in this respect. It is important to evaluate the performance of a new descriptor also with smaller sample sizes, as was done, for example, in Ref. [29] by cropping 41×41 patches from the original 200×200 pixel CUReT textures.

There is also a need for more research on developing approaches that are robust to lighting variations. The use of Gabor filtering [11] or a preprocessing by Tan and Triggs [12] is an example of the state of the art for face recognition, but this approach

also tends to smooth small details from the images. Some of the recent works have successfully used image gradient orientation computation prior to feature extraction, as gradient orientations are known to be less sensitive to lighting variations than the original images.

Most of the past research has focused on improving the robustness of LBP for some specific tasks, like preprocessing prior to LBP computation (e.g., Gabor filtering and gradient computation), using different methods for encoding (e.g., LTP), reducing the number of neighbors in multiscale analysis (e.g., LQP), and processing LBP codes in different ways (e.g., LBP-HF). A future direction would be to consider different stages of feature computation together to optimize the performance. A good example of this direction is by Lei et al. [140], who obtained excellent results in face recognition by combining preprocessing by learning-based linear filtering, soft sampling to find the best set of neighbors for LBP computation, and clustering to find the optimal way for encoding.

Combining different types of complementary local operators is another way to go ahead, as, for example, the recent works on human detection and face recognition show. Approaches combining the strengths of LBP and HOG (or SIFT), for example, have led to increased performance.

Research on spatiotemporal LBP variants has not been as active as in the spatial domain. One would expect to see much more research in this area. Different effective ways of obtaining information in the temporal domain combined with novel ideas proposed for analyzing images in the spatial domain could be one way to progress. The robustness of spatiotemporal operators to different variations is largely unexplored. One example of this kind of work is by Zhao et al. [51], who demonstrated that their spatiotemporal LBP-HF features are robust with respect to view variations.

The use of LBP for analyzing 1-D signals has not been much studied, but has potential for a large number of novel applications, as the recent examples presented above demonstrate. In these works, quite primitive LBP-based analysis was used both for 1-D signal analysis and for 2-D analysis of spectrograms derived from 1-D signals. Borrowing ideas from the recent developments in spatial domain LBP variants could lead to much more powerful methods. The results on analyzing 2-D spectrograms also suggest that LBP variants could be very useful in a wide variety of applications of exploratory data analysis, in which relations between neighboring data elements should be considered.

With the introduction of Kinect and emerging more accurate sensors for depth sensing, the interest on processing depth images (3-D) and depth image sequences (4-D) has been growing, with applications, for example, in action and gesture recognition, face recognition, and recognition of facial AUs. Some LBP variants for depth images have been developed especially for face analysis, but research on 4-D data is largely unexplored. There is also much space for new developments in 3-D volume image analysis widely used in medicine. Simple 3-D variants of ordinary LBP have provided very promising improvements, but much more could be possible with more sophisticated solutions.

The databases used for comparing texture descriptors should be reconsidered. A good practice in some recent papers is to use both texture and face databases and datasets from other application domains. For 2-D faces, there are many good options available like Labeled Faces in the Wild and Face Recognition Grand Challenge (FRGC v2.0), representing different challenges. A problem with most of the used texture databases is that the best methods currently obtain over 95% accuracy. KTH-TIPS-2 represents a more difficult problem in this respect. It would be valuable to have extensive benchmarks with challenging databases for the most promising variants, as was done for texture retrieval in Ref. [141].

9.6 CONCLUSIONS

Due to its advantages, that is, flexibility, invariance to monotonic gray-level changes, and computational simplicity, LBP is a very powerful descriptor to represent local structures in images. A large number of variants have been designed aiming to obtain improved performance and/or robustness in one or more aspects of the original LBP. We have divided the extensions and modifications into 10 categories and introduced some representative variants in each of them.

A number of new effective local image descriptors, including LPQ, POEMs, and LEPs, have also been inspired by LBP. They provide the ability, for example, to deal with blur or noise and can be jointly used with LBP to complement each other.

Even though it is commonly agreed that multiscale analysis and joint use of complementary descriptors can improve the performance, a downside is the large dimensionality of the produced feature vector. To obtain a small set of the most discriminative LBP-based features, different feature selection and learning strategies were also discussed. Extensions of LBP to spatiotemporal domain extend the applicability of LBP from static images to dynamic video sequences. Different variants of the original spatiotemporal LBP were introduced.

Moreover, the future challenges of LBP were discussed. Among these are an improved robustness to view, scale, and lighting variations, optimization of different stages of feature computation, and combinations of different descriptors. There is also a need for further research on problems such as spatiotemporal LBPs and analysis of 1-D signals, depth, 3-D volume images, and 4-D depth image sequences.

REFERENCES

[1] T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture measures with classification based on feature distributions, Pattern Recogn. 29 (1) (1996) 51–59.
[2] T. Ojala, M. Pietikäinen, Unsupervised texture segmentation using feature distributions, Pattern Recogn. 32 (1999) 477–486.
[3] T. Ojala, K. Valkealahti, E. Oja, M. Pietikäinen, Texture discrimination with multidimensional distributions of signed gray-level differences, Pattern Recogn. 34 (3) (2001) 727–739.

[4] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Anal. Mach. Intell. 24 (7) (2002) 971–987.

[5] T. Ahonen, A. Hadid, M. Pietikäinen, Face description with local binary patterns: application to face recognition, IEEE Trans Pattern Anal. Mach. Intell. 28 (12) (2006) 2037–2041.

[6] G. Zhao, M. Pietikäinen, Dynamic texture recognition using local binary patterns with an application to facial expressions, IEEE Trans. Pattern Anal. Mach. Intell. 29 (6) (2007) 915–928.

[7] M. Pietikäinen, A. Hadid, G. Zhao, T. Ahonen, Computer Vision Using Local Binary Patterns, Springer, Berlin, Heidelberg, 2011.

[8] D. Huang, C. Shan, A. Mohsen, Y. H. Wang, L. Chen, Local binary patterns and its applications on facial image analysis: a survey, IEEE Trans. Syst. Man Cybern. C Appl. Rev. 41 (6) (2011) 765–781.

[9] S. Brahnam, L. C. Jain, L. Nanni, A. Lumini (Eds.), Local Binary Patterns: New Variants and Applications, Springer, Berlin, Heidelberg, 2013.

[10] Z. Guo, L. Zhang, D. Zhang, A completed modeling of local binary pattern operator for texture classification, IEEE Trans. Image Process. 19 (2010) 1657–1663.

[11] W. Zhang, S. Shan, W. Gao, X. Chen, H. Zhang, Local Gabor binary pattern histogram sequence (LGBPHS): a non-statistical model for face representation and recognition, in: Proceedings of IEEE International Conference on Computer Vision, vol. 1, 2005, pp. 786–791.

[12] X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, IEEE Trans. Image Process. 19 (6) (2010) 1635–1650.

[13] C. H. Yao, S. Y. Chen, Retrieval of translated, rotated and scaled color textures, Pattern Recogn. 36 (4) (2003) 913–929.

[14] X. Li, W. Hu, Z. Zhang, H. Wang, Heat kernel based local binary pattern for face representation, IEEE Signal Process. Lett. 17 (2010) 308–311.

[15] S. Liao, A. C. S. Chung, Face recognition by using elongated local binary patterns with average maximum distance gradient magnitude, in: Proceedings of Asian Conference on Computer Vision, 2007, pp. 672–679.

[16] L. Nanni, A. Lumini, S. Brahnam, Local binary patterns variants as texture descriptors for medical image analysis, Artif. Intell. Med. 49 (2010) 117–125.

[17] Y. He, N. Sang, C. Gao, Multi-structure local binary patterns for texture classification, Pattern Anal. Appl. 16 (4) (2012) 595–607.

[18] L. Wolf, T. Hassner, Y. Taigman, Effective unconstrained face recognition by combining multiple descriptors and learned background statistics, IEEE Trans. Pattern Anal. Mach. Intell. 33 (10) (2011) 1978–1990.

[19] S. A. Orjuela, J. P. Yanez Puentes, P. Philips, The geometric local texture patterns (GLTP), in: S. Brahnam, L. C. Jain, L. Nanni, A. Lumini (Eds.), Local Binary Patterns: New Variants and Applications, Springer, Berlin, Heidelberg, 2013, pp. 85–112.

[20] K. Wang, C.-E. Bichot, C. Chu, B. Li, Pixel to patch sampling structure and local neighboring intensity relationship patterns for texture classification, IEEE Signal Process. Lett. 20 (9) (2013) 853–856.

[21] J. Ylioinas, A. Hadid, Y. Guo, M. Pietikäinen, Efficient image appearance description using dense sampling based local binary patterns, in: Proceedings of Asian Conference on Computer Vision, 2013, pp. 375–388.

[22] A. Hafiane, G. Seetharam, B. Zavidovique, Median binary pattern for texture classification, in: Proceedings of International Conference on Image Analysis and Recognition, 2007, pp. 387–398.

[23] H. Jin, Q. Liu, H. Lu, X. Tong, Face detection using improved LBP under Bayesian framework, in: Proceedings of International Conference on Image and Graphics, 2004, pp. 306–309.

[24] M. Heikkilä, M. Pietikäinen, A texture-based method for modeling the background and detecting moving objects, IEEE Trans. Pattern Anal. Mach. Intell. 28 (4) (2006) 657–662.

[25] M. Heikkilä, M. Pietikäinen, C. Schmid, Description of interest regions with local binary patterns, Pattern Recogn. 42 (3) (2009) 425–436.

[26] T. Ahonen, M. Pietikäinen, Soft histograms for local binary patterns, in: Proceedings of Finnish Signal Processing Symposium, 2007, 4 p.

[27] S. Katsigiannis, E. Keramidas, D. Maroulis, FLBP: fuzzy local binary patterns, in: S. Brahnam, L. C. Jain, L. Nanni, A. Lumini (Eds.), Local Binary Patterns: New Variants and Applications, Springer, Berlin, Heidelberg, 2013, pp. 149–175.

[28] S. Liao, G. Zhao, V. Kellokumpu, M. Pietikäinen, S. Z. Li, Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2010, 8 p.

[29] J. Ylioinas, X. Hong, M. Pietikäinen, Constructing local binary pattern statistics by soft voting, in: Proceedings of Scandinavian Conference on Image Analysis, 2013, pp. 119–130.

[30] J. Trefny, J. Matas, Extended set of local binary pattern for rapid object detection, in: Proceedings of Computer Vision Winter Workshop, 2010.

[31] L. Liu, L. Zhao, Y. Long, G. Kuang, P. Fieguth, Extended local binary patterns for texture classification, Image Vision Comput. 30 (2012) 86–99.

[32] B. Zhang, Y. Gao, S. Zhao, J. Liu, Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor, IEEE Trans. Image Process. 19 (2) (2010) 533–544.

[33] T. Ahonen, M. Pietikäinen, Image description using joint distribution of filter bank responses, Pattern Recogn. Lett. 34 (2009) 368–376.

[34] Z. Guo, Q. Li, J. You, D. Zhang, W. Liu, Local directional derivative pattern for rotation invariant texture classification, Neural Comput. Appl. 8 (21) (2012) 1893–1904.

[35] S. ul Hussain, B. Triggs, Visual recognition using local quantized patterns, in: Proceedings of European Conference on Computer Vision, Part II, 2012, pp. 716–729.

[36] X. Huang, G. Zhao, X. Hong, M. Pietikäinen, W. Zheng, Texture description with completed local quantized patterns, in: Proceedings of Scandinavian Conference on Image Analysis, 2013, pp. 1–10.

[37] N. Jiang, J. Xu, S. Goto, Pedestrian detection using gradient local binary patterns, IEICE Trans. Fund. E95-A (8) (2012) 1280–1287.

[38] T. Mäenpää, M. Pietikäinen, Multi-scale binary patterns for texture analysis, in: Proceedings of Scandinavian Conference on Image Analysis, 2003, pp. 885–892.

[39] S. Liao, X. Zhu, Z. Lei, L. Zhang, S. Z. Li, Learning multi-scale block local binary patterns for face recognition, in: Proceedings of International Conference on Biometrics, 2007, pp. 828–837.

[40] M. Turtinen, M. Pietikäinen, Contextual analysis of textured scene images, in: Proceedings of British Machine Vision Conference, 2006, pp. 849–858.

[41] Y. He, N. Sang, C. Gao, Pyramid-based multi-structure local binary pattern for texture classification, in: Proceedings of Asian Conference on Computer Vision, vol. 3, 2010, pp. 1435–1446.

[42] X. Qian, X. Hua, P. Chen, L. Ke, PLBP: an effective local binary patterns texture descriptor with pyramid representation, Pattern Recogn. 44 (10/11) (2011) 2502–2515.

[43] T. Song, H. Li, WaveLBP based features for image classification, Pattern Recognit. Lett. 34 (2013) 1323–1328.

[44] L. Liu, Y. Long, P. Fieguth, S. Lao, G. Zhao, BRINT: binary rotation invariant and noise tolerant texture classification, IEEE Trans. Image Process. 23 (7) (2014) 3071–3084.

[45] M. Pietikäinen, T. Ojala, Z. Xu, Rotation-invariant texture classification using feature distributions, Pattern Recogn. 33 (2000) 43–52.

[46] Z. H. Guo, L. Zhang, D. Zhang, Rotation invariant texture classification using adaptive LBP with directional statistical features, in: Proceedings of International Conference on Image Processing, 2010, pp. 285–288.

[47] O. Garcia-Olalla, E. Alegre, R. Fernandez-Robles, M. T. Garcia-Ordas, Vitality assessment of boar sperm using an adaptive LBP based on oriented deviation, in: Proceedings of ACCV 2012 Workshops, 2013, pp. 61–72.

[48] Z. H. Guo, L. Zhang, D. Zhang, Rotation invariant texture classification using LBP variance (LBPV) with global matching, Pattern Recogn. 43 (3) (2010) 706–719.

[49] L. Zhang, D. Zhang, Z. Guo, D. Zhang, Monogenic-LBP: a new approach for rotation invariant texture classification, in: Proceedings of International Conference on Image Processing, 2010, pp. 2677–2770.

[50] Y. Zhao, D.-S. Huang, W. Jia, Completed local binary count for rotation invariant texture classification, IEEE Trans. Image Process. 21 (10) (2012) 4492–4497.

[51] G. Zhao, T. Ahonen, J. Matas, M. Pietikäinen, Rotation-invariant image and video description with local binary pattern features, IEEE Trans. Image Process. 21 (4) (2012) 1465–1477.

[52] Z. Li, G. Liu, Y. Yang, J. You, Scale- and rotation-invariant local binary pattern using scale-adaptive texton and subuniform-based circular shift, IEEE Trans. Image Process. 21 (4) (2012) 2130–2140.

[53] X. Qi, R. Xiao, C.-G. Li, Y. Qiao, J. Guo, X. Tang, Pairwise rotation invariant co-occurrence local binary pattern, IEEE Trans. Pattern Anal. Mach. Intell. 36 (11) (2014) 2199–2213.

[54] X. Qi, Y. Qiao, C.-G. Li, J. Guo, Multi-scale joint encoding of local binary patterns for texture and material classification, in: Proceedings of British Machine Vision Conference, 2013.

[55] R. Nosaka, Y. Ohkava, K. Fukui, Feature extraction based on co-occurrence of adjacent local binary patterns, in: Proceedings of PSIVT, Part II, 2011, pp. 82–91.

[56] R. Nosaka, C. H. Suryanto, K. Fukui, Rotation-invariant co-occurrence among adjacent LBPs, in: Proceedings of ACCV Workshops, 2013, pp. 15–25.

[57] W. Louis, K. N. Plataniotis, Co-occurrence of local binary patterns features for frontal face detection in surveillance applications, EURASIP J. Image Video Process. (2011), doi:10.1155/2011/745487.

[58] T. Mäenpää, M. Pietikäinen, Classification with color and texture: jointly or separately? Pattern Recogn. 37 (2004) 1629–1640.

[59] C. Zhu, C. E. Bichot, L. Chen, Multi-scale color local binary patterns for visual object classes recognition, in: Proceedings of International Conference on Pattern Recognition, 2010, pp. 3065–3068.

[60] C. Zhu, C.-E. Bichot, L. Chen, Image region description using orthogonal combination of local binary patterns enhanced with color information, Pattern Recogn. 46 (2013) 1949–1963.

[61] X. Qi, Y. Qiao, C.-G. Li, J. Guo, Exploring cross-channel texture correlation for color texture classification, in: Proceedings of British Machine Vision Conference, 2013.

[62] S. Banerji, A. Sinha, C. Liu, New image descriptors based on color, texture, shape, and wavelets for object and scene image classification, Neurocomputing 117 (2013) 173–185.

[63] G. Kylberg, I.-M. Sintorn, Evaluation of noise robustness for local binary pattern descriptors in texture classification, EURASIP J. Image Video Process. 17 (2013), 1-20.

[64] J. Chen, V. Kellokumpu, G. Zhao, M. Pietikäinen, RLBP: robust local binary pattern, in: Proceedings of British Machine Vision Conference, 2013.

[65] J. Ren, X. Jiang, J. Yuan, Noise-resistant local binary pattern with an embedded error-correction mechanism, IEEE Trans. Image Process. 22 (10) (2013) 4049–4060.

[66] Y. Zhao, W. Jia, R.-X. Hu, H. Min, Completed robust local binary pattern for texture classification, Neurocomputing 106 (2013) 68–76.

[67] S. Liao, A. C. S. Chung, Texture classification by using advanced local binary patterns and spatial distribution of dominant patterns, in: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 1, 2007, pp. 1221–1224.

[68] Y. Guo, G. Zhao, M. Pietikäinen, Texture classification using a linear configuration model based descriptor, in: Proceedings of British Machine Vision Conference, 2011.

[69] F. M. Khellah, Texture classification using dominant neighborhood structure, IEEE Trans. Image Process. 20 (11) (2011) 3270–3279.

[70] X. Hong, G. Zhao, M. Pietikäinen, X. Chen, Combining LBP difference and feature correlation for texture description, IEEE Trans. Image Process. 23 (6) (2014) 2557–2568.

[71] G. A. Papagostas, E. D. E. Koulouriotis, E. Karakasis, V. D. Tourassis, Moment-based local binary patterns: a novel descriptor for invariant pattern recognition applications, Neurocomputing 99 (2013) 358–371.

[72] C. H. Chan, M. A. Tahir, J. Kittler, M. Pietikäinen, Multiscale local phase quantization for robust component-based face recognition using kernel fusion of multiple descriptors, IEEE Trans. Pattern Anal. Mach. Intell. 35 (5) (2013) 1164–1177.

[73] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, in: Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition, vol. 2, 2005, pp. 886–893.

[74] X. Wang, T. X. Han, S. Yan, An HOG-LBP human detector with partial occlusion handling, in: Proceedings of IEEE International Conference on Computer Vision, 2009, pp. 32–39.

[75] S. Yan, S. Shan, X. Chen, W. Gao, Locally assembled binary (LAB) feature with feature-centric cascade for fast and accurate face detection, in: Proceedings of IEEE International Conference on Computer Vision, 2008, 8 p.

[76] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikäinen, X. Chen, W. Gao, WLD: a robust local image descriptor, IEEE Trans. Pattern Anal. Mach. Intell. 32 (9) (2010) 1705–1720.

[77] J. Chen, G. Zhao, M. Salo, E. Rahtu, M. Pietikäinen, Automatic dynamic texture segmentation using local descriptors and optical flow, IEEE Trans. Image Process. 22 (1) (2013) 326–339.

[78] F. Liu, Z. Tang, J. Tang, WLBP: Weber local binary pattern for local image description, Neurocomputing 120 (2013) 325–335.

[79] B. Jun, I. Choi, D. Kim, Local transform features and hybridization for accurate face and human detection, IEEE Trans. Pattern Anal. Mach. Intell. 35 (8) (2013) 1423–1436.

[80] D. T. Nguyen, P. O. Ogunbona, W. Li, A novel shape-based non-redundant local binary pattern descriptor for object detection, Pattern Recogn. 46 (2013) 1485–1500.

[81] Y. Ma, L. Deng, X. Chen, N. Guo, Integrating orientation cue with EOH-OLBP based multi-level features for human detection, IEEE Trans. Circ. Syst. Video Tech. 23 (10) (2013) 1755–1766.

[82] T. Mäenpää, T. Ojala, M. Pietikäinen, M. Soriano, Robust texture classification by subsets of local binary patterns, in: Proceedings of International Conference on Pattern Recognition, 2000, pp. 947–950.

[83] R. S. Smith, T. Windeatt, Facial expression detection using filtered local binary pattern features with ECOC classifiers and Platt scaling, in: Proceedings of JMLR Workshop on Applications of Pattern Analysis, vol. 11, 2010, pp. 111–118.

[84] S. Liao, M. W. K. Law, A. C. S. Chung, Dominant local binary patterns for texture classification, IEEE Trans. Image Process. 18 (5) (2009) 1107–1118.

[85] Y. Guo, G. Zhao, M. Pietikäinen, Discriminative features for texture description, Pattern Recogn. 45 (10) (2012) 3834–3843.

[86] G. Zhang, X. Huang, S. Li, Y. Wang, X. Wu, Boosting local binary pattern (LBP)-based face recognition, in: Proceedings of Advances in Biometric Authentication, 2005, pp. 179–186.

[87] C. Shan, T. Gritti, Learning discriminative LBP-histogram bins for facial expression recognition, in: Proceedings of British Machine Vision Conference, 2008, 10 p.

[88] C. K. Heng, S. Yokomitsu, Y. Matsumoto, H. Tamura, Shrink boost for selecting multi-LBP histogram features in object detection, in: Proceedings of International Conference on Computer Vision and Pattern Recognition, 2012, pp. 3250–3257.

[89] S. Shan, W. Zhang, Y. Su, X. Chen, W. Gao, Ensemble of piecewise FDA based on spatial histogram of local (Gabor) binary patterns for face recognition, in: Proceedings of International Conference on Pattern Recognition, vol. 4, 2006, pp. 606–609.

[90] D. Zhao, Z. Lin, X. Tang, Laplacian PCA and its applications, in: Proceedings of International Conference on Computer Vision, 2007, 8 p.

[91] S. ul Hussain, B. Triggs, Feature sets and dimensionality reduction for visual object detection, in: Proceedings of British Machine Vision Conference, 2010, 10 p.

[92] L. Nanni, S. Brahnam, A. Lumini, A simple method for improving local binary patterns by considering non-uniform patterns, Pattern Recogn. 45 (2012) 3844–3851.

[93] J. Ren, X. Jiang, J. Yuan, Learning binarized pixel-difference pattern for scene recognition, in: Proceedings of International Conference on Image Processing, 2013, 5 p.

[94] J. Wu, J. H. Rehg, CENTRIST: a visual descriptor for scene categorization, IEEE Trans. Pattern Anal. Mach. Intell. 33 (8) (2011) 1489–1501.

[95] D. Maturna, A. Soto, D. Mery, Face recognition with decision tree-based local binary patterns, in: Proceedings of Asian Conference on Computer Vision, 2010.

[96] E. Rahtu, J. Heikkilä, V. Ojansivu, T. Ahonen, Local phase quantization for blur-insensitive image analysis, Image Vis. Comput. 30 (8) (2012) 501–512.

[97] J. Päivärinta, E. Rahtu, J. Heikkilä, Volume local phase quantization for blur-insensitive dynamic texture classification, in: Proceedings of Scandinavian Conference on Image Analysis, 2011, pp. 360–369.

[98] H. Lategahn, S. Gross, S. Stehle, T. Aach, Texture classification by modeling joint distributions of local patterns with Gaussian mixtures, IEEE Trans. Image Process. 19 (6) (2010) 1548–1557.

[99] N.-S. Vu, A. Caplier, Enhanced patterns of oriented edge magnitudes for face recognition and image matching, IEEE Trans. Image Process. 21 (3) (2012) 1352–1365.

[100] G. Sharma, S. ul Hussain, F. Jurien, Local higher-order statistics (LHS) for texture categorization and facial analysis, in: Proceedings of European Conference on Computer Vision, Part VII, 2012, pp. 1–12.

[101] J. Zhang, J. Liang, H. Zhao, Local energy pattern for texture classification using self-adaptive quantization thresholds, IEEE Trans. Image Process. 22 (1) (2013) 31–42.

[102] J. He, H. Ji, X. Yang, Rotation invariant texture descriptor using local shearlet-based energy histograms, IEEE Signal Process. Lett. 30 (9) (2013) 905–908.

[103] R. Maani, S. Kalra, Y.-H. Yang, Rotation invariant local frequency descriptors for texture classification, IEEE Trans. Image Process. 22 (6) (2013) 2409–2419.

[104] R. Zabih, J. Woodfill, Non-parametric local transforms for computing visual correspondences, in: Proceedings of the Third European Conference on Computer Vision, Stockholm, vol. 2, May 1994, pp. 151–158.

[105] M. Crosier, L. D. Griffin, Using basic image features for texture classification, Int. J. Comput. Vision 88 (2010) 447–460.

[106] J. Kannala, E. Rahtu, BSIF: binarized statistical image features, in: Proceedings of International Conference on Pattern Recognition, 2012, pp. 1363–1366.

[107] M. Calonder, V. Lepetit, M. Özuysal, T. Trzcinski, C. Strecha, P. Fua, BRIEF: computing a local binary descriptor very fast, IEEE Trans. Pattern Anal. Mach. Intell. 34 (7) (2012) 1281–1298.

[108] J. Heinly, E. Dunn, J.-M. Frahm, Comparative evaluation of binary features, in: Proceedings of European Conference on Computer Vision, Part II, 2012, pp. 759–773.

[109] O. Miksik, K. Mikolajczyk, Evaluation of local detectors and descriptors for fast feature matching, in: Proceedings of International Conference on Pattern Recognition, 2012, pp. 2681–2684.

[110] G. Zhao, M. Pietikäinen, Boosted multi-resolution spatiotemporal descriptors for facial expression recognition, Pattern Recognit. Lett. 30 (12) (2009) 1117–1127.

[111] L. Nanni, S. Brahnam, A. Lumini, Local ternary patterns from three orthogonal planes for human action classification, Expert Syst. Appl. 38 (2011) 5125–5128.

[112] R. Mattivi, L. Shao, Human action recognition using LBP-TOP as sparse spatio-temporal feature descriptor, in: Proceedings of International Conference on Computer Analysis of Images and Patterns, 2009, pp. 740–747.

[113] Z. Lei, S. Liao, M. Pietikäinen, S. Z. Li, Face recognition by exploring information jointly in space, scale and orientation, IEEE Trans. Image Process. 20 (1) (2011) 247–256.

[114] G. Zhao, X. Huang, Y. Gizatdinova, M. Pietikäinen, Combining dynamic texture and structural features for speaker identification, in: Proceedings of ACM Multimedia Workshop Multimedia in Forensics, Security and Intelligence, 2010, pp. 93–98.

[115] B. Goswami, C. H. Chan, J. Kittler, B. Christmas, Local ordinal contrast pattern histograms for spatio-temporal, lip-based speaker authentication, in: Proceedings of IEEE International Conference on Biometrics: Theory, Applications and Systems, 2010, 6 p.

[116] X. Huang, G. Zhao, M. Pietikäinen, W. Zheng, Spatiotemporal local monogenic binary patterns for facial expression recognition, IEEE Signal Process. Lett. 19 (5) (2012) 243–246.

[117] J. Ruiz-Hernandez, M. Pietikäinen, Encoding local binary patterns using the re–parametrization of the second order Gaussian jet, in: Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition, 2013, 6 p.

[118] Y. Guo, G. Zhao, Z. Zhou, M. Pietikäinen, Video texture synthesis with multi-frame LBP-TOP and diffeomorphic growth model, IEEE Trans. Image Process. 22 (10) (2013) 3879–3891.

[119] P. Yang, Q. Liu, D. Metaxas, Boosting encoded dynamic features for facial expression recognition, Pattern Recognit. Lett. 30 (2) (2009) 132–139.

[120] J. Yang, S. Wang, Z. Lei, Y. Zhao, S. Z. Li, Spatio-temporal LBP based moving object segmentation in compressed domain, in: Proceedings of IEEE International Conference on Advanced Video and Signal-Based Surveillance, 2012, pp. 252–257.

[121] G. Zhao, M. Pietikäinen, Visual speaker identification with spatiotemporal directional features, in: Proceedings of International Conference on Image Analysis and Recognition, 2013, pp. 1–10.

[122] Y. Huang, Y. Wang, T. Tan, Combining statistics of geometrical and correlative features for 3D face recognition, in: Proceedings of British Machine Vision Conference, 2006, pp. 879–888.

[123] T. Huynh, R. Min, J.-L. Dugelay, An efficient LBP-based descriptor for facial depth images applied to gender recognition using RGB-D face data, in: Proceedings of ACCV Workshops, 2013, pp. 133–145.

[124] G. Sandbach, S. Zafeiriou, M. Pantic, Local normal binary patterns for 3D facial action unit detection, in: Proceedings of IEEE International Conference on Image Processing, 2012.

[125] N. Bayramoglu, G. Zhao, M. Pietikäinen, CS-3DLBP and geometry based person independent 3D facial action unit detection, in: Proceedings of IEEE/IAPR International Conference on Biometrics, 2013.

[126] H. Tang, B. Yin, Y. Sun, Y. Hu, 3D face recognition using local binary patterns, Signal Proc. 93 (2013) 2190–2198.

[127] T. Fang, X. Zhao, S. K. Shah, I. Kakadiaris, 4D facial expression recognition, in: Proceedings of ICCV Workshops, 2011, pp. 1594–1601.

[128] M. Reale, X. Zhang, L. Yin, Nebula feature: a space-time feature for posed and spontaneous 4D facial behavior analysis, in: Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition, 2013, 8 p.

[129] J. Banerjee, A. Moelker, W. J. Niessen, T. van Walsum, 3D LBP-based rotationally invariant region description, in: Proceedings of ACCV 2012 Workshops, 2013, pp. 26–37.

[130] J. Fehr, H. Burkhardt, 3D rotation invariant local binary patterns, in: Proceedings of International Conference on Pattern Recognition, 2008, 4 p.

[131] P. Morgado, M. Silveira, J. S. Marques, Diagnosis of Alzheimers disease using 3D local binary patterns, Comput. Methods Biomech. Biomed. Eng Imaging Vis. 1 (1) (2013) 2–12.

[132] A. Burner, R. Donner, M. Mauerhoefer, M. Holzer, F. Kainberger, G. Langs, Texture bags: anomaly retrieval in medical images based on local 3D-texture similarity, in: Proceedings of MICCAI, 2011.

[133] N. Chatlani, J. J. Soraghan, Local binary patterns for 1-D signal processing, in: Proceedings of European Signal Processing Conference, 2010, pp. 95–99.

[134] Q. Zhu, N. Chatlani, J. S. Soraghan, 1-D local binary patterns based VAD used in HMM-based improved speech recognition, in: Proceedings of European Signal Processing Conference, 2012, pp. 1633–1637.

[135] L. Houam, A. Hafiane, A. Boukrouche, E. Lespessailles, R. Jenanna, One dimensional local binary pattern for bone texture classification, Pattern Anal. Appl. 17 (1) (2012) 179–193.

[136] N. Lazic, P. Aarabi, Spoken term detection using visual spectrogram matching, in: Proceedings of IEEE International Symposium on Multimedia, 2008, pp. 637–642.

[137] Y. M. G. Costa, L. S. Oliveira, A. L. Koerich, F. Gouyon, J. G. Martins, Music genre classification using LBP textural features, Signal Proc. 92 (2012) 2723–2737.

[138] M. Esfahanian, H. Zhuang, N. Erdol, Using local binary patterns as features for classification of dolphin calls, JASA Express Lett. 134 (1) (2013) 105–111.

[139] S. Moore, R. Bowden, Local binary patterns for multi-view facial expression recognition, Comput. Vis. Image Understand. 115 (5) (2011) 541–558.

[140] Z. Lei, M. Pietikäinen, S. Z. Li, Learning discriminant face descriptor, IEEE Trans. Pattern Anal. Mach. Intell. 36 (2) (2014) 289–302.

[141] N. P. Doshi, G. Schaefer, A comprehensive benchmark of local binary pattern algorithms for texture retrieval, in: Proceedings of International Conference on Pattern Recognition, 2012, pp. 2760–2763.