

# LBP-Based Edge-Texture Features for Object Recognition

Amit Satpathy, *Member, IEEE*, Xudong Jiang, *Senior Member, IEEE*, and How-Lung Eng, *Member, IEEE*

**Abstract**—This paper proposes two sets of novel edge-texture features, Discriminative Robust Local Binary Pattern (DRLBP) and Ternary Pattern (DRLTP), for object recognition. By investigating the limitations of Local Binary Pattern (LBP), Local Ternary Pattern (LTP) and Robust LBP (RLBP), DRLBP and DRLTP are proposed as new features. They solve the problem of discrimination between a bright object against a dark background and vice-versa inherent in LBP and LTP. DRLBP also resolves the problem of RLBP whereby LBP codes and their complements in the same block are mapped to the same code. Furthermore, the proposed features retain contrast information necessary for proper representation of object contours that LBP, LTP, and RLBP discard. Our proposed features are tested on seven challenging data sets: INRIA Human, Caltech Pedestrian, UIUC Car, Caltech 101, Caltech 256, Brodatz, and KTH-TIPS2-a. Results demonstrate that the proposed features outperform the compared approaches on most data sets.

**Index Terms**—Object recognition, local binary pattern, local ternary pattern, feature extraction, texture.

## I. INTRODUCTION

CATEGORY recognition and detection are 2 parts of object recognition. The objective of category recognition is to classify an object into one of several predefined categories. The goal of detection is to distinguish objects from the background. There are various object recognition challenges. Typically, objects have to be detected against cluttered, noisy backgrounds and other objects under different illumination and contrast environments. Proper feature representation is a crucial step in an object recognition system as it improves performance by discriminating the object from the background or other objects in different lightings and scenarios. Furthermore, a good feature also simplifies the classification framework.

Object recognition features are categorized into two groups - sparse and dense representations [7]. For *sparse* feature representations, interest-point detectors are used to identify structures such as corners and blobs on the object. A feature

is created for the image patch around each point. Popular feature representations include Scale-Invariant Feature Transform (SIFT) [16], [30], Speeded Up Robust Feature [3], Local Steering Kernel [49], Principal Curvature-Based Regions [9], Region Self-Similarity features [33], [50], Sparse Color [51] and the sparse parts-based representation [1]. A comprehensive evaluation of sparse features can be found in [34] and [35].

*Dense* feature representations, which are extracted at fixed locations densely in a detection window, are gaining popularity as they describe objects richly compared to sparse feature representations. Various feature representations such as Wavelet [40], Haar-like features [55], Histogram of Oriented Gradients (HOG) [8], [56], Extended Histogram of Gradients [44]–[46], [48], Feature Context [57], Local Binary Pattern (LBP) [2], [22], [47], Local Ternary Pattern (LTP) [52], Geometric-blur [59] and Local Edge Orientation Histograms [25] have been proposed over recent years. Dense SIFT has also been proposed to alleviate the sparse representation problems [4], [24], [53].

LBP is the most popular texture classification feature [18], [20], [21], [27], [38], [41], [42], [62]. It has also shown excellent face detection performance [2], [19], [26], [52], [61]. It is robust to illumination and contrast variations as it only considers the signs of the pixel differences. Histogramming LBP codes makes the descriptor resistant to translations within the histogramming neighbourhood. However, it is sensitive to noise and small fluctuations of pixel values. To handle this, Local Ternary Pattern (LTP) has been proposed [52]. In comparison to LBP, it has 2 thresholds which creates 3 different states as compared to 2 in LBP. It is more resistant to noise and small pixel value variations compared to LBP. Like LBP, it has also been used for texture classification and face detection [13], [23], [28], [42], [52], [60].

However, for object recognition, LBP and LTP present two issues. They differentiate a bright object against a dark background and vice versa. This increases the object intra-class variations which is undesirable for most object recognitions. Nguyen *et al.* [37] propose Robust LBP (RLBP) to map a LBP code and its complement to the minimum of both to solve the problem. However, in the *same block*, RLBP also maps to the same value. For some different local structures, a similar feature is obtained. Hence, it is unable to differentiate them.

Different objects have different shapes and textures. It is therefore desirable to represent objects using both texture and edge information. However, in order to be robust to illumination and contrast variations, LBP, LTP and RLBP do not differentiate between a weak contrast local pattern and

Manuscript received March 9, 2013; revised September 23, 2013 and December 23, 2013; accepted February 20, 2014. Date of publication March 11, 2014; date of current version March 26, 2014. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Rafael Molina.

A. Satpathy is with the Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore 138632 (e-mail: satpathya@i2r.a-star.edu.sg).

X. Jiang is with the School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore 639798 (e-mail: exdjiang@ntu.edu.sg).

H.-L. Eng is with Zweec Analytics, Singapore 139950 (e-mail: howlungeng@zweec.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2014.2310123

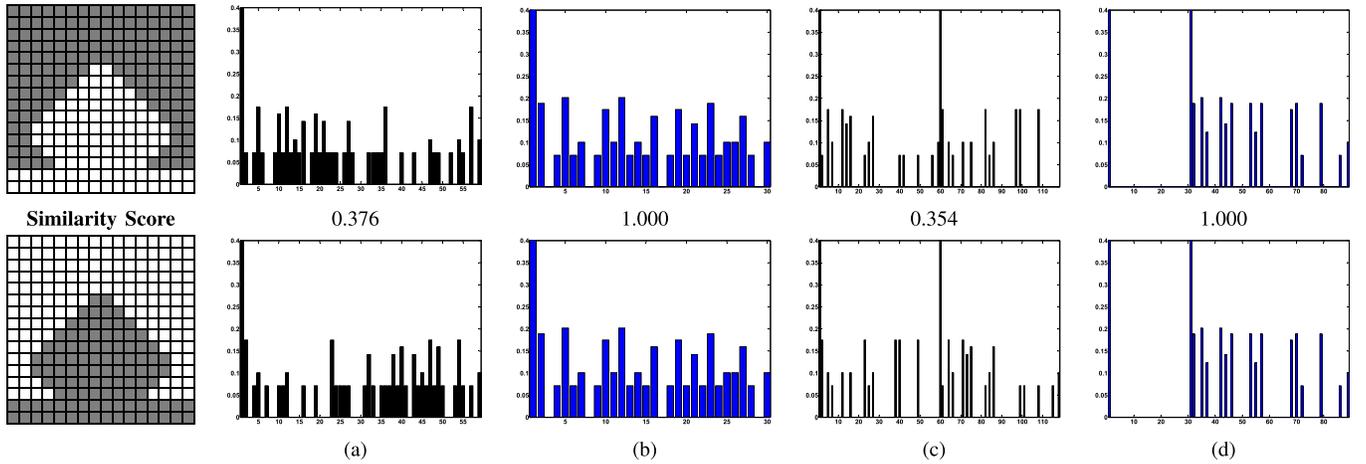


Fig. 1. Problem of Local Binary Patterns (LBP) and Local Ternary Patterns (LTP) and their solutions by Robust LBP (RLBP) and Robust LTP (RLTP). The similarity score between histograms is measured by histogram intersection. (a) LBP. (b) RLBP. (c) LTP. (d) RLTP.

a similar strong one. They only capture texture information. Object contours, which also contain discriminatory information, tend to be situated in strong contrast regions. Therefore, by discarding contrast information, contours may not be effectively represented.

In this paper, we propose two sets of novel edge-texture features, Discriminative Robust LBP (DRLBP) and LTP. The proposed features solve the issues of LBP, LTP and RLBP. They alleviate the intensity reversal problem of object and background. Furthermore, DRLBP discriminates local structures that RLBP misrepresent. In addition, the proposed features retain the contrast information of image patterns. They contain *both* edge and texture information which is desirable for object recognition. We present comprehensive experimental results on 7 data sets - INRIA Human, Caltech Pedestrian, UIUC Car, Caltech 101, Caltech 256, Brodatz and KTH-TIPS2-a. Results indicate that the proposed features outperform LBP, LTP and RLBP and perform better than other approaches in comparison on data sets.

## II. DISCRIMINATIVE ROBUST LOCAL BINARY AND TERNARY PATTERNS

### A. Limitations of LBP, LTP and Robust LBP

The LBP [38] code at location  $(x, y)$  is computed as follows:

$$LBP_{x,y} = \sum_{b=0}^{B-1} s(p_b - p_c)2^b, \quad (1)$$

$$s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

where  $p_c$  is the pixel value at  $(x, y)$ ,  $p_b$  is the pixel value estimated using bilinear interpolation from neighbouring pixels in the  $b$ -th location on the circle of radius  $R$  around  $p_c$  and  $B$  is the total number of neighbouring pixels. A  $2^B$ -bin block histogram is computed. There are some patterns that occur more frequently than others and the number of state

transitions between 0 and 1 for them are at most two [38]. Such patterns are called uniform patterns and the rest as non-uniform. By giving each uniform pattern a bin and collating all non-uniform patterns into a single bin, the bin number is reduced. For  $B = 8$ , it is reduced from 256 to 59.

In [38], another LBP variant, rotation-invariant LBP, is proposed for texture classification. However, our focus is on object recognition. Different objects exhibit different shapes that are captured by the orientation information. It is beneficial to retain these inter-class variations to model the objects. Using rotation-invariant features will significantly reduce the inter-class variations and oversimplify the object models. Therefore, in our work, rotation-invariance is not considered.

LBP is invariant to monotonic intensity changes. Hence, it is robust to illumination and contrast variations. However, it is sensitive to noise and small pixel value fluctuations. Therefore, LTP [52] has been proposed to handle this situation. The LTP code at location  $(x, y)$  is computed as follows:

$$LTP_{x,y} = \sum_{b=0}^{B-1} s'(p_b - p_c)3^b, \quad (2)$$

$$s'(z) = \begin{cases} 1, & z \geq T \\ 0, & -T < z < T \\ -1, & z \leq -T \end{cases}$$

where  $T$  is a user-defined threshold. As defined by  $s'(z)$ , LTP has 3 states while LBP has two. A  $3^B$ -bin block histogram is computed. For  $B = 8$ , the histogram has 6561 bins which is very high-dimensional. Hence, in [52], the authors propose to split the LTP code into its “upper” and “lower” LBP codes. The “upper” code,  $ULBP$ , is computed as follows:

$$ULBP = \sum_{b=0}^{B-1} f(p_b - p_c)2^b, \quad (3)$$

$$f(z) = \begin{cases} 1, & z \geq T \\ 0, & \text{otherwise} \end{cases}$$

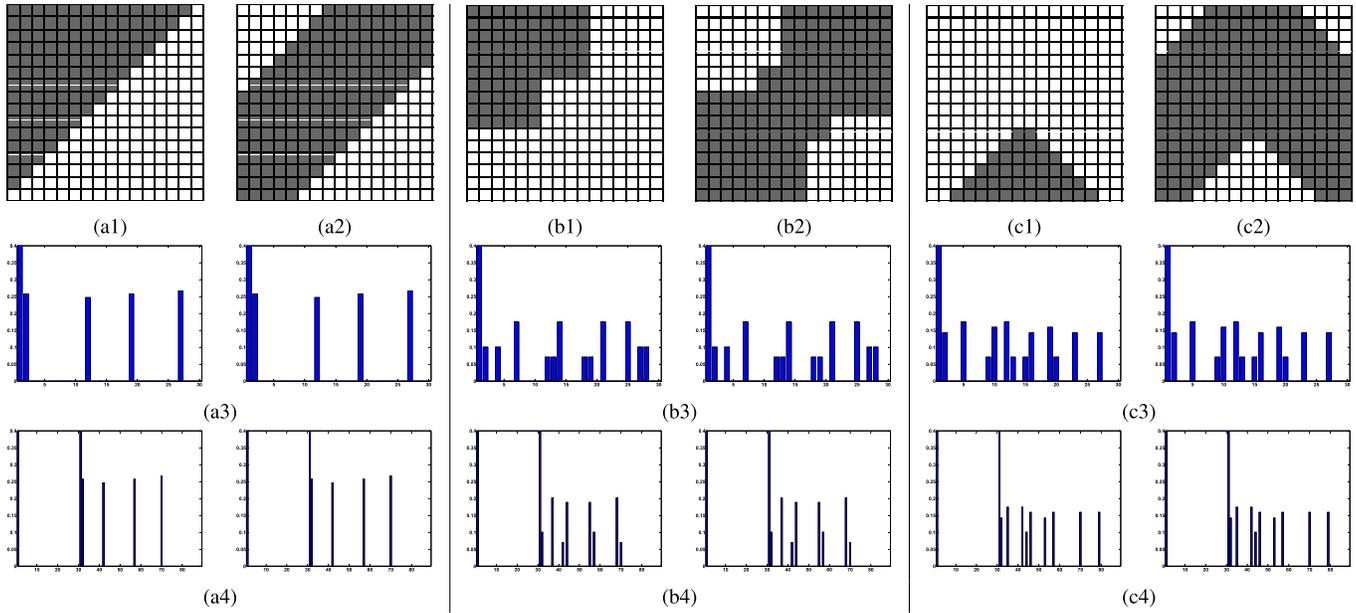


Fig. 2. Problem of Robust LBP (RLBP) and Robust LTP (RLTP). 6 local structures are shown in the first row. The second row shows the RLBP features for each structure. The third row shows the RLTP features for each structure. RLBP and RLTP make the different structures in (a1) and (a2) similar as shown in (a3) and in (a4) and different structures in (b1) and (b2) similar as shown in (b3) and (b4). A similar situation can be observed in (c).

The “lower” code, *LLBP*, is computed as follows:

$$LLBP = \sum_{b=0}^{B-1} f'(p_b - p_c)2^b, \quad (4)$$

$$f'(z) = \begin{cases} 1, & z \leq -T \\ 0, & \text{otherwise} \end{cases}$$

By doing so, the bin number is reduced from 6561 to 512. Using uniform LBP codes, it is further reduced to 118.

An issue with LBP and LTP is that they differentiate a bright object against a dark background and vice-versa. This differentiation makes the object intra-class variation larger. In Fig. 1(a) and (c), the 2 situations in a block are represented by LBP and LTP. As seen, the LBP and LTP features for the 2 situations are different.

To solve the above mentioned problem of LBP, the authors in [37] propose mapping a LBP code and its complement to the minimum of the two. For instance, “1101 0101” and its complement, “0010 1010”, become the same code “0010 1010” in the mapping. The states are changed from 0 to 1 or 1 to 0 during this mapping. By doing so, the code is *robust* to the reversal in intensity between the background and the objects. Based on this, we name this code as *Robust LBP* (RLBP). RLBP is computed as follows:

$$RLBP_{x,y} = \min \left\{ LBP_{x,y}, 2^B - 1 - LBP_{x,y} \right\}, \quad (5)$$

where  $LBP_{x,y}$  is as defined in Eq (1) and  $2^B - 1 - LBP_{x,y}$  is the complement code. Since the mapping halves the number of codes, the RLBP bin number is 128 for  $B = 8$ . Using uniform codes, it is reduced to 30. Fig. 1(b) illustrates how RLBP solves the issue of intensity reversal of object and background. It is seen that for both situations, the RLBP feature is the same.

To solve the problem of brightness reversal of object and background, RLBP maps all LBP codes to the minimum of the code and its complement. However, this mapping makes it difficult to differentiate some dissimilar local structures. It is possible for 2 different structures to have similar or even same features. This is illustrated in row 2 of Fig. 2. This problem is caused by merging the complement codes *in the same block*.

### B. The Proposed Discriminative Robust Local Binary Pattern

An object has 2 distinct cues for differentiation from other objects - the object surface texture and the object shape formed by its boundary. The boundary often shows much higher contrast between the object and the background than the surface texture. Differentiating the boundary from the surface texture brings additional discriminatory information because the boundary contains the shape information. However, in order to be robust to illumination and contrast variations, LBP does not differentiate between a weak contrast local pattern and a strong contrast one. It mainly captures the object texture information. The histogramming of LBP codes only considers the frequencies of the codes i.e. the weight for each code is the same. This makes it difficult to differentiate a weak contrast local pattern and a strong contrast one.

To mitigate this, we propose to fuse edge and texture information in a single representation by modifying the way the codes are histogrammed. Instead of considering the code frequencies, we assign a weight,  $\omega_{x,y}$ , to each code which is then voted into the bin that represents the code. The weight we choose is the pixel gradient magnitude which is computed as follows. Following [8], the square root of the pixels is taken. Then, the first order gradients are computed. The gradient magnitude at each pixel is then computed as  $\omega_{x,y} = \sqrt{I_x^2 + I_y^2}$  where  $I_x$  and  $I_y$  are the first-order derivatives in the  $x$  and

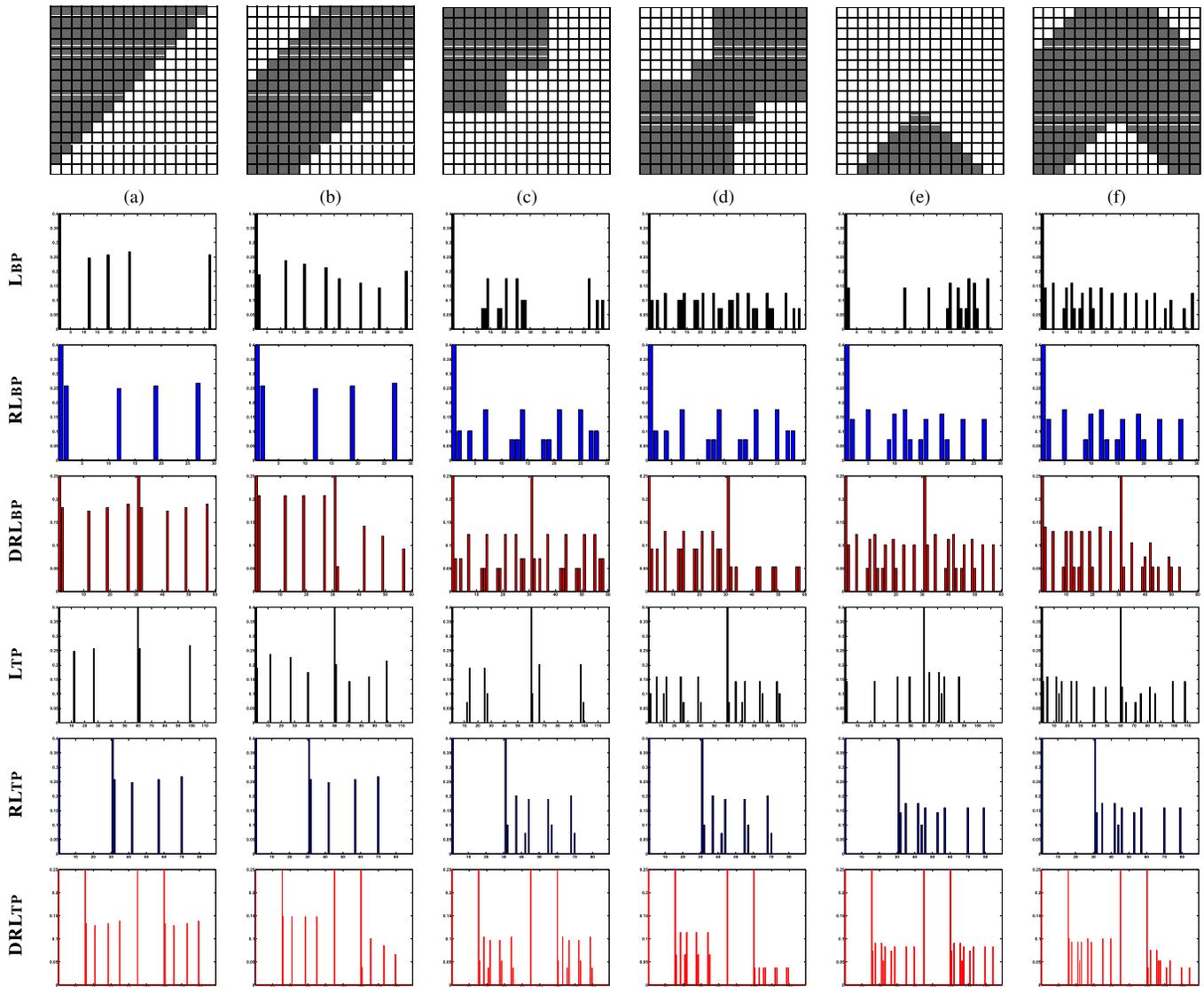


Fig. 3. DRLBP and DRLTP representations of local structures in Fig. 2. DRLBP and DRLTP differentiate the local structure pairs, {(a),(b)}, {(c),(d)} and {(e),(f)} misrepresented by RLBP and RLTP.

$y$  directions.  $\omega_{x,y}$  is then used to weigh the LBP code. The stronger the pixel contrast, the larger the weight assigned to the pixel LBP code. In this way, if a LBP code covers both sides of a strong edge, its gradient magnitude will be much larger and by voting this into the bin of the LBP code, we take into account if the pattern in the local area is of a strong contrast. Thus, the resulting feature will contain both edge and texture information in a single representation. The value of the  $i^{\text{th}}$  weighted LBP bin of a  $M \times N$  block is as follows:

$$h_{lbp}(i) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \omega_{x,y} \delta(LBP_{x,y}, i), \quad (6)$$

$$\delta(m, n) = \begin{cases} 1, & m = n \\ 0, & \text{otherwise} \end{cases}$$

The RLBP histogram is created from (6) as follows:

$$h_{rlbp}(i) = h_{lbp}(i) + h_{lbp}(2^B - 1 - i), \quad 0 \leq i < 2^{B-1} \quad (7)$$

where  $h_{rlbp}(i)$  is the  $i^{\text{th}}$  RLBP bin value. To mitigate the RLBP issue in Fig. 2, consider the absolute difference between the bins of a LBP code and its complement to form Difference of LBP (DLBP) histogram as follows:

$$h_{dlbp}(i) = |h_{lbp}(i) - h_{lbp}(2^B - 1 - i)|, \quad 0 \leq i < 2^{B-1} \quad (8)$$

where  $h_{dlbp}(i)$  is the  $i^{\text{th}}$  DLBP bin value. The number of DLBP bins is 128 for  $B = 8$ . Using uniform codes, it is reduced to 30. For blocks that contain structures with both LBP codes and their complements, DLBP assigns small values to the mapped bins. It differentiates these structures from those having no complement codes within the block.

The 2 histogram features, RLBP and DLBP, are concatenated to form *Discriminative Robust LBP* (DRLBP) as follows:

$$h_{drlbp}(j) = \begin{cases} h_{rlbp}(j), & 0 \leq j < 2^{B-1} \\ h_{dlbp}(j - 2^{B-1}), & 2^{B-1} \leq j < 2^B \end{cases} \quad (9)$$

For  $B = 8$ , the number of bins is 256 (128 + 128). Using uniform codes, it is reduced to 60 (30 + 30).

TABLE I  
SIMILARITY SCORES FOR LOCAL STRUCTURE PAIRS IN FIG. 3

Approach	Similarity Score		
	{(a),(b)}	{(c),(d)}	{(e),(f)}
LBP	0.280	0.366	0.445
RLBP	1.000	1.000	1.000
DRLBP	0.109	0.169	0.111
LTP	0.280	0.354	0.354
RLTP	1.000	1.000	1.000
DRLTP	0.109	0.148	0.112

DRLBP contains *both* edge and texture information. Fig. 3 illustrates how DRLBP produces different features for the structures shown in Fig. 2. Table I shows the similarity scores between the features using histogram intersection. Hence, DRLBP represents objects more discriminatively than RLBP. It also resolves the issue of intensity reversal of object and background as shown in Fig. 5.

### C. The Proposed Discriminative Robust Local Ternary Pattern

1) *Robust Local Ternary Pattern*: LBP is sensitive to noise and small pixel value fluctuations [52]. LTP solves this using 2 thresholds to generate codes. It is more resistant to small pixel value variations and noise compared to LBP. However, it also has the same problem as LBP whereby it differentiates a bright object against a dark background and vice-versa. RLBP [37] solves this problem for LBP by mapping a LBP code and its complement to the minimum of the two.

However, RLBP cannot be applied to *ULBP* and *LLBP* of LTP. For a pair of object/background intensity inverted patterns, their *ULBP* codes are not complements. Similarly, their *LLBP* codes are also not complements. This is illustrated in Fig. 4 where 2 different cases of object/background inverted intensity patterns are shown. In Fig. 4(a1) and (a2), a case illustrating a neighbourhood, where all 3 LTP states occur, is shown. From the two LTP codes, it is observed that the 2 patterns are simply intensity inverted. However, their corresponding *ULBP* codes are *not* complements. Similarly, their corresponding *LLBP* codes are also not complements. A similar situation is observed in (b1) and (b2) where only 2 LTP states are present. The *ULBP* and *LLBP* codes are not complements. Hence, RLBP cannot be applied to *ULBP* and *LLBP* to obtain a feature that is robust to the reversal in intensity between the objects and background.

In order to alleviate this problem of LTP, we need to analyze the 3-state LTP definition in (2): 1, 0 and  $-1$ . The state of 0 represent regions of small variations, noise and uniform regions. It will not change when there is an inversion of brightness between the background and objects as the variations remain the same. Therefore, for a pair of brightness inverted object/background patterns, only the state of  $-1$  is inverted to 1 and vice-versa. Hence, for every LTP code, we can find its corresponding inverted code. For instance, “ $-1-100\ 1100$ ” has an inverted code “ $1100\ -1-100$ ”. If both codes are mapped to a same bin, a feature that is robust to the reversal in intensity between the objects and background can be obtained.

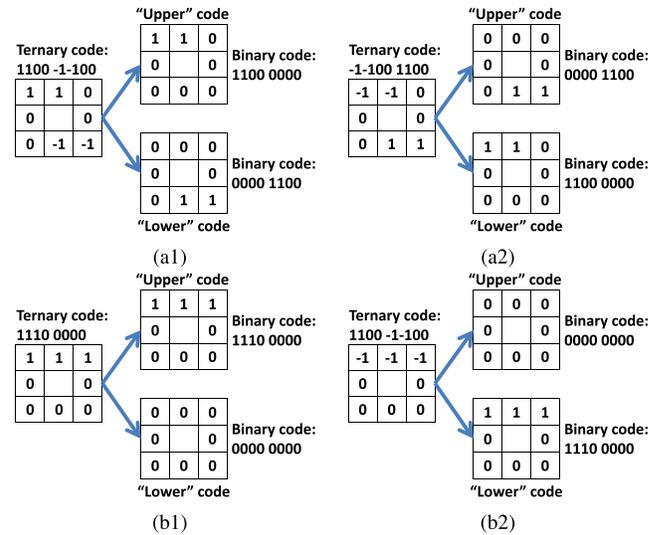


Fig. 4. Illustration of *ULBP* and *LLBP* codes of LTP for 2 situations where the intensities are reversed. It can be seen that the *ULBP* and *LLBP* codes are reversed for the 2 situations.

In this paper, the maximum of a LTP code and its inverted representation is chosen. We name it as *Robust LTP* (RLTP). Mathematically, RLTP is formulated as follows:

$$RLTP_{x,y} = \max \{ LTP_{x,y}, -LTP_{x,y} \} \quad (10)$$

The RLTP code can then be split into “upper” and “lower” LBP codes. The “upper” code, *URLBP*, is expressed as follows:

$$URLBP = \sum_{b=0}^{B-1} h(RLTP_{x,y,b}) 2^b, \quad (11)$$

$$h(z) = \begin{cases} 1, & z = 1 \\ 0, & \text{otherwise} \end{cases}$$

where  $RLTP_{x,y,b}$  represents the RLTP state value at the  $b$ -th location. The “lower” code, *LRLBP*, is computed as follows:

$$LRLBP = \sum_{b=0}^{B-1} h'(RLTP_{x,y,b}) 2^b, \quad (12)$$

$$h'(z) = \begin{cases} 1, & z = -1 \\ 0, & \text{otherwise} \end{cases}$$

The most significant bit of *LRLBP* is 0 as the state at  $(B-1)$ -th location of RLTP is either 0 or 1. Fig. 1(d) illustrates how RLTP alleviates the brightness reversal problem of object and background. It is observed that for the two situations, the RLTP features are the same.

However, similar to RLBP in Section II-A, RLTP also maps a LTP code and its inverted representation in the same block to the same value. This is illustrated in Fig. 2 in the last row.

2) *Discriminative Robust Local Ternary Patterns*: LTP and RLTP are also robust to illumination and contrast variations and only capture texture information. Hence, the weighting scheme in Section II-B is also used. The  $k^{th}$  weighted LTP bin value of a  $M \times N$  image block is as follows:

$$h_{ltp}(k) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \omega_{x,y} \delta(LTP_{x,y}, k), \quad (13)$$

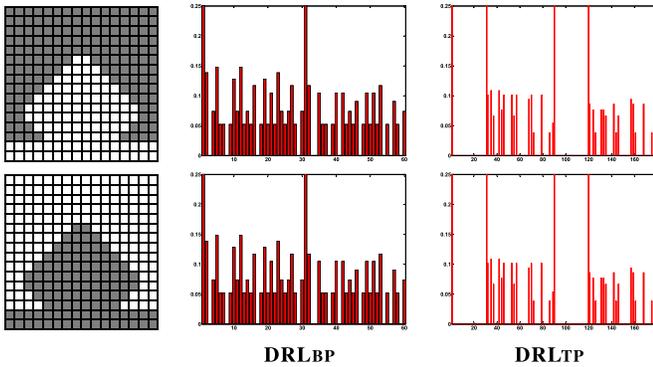


Fig. 5. Same DRLBPs and DRLTPs are produced for the two intensity reversed patterns in Fig. 1. The similarity values using histogram intersection is 1 for both features.

The RLTP histogram is created from (13) as follows:

$$h_{rltp}(k) = \begin{cases} h_{ltp}(k), & k = 0 \\ h_{ltp}(k) + h_{ltp}(-k), & 0 < k < \frac{3^B+1}{2} \end{cases} \quad (14)$$

where  $h_{rltp}(k)$  is the  $k^{th}$  RLTP bin value.

The absolute difference between the bins of a LTP code and its inverted representation is taken to form Difference of LTP (DLTP) histogram as follows:

$$h_{dltp}(k) = |h_{ltp}(k) - h_{ltp}(-k)|, \quad 0 < k < \frac{3^B+1}{2} \quad (15)$$

where  $h_{dltp}(k)$  is the  $k^{th}$  DLTP bin value. RLTP and DLTP are concatenated to form *Discriminative Robust LTP* (DRLTP) as follows:

$$h_{drltp}(l) = \begin{cases} h_{rltp}(l), & 0 \leq l < \frac{3^B+1}{2} \\ h_{dltp}(l - \frac{3^B-1}{2}), & \frac{3^B+1}{2} \leq l < 3^B \end{cases} \quad (16)$$

Using (11) and (12), the “upper” and “lower” LBP histograms of DRLTP are computed. Similar to DRLBP, DRLTP contains *both* edge and texture information. Fig. 3 illustrates how DRLTP produces different features for the structures in Fig. 2. Table I shows the similarity scores between the features using histogram intersection. It also resolves the issue of brightness reversal of object and background as shown in Fig. 5.

#### D. Efficient Computation of DRLTP Using $ULBP$ and $LLBP$

Using LTP to find RLTP, DLTP and DRLTP is computationally intensive and requires a large storage requirement. For  $B = 8$ , the number of LTP codes is 6561. In order to generate the RLTP and DLTP histograms from the LTP histogram, there are 3280 addition and subtraction operations respectively. This is followed by 8 addition operations for each RLTP and DLTP code to find the “upper” LBP code and 8 addition operations to find the “lower” LBP code. If the “upper” and “lower” LBP codes of RLTP and DLTP can be produced directly from the split LBP codes of LTP, the computational complexity and storage requirements will be greatly reduced.

The behaviours of  $ULBP$  (3) and  $LLBP$  (4) for object/background intensity inverted situations are analyzed as follows. Suppose there is a bright object against a dark

background. Consider a neighbourhood with an object boundary. Assume that the centre pixel resides in the background. The differences between the object pixel values and the centre pixel value are larger than the threshold,  $T$ . The differences between the background pixel values and the centre pixel value are in between  $T$  and  $-T$ . The  $ULBP$  bits corresponding to the object are 1 while the rest are 0. The  $LLBP$  bits are all 0. If the brightness is now inverted for the situation, all  $ULBP$  bits are 0 and the  $LLBP$  bits corresponding to the object are 1 while the rest are 0. The brightness inversion turns  $LLBP$  into  $ULBP$  and  $ULBP$  into  $LLBP$ .

Now, assume that the centre pixel does not belong to the background or object. Instead, it has a value between the bright object and dark background pixel values. The absolute differences of the object and the centre pixel and the background and the centre pixel are larger than  $T$ . The  $ULBP$  bits corresponding to the object are 1 while the rest are 0. The  $LLBP$  bits corresponding to the background are 1 while the rest are 0. If the intensity is now inverted for the situation, the  $ULBP$  bits corresponding to the background are all 1 while the rest are 0. Similarly, the  $LLBP$  bits corresponding to the object are 1 while the rest are 0. Again, the intensity inversion turns  $LLBP$  into  $ULBP$  and  $ULBP$  into  $LLBP$ .

From the above analysis, we find that the  $ULBP$  and  $LLBP$  codes for object/background intensity inverted situations are exchanged. If they are rearranged such that the “upper” and “lower” codes for both situations are the same, RLTP is achieved. This can be done as follows. For any LTP code, the  $URLBP$  code is defined as follows:

$$URLBP = \max\{ULBP, LLBP\}, \quad (17)$$

The  $LRBP$  code is defined as follows:

$$LRLBP = \min\{ULBP, LLBP\}, \quad (18)$$

By producing  $URLBP$  and  $LRLBP$  codes for any LTP code, RLTP is obtained in the split LBP code representation. For the situation where  $ULBP = 0$  and  $LLBP = 0$ , only 1 LBP result is considered and assigned to  $LRLBP$ . In Fig. 4(a) and (b), for each case, the LBP codes of the 2 intensity inverted LTP codes are reversed. For instance, in Fig. 4(a1), the  $ULBP$  code is the  $LLBP$  code in (a2). Similarly, the  $LLBP$  code is the  $ULBP$  code in (a2). By following (17) and (18), we can obtain the  $URLBP$  and  $LRLBP$  easily from  $ULBP$  and  $LLBP$  for both cases.

The  $s^{th}$   $URLBP$  bin value,  $0 < s < 2^B$ , is generated from  $ULBP$  and  $LLBP$  codes as follows:

$$h_{urlbp}(s) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \omega_{x,y} \delta(\max(ULBP, LLBP), s), \quad (19)$$

The  $t^{th}$   $LRLBP$  bin value,  $0 \leq t < 2^{B-1}$ , is as follows:

$$h_{lrlbp}(t) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \omega_{x,y} \delta(\min(ULBP, LLBP), t), \quad (20)$$

The split LBP histograms,  $UDLBP$  and  $LDLBP$ , of DLTP are also generated from the  $ULBP$  and  $LLBP$  codes. For every LTP code whose  $ULBP$  and  $LLBP$  are swapped,

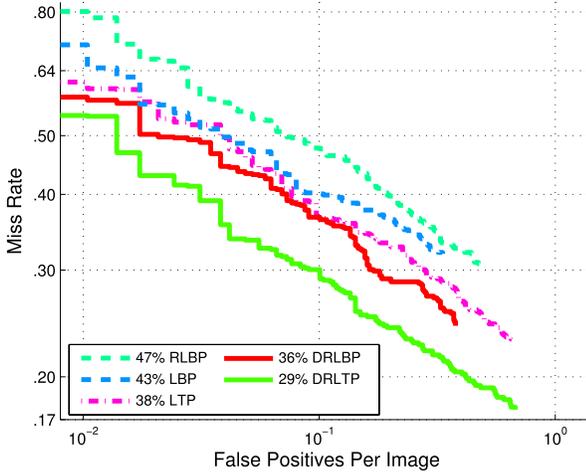


Fig. 6. Performance of DRLBP and DRLTP against LBP, LTP and RLBP [Best viewed in colour]. DRLTP outperforms all other methods.

TABLE II

RECOGNITION ACCURACY OF DRLBP AND DRLTP AGAINST LTP, LBP AND RLBP ON CALTECH 101 DATA SET

Method	Performance(%) (15 training and test)	Performance(%) (30 training and test)	Performance(%) (90% training and 10% test)
DRLTP	72.59 ± 1.58	80.41 ± 0.98	93.10 ± 0.81
DRLBP	57.64 ± 1.94	66.95 ± 0.62	89.13 ± 1.30
LTP	55.71 ± 1.03	66.00 ± 1.01	74.65 ± 1.59
LBP	52.63 ± 2.23	63.23 ± 1.22	79.04 ± 1.30
RLBP	56.39 ± 2.12	65.86 ± 0.51	83.84 ± 1.83

the corresponding  $UDLBP$  and  $LDLBP$  bin values are decremented by  $\omega_{x,y}$  accordingly. Otherwise, the bins are incremented by  $\omega_{x,y}$ .

The  $s^{th}$   $UDLBP$  bin value is as follows:

$$h_{udlbp}(s) = \left| \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \omega_{x,y} \delta'(\lambda(ULBP, LLBP), s) \right|, \quad (21)$$

$$\lambda(p, q) = \begin{cases} p, & p > q \\ -q, & p < q \end{cases}$$

$$\delta'(m, n) = \begin{cases} 1, & m = n, m > 0 \\ -1, & |m| = n, m < 0 \\ 0, & \text{otherwise} \end{cases}$$

$\lambda(\bullet)$  determines whether the  $ULBP$  and  $LLBP$  codes are being swapped. If a swap occurs, the negative maximum code is assigned to the result.  $\delta'(\bullet)$  checks the value output from  $\lambda$  with  $s$ . If the value is positive and matches  $s$ , the  $s^{th}$  bin value is incremented. Otherwise, it is decremented. The  $t^{th}$   $LDLBP$  bin value is as follows:

$$h_{ldlbp}(t) = \left| \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \omega_{x,y} \delta''(\lambda'(ULBP, LLBP), t) \right|, \quad (22)$$

$$\lambda'(p, q) = \begin{cases} q, & p \geq q \\ -p, & p < q \end{cases}$$

$$\delta''(m, n) = \begin{cases} 1, & m = n, m \geq 0 \\ -1, & |m| = n, m < 0 \\ 0, & \text{otherwise} \end{cases}$$

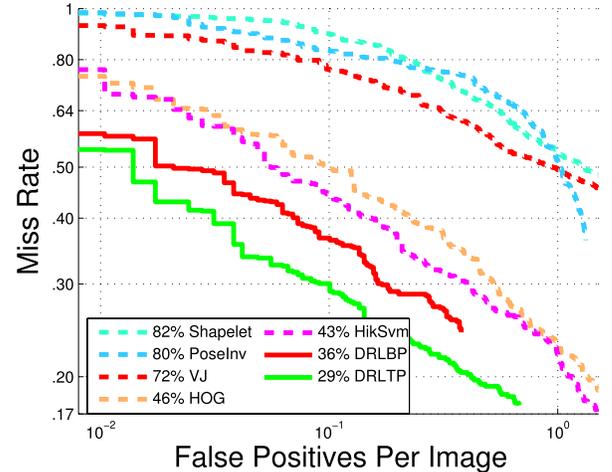


Fig. 7. Performance of DRLBP and DRLTP against existing state-of-the-art methods of dense representation [Best viewed in colour]. DRLTP outperforms all other methods.

$\lambda'(\bullet)$  determines whether the  $ULBP$  and  $LLBP$  codes are being swapped. If a swap occurs, the negative minimum code is assigned to the result.  $\delta''(\bullet)$  checks the value output from  $\lambda'$  with  $t$ . If the value is zero or positive and matches  $t$ , the  $t^{th}$  bin value is incremented. Otherwise, it is decremented. The  $URLBP$ ,  $LRLBP$ ,  $UDLBP$  and  $LDLBP$  histograms are then concatenated to form DRLTP.

A property of  $ULBP$  and  $LLBP$  is that any  $ULBP$  bit that is 1 must be 0 in  $LLBP$ . Similarly, any  $LLBP$  bit that is 1 must be 0 in  $ULBP$ . This property of  $ULBP$  and  $LLBP$  also applies to  $URLBP$  and  $LRLBP$ . Since RLTP is obtained from a max function, the most significant bit i.e. left-most bit of  $LRLBP$  must be 0. Thus, the number of possible codes of  $LRLBP$  is almost halved compared to  $URLBP$ . For  $B = 8$ , the  $URLBP$  have values in the range from 1 to 255.  $LRLBP$  only have values in the range from 0 to 127. Hence, the uniform  $URLBP$  bin number is 58 and that of uniform  $LRLBP$  is 30 (the bin representing 0 is in  $LRLBP$ ). Hence, for a block, the total RLTP bin number using  $URLBP$  and  $LRLBP$  is 88. Similarly, for DLTP using  $UDLBP$  and  $LDLBP$ , the total bin number is also 88. Overall, the DRLTP bin number is 176.

### III. EXPERIMENTAL EVALUATION

7 challenging data sets are used for evaluation. The detection performance is evaluated on 3 data sets - INRIA [8], Caltech Pedestrian Benchmark Data Set [10] and UIUC Car [1]. Results are reported for the INRIA and Caltech data sets using the per-image methodology [10]. Results are reported for UIUC Car data set using performance at Equal Error Rate (EER). The classification performance is evaluated on 2 data sets - Caltech 101 [11], [12] and Caltech 256 [17]. All the classification experiments for each data set are repeated 10 times with randomly selected training and test images (15 and 30). The average of per-class recognition rates is recorded for each run. The mean and standard deviation of the results from individual runs is reported as the final results. The framework of [24] is used in the classification experiments

where the SIFT features are replaced with the proposed features. The texture classification performance is evaluated on 2 data sets - Brodatz [5], [39] and KTH-TIPS2-a [6].

A  $8 \times 8$  pixels block size is used for UIUC Car data set. For INRIA, Caltech Pedestrian, Caltech 101 and Caltech 256, a  $16 \times 16$  pixels block size is used. A 50% overlap of blocks is considered. The histograms are normalized using L1-norm. The square root of the bins are then taken. Linear SVM classifier is used for these data sets. For the classification experiments, the linear SVM classifiers are trained using a one-versus-all rule i.e. a classifier is trained to separate each class from the rest and a test image is assigned the label of the classifier with the highest response.

For the texture classification experiments, we follow the procedures in [7] for training and testing on Brodatz and KTH-TIPS2-a. Global features are used for Brodatz and KTH-TIPS2-a i.e. the entire image is represented by a single histogram. The DRLBP and DRLTP histograms are normalized first using L2-norm followed by L1-norm. Similar to [7], we use a 3-nearest neighbor classifier with normalized histogram intersection as the distance measure between features.

For all data sets, a circular neighbourhood of radius 1 ( $R$ ) and 8 ( $B$ ) pixels is considered. The uniform pattern representation is used. For LTP and DRLTP in our experiments, the threshold,  $T$ , is 3 for INRIA and Caltech Pedestrian, 9 for UIUC Car, Caltech 101 and Caltech 256, 15 for Brodatz and 5 for KTH-TIPS2-a.

#### A. Performance Comparison of DRLBP and DRLTP Against LBP, LTP and RLBP

We compare the performance of DRLBP and DRLTP against LBP, LTP and RLBP on INRIA for detection and on Caltech 101 for classification. The INRIA training set contains 2416 cropped positive images and 1218 uncropped negative images. The sliding image window size is  $128 \times 64$  pixels. We randomly take 10 samples from each negative image to obtain a total of 12180 negative samples for training the linear SVM classifier. Bootstrapping is then performed across multiple scales at a scale step of 1.05 to obtain hard negatives which are added to the original training set for retraining. This training procedure is exactly the same as described in [8] and [10].

The INRIA test set consist of 288 images. The images are scanned over multiple scales at a scale step of 1.05. The window stride is 8 pixels in the  $x$  and  $y$  directions. These parameters are the same as those in [10]. The miss rate (MR) against false positives per image (FPPI) (using log-log plots) is plotted to compare between different detectors. The *log-average miss rate* (LAMR) [10] is used to summarize the detector performance which is computed by averaging the miss rates at nine evenly spaced FPPI rates in the range  $10^{-2}$  to  $10^0$ . If any of the curves end before reaching  $10^0$ , the minimum miss rate achieved is used [10].

From Fig. 6, it is seen that our proposed features outperform its predecessors. RLBP underperforms LBP as there is a loss of information due to the mapping of LBP codes and their complements to the same code. DRLBP outperforms RLBP

TABLE III  
RECOGNITION ACCURACY AT EQUAL ERROR RATE ON UIUC CARS DATA SET. THE RESULTS OF HIKSVM FOR TEST SET II IS NOT PROVIDED IN [31]

Method	Performance(%)	
	Test Set I	Test Set II
<b>DRLTP</b>	<b>99.5</b>	<b>96.4</b>
<b>DRLBP</b>	<b>98.3</b>	<b>88.5</b>
HIKSVm [31]	98.5	-
Agarwal <i>et al.</i> [1]	79.0	45.0
Mutch <i>et al.</i> [36]	99.9	90.6
Hae <i>et al.</i> [49]	88.1	77.7
Gall <i>et al.</i> [14]	98.5	98.6

and LBP. LTP outperforms LBP thanks to its robustness to noise and small pixel value fluctuations. Similarly, DRLTP outperforms LTP. Overall, DRLTP performs the best at 29%.

The two columns of Table II show the classification performances of DRLBP and DRLTP against LBP, LTP and RLBP on Caltech 101 using respective 15 and 30 training samples per class where the number of test samples are up to the number of training samples for each class. Again, it is seen that our proposed features outperform its predecessors. DRLTP has a recognition rate of 72.59% while LTP has a recognition rate of 55.71% for the 15 training and test images case. This shows a significant gain of 17%. Furthermore, for 30 training and test images case, the gain is 14%. Similarly, DRLBP has a gain of 1% and approximately 3% in comparison to RLBP and LBP for both cases. Furthermore, we also perform another experiment using 90% of the samples per class as training data with the remaining 10% as test data. The third column of the table shows the results. There is a significant improvement in performance for all features. This is expected as there are more samples available for training which improves classification performance. DRLTP still gives the best performance at 93.1%.

In the subsequent sections, only DRLBP and DRLTP will be compared against some other state-of-the-art approaches.

#### B. Comparisons With Other Approaches on INRIA Data Set

We compare the DRLBP and DRLTP performance with VJ [55], SHAPELET [43], POSEINV [29], HIKSVM [31] and HOG [8]. The results of all compared detectors are given in [10]. These detectors are optimized, trained and tested in [10] by their respective authors. From Fig. 7, DRLTP achieves a LAMR of 29% which is significantly lower than all compared state-of-the-art methods. DRLBP has a LAMR of 36%.

#### C. Comparisons With Other Approaches on Caltech Data Set

The Caltech Pedestrian data set [10] contains color video sequences and pedestrians with a wide range of scales and scene variations. It has been created from a recorded video on a car moving through some densely populated human areas. It contains motion, blur and noise artifacts, and has various stages of occlusion. The data set is divided into 11 sessions. The first 6 sessions are designated as the training set with the remaining 5 as the test set.

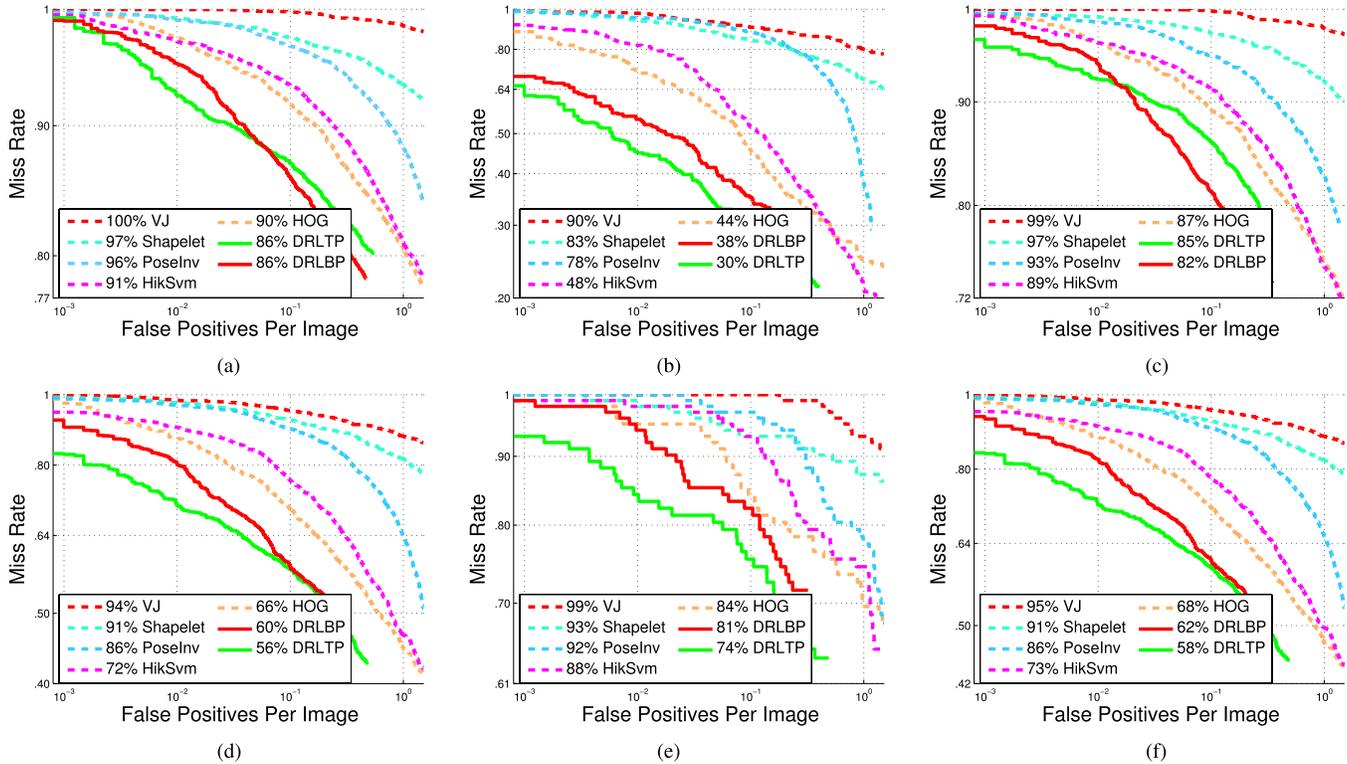


Fig. 8. Evaluation results under six different conditions on the test set of the Caltech Pedestrian Data Set [Best viewed in colour]. (a) DRLBP and DRLTP rank first in overall performance of all detectors on all annotated pedestrians. (b) DRLTP ranks first in performance on unoccluded pedestrians over 80 pixels (near scale). (c) DRLBP ranks first in performance on unoccluded pedestrians between 30-80 pixels. (d) DRLTP ranks first in performance on unoccluded pedestrians over 50 pixels tall. (e) Even under partial occlusion, DRLTP performs the best among all other methods. DRLBP ranks second. (f) DRLTP ranks first in performance on 50-pixel or taller, unoccluded or partially occluded pedestrians (reasonable).

Results were reported in [10] using detectors trained on other data sets like INRIA for detection on their test set. Our results are also presented in a similar manner where detectors are trained using the INRIA data set and tested on the test sessions. The scale step is 1.05. The window stride is 8 pixels in the  $x$  and  $y$  directions. To detect humans at smaller scales, the original images are upsampled. Only every 30<sup>th</sup> frame is evaluated. The settings used are the same as [10].

The detectors compared with our implemented detectors are the same as those in Section III-B. The results of the compared detectors are given in [10]. These detectors are optimized, trained and tested in [10] by their respective authors. The performance is analyzed under six conditions [10] as shown in Fig. 8. The results are discussed under each condition as follows.

**Overall:** Fig. 8(a) plots the performance on all test sessions for every annotated pedestrian. DRLBP and DRLTP rank first at 86%. HOG is second at 90%. At lower FPPIs, DRLTP performs better than DRLBP. For detection systems that require low false positives with low MRs, DRLTP is preferred.

**Scale:** Fig. 8(b) plots the performance on unoccluded pedestrians with heights over 80 pixels. DRLTP performs the best at 30% with DRLBP second at 38%. Fig. 8(c) plots the performance on unoccluded pedestrians with heights between 30-80 pixels. DRLBP ranks first at 82% followed by DRLTP at 85%. At lower FPPIs, DRLTP performs better than DRLBP.

**Occlusion:** Fig. 8(d) plots the performance on unoccluded pedestrians with heights over 50 pixels. DRLTP ranks first at

TABLE IV  
RECOGNITION ACCURACY ON CALTECH 101 DATA SET

Approach	Performance(%) (15 test)	Performance(%) (30 test)
<b>DRLTP</b>	<b>72.59 ± 1.58</b>	<b>80.41 ± 0.98</b>
<b>DRLBP</b>	<b>57.64 ± 1.94</b>	<b>66.95 ± 0.62</b>
Lazebnik <i>et al.</i> [24]	56.40	64.40 ± 0.80
HikSVM [31]	50.15 ± 0.61	56.49 ± 0.78
Feature Context [57]	69.63 ± 0.84	77.09 ± 0.74
Boiman <i>et al.</i> [4]	65.00 ± 1.14	70.04
Gehler <i>et al.</i> [15]	70.00	77.70 ± 0.30
Yang <i>et al.</i> [58]	67.00 ± 0.45	73.20 ± 0.54

56% and DRLBP ranks second at 60%. Fig. 8(e) plots the performance on partially occluded (1-35% occluded) pedestrians with heights over 50 pixels. DRLTP ranks first at 74% and DRLBP ranks second at 81%.

**Reasonable:** Fig. 8(f) shows evaluation of performance on pedestrians that are over 50 pixels tall under no or partial occlusion (reasonable condition). DRLTP ranks first at 58% and DRLBP ranks second at 62%.

#### D. Comparisons With Other Approaches on UIUC Car

The data set contains side views of cars taken with cameras. The cars are of different resolutions and contain instances of partial occlusion, low contrast and highly noisy and textured background. The training set contains 550 cars and 500 non-car images of 40 × 100 pixels.

TABLE V  
RECOGNITION ACCURACY ON CALTECH 256 DATA SET

Approach	Performance(%) (15 test)	Performance(%) (30 test)
<b>DRLTP</b>	<b>72.34 ± 0.47</b>	<b>81.89 ± 0.53</b>
<b>DRLBP</b>	<b>51.14 ± 1.37</b>	<b>60.80 ± 0.51</b>
Boiman <i>et al.</i> [4]	35.00	42.00
Gehler <i>et al.</i> [15]	34.00	45.80
Yang <i>et al.</i> [58]	27.73 ± 0.51	34.02 ± 0.35

There are 2 test sets. The first set contains 170 images with 200 cars of the same size as those in the training set i.e. single scale. The second set contains 108 images with 139 cars of different sizes ranging from 0.8 to 2 times the size of the cars in the training set i.e. multi scale. During testing, the window stride is 5 pixels horizontally and 2 pixels vertically.

Table. III presents the DRLBP and DRLTP recognition accuracy at Equal Error Rate (EER) against some other methods for both test sets. In comparison with the other state-of-the-art methods, DRLTP is second to [14] for multi scale test set and to [36] for the single scale test set. In [14], hybrid features with complex classification architecture is used. In [36], the features are created from multiple convolutions with 8 filters (4 Gabor and 4 3-D Max filters) over 10 scales. These are much more computationally intensive compared to our method and yet, we achieve a comparable performance.

#### E. Comparisons With Other Approaches on Caltech 101

Caltech 101 contains 101 different object classes like animals, vehicles, etc. with significant variance in shape and an additional “background” class. The number of images per class vary from 31 to 800. The classifiers are tested on 15 and 30 random test samples per class. Table IV shows the performance of the proposed features against some other approaches. DRLTP outperforms all approaches for both 15 and 30 test samples.

#### F. Comparisons With Other Approaches on Caltech 256

Caltech 256 contains 256 different object classes with higher intra-class variability and object location variability than Caltech 101 and an additional “background” class. The number of images per class vary from 80 to 827. The classifiers are tested on 15 and 30 random test samples per class. Table V shows the performance of the proposed features against some other approaches. DRLBP and DRLTP outperform all approaches for both 15 and 30 test samples.

#### G. Comparisons With Other Approaches on Brodatz

The Brodatz data set consist of 32 texture categories. Each texture image is grayscale and  $256 \times 256$  pixels in size. It is divided into 16 disjoint samples of  $64 \times 64$  pixels. Three additional samples are generated from each sample as follows: 1) a sample rotated by 90 degrees, 2) a  $64 \times 64$  scaled sample obtained from the  $45 \times 45$  pixels window centered in the original sample, and 3) a sample that is both rotated and scaled. Hence, the entire data set is comprised of 2,048 samples.

TABLE VI  
RECOGNITION ACCURACY ON BRODATZ TEXTURES

Approach	Performance(%)
<b>DRLTP</b>	<b>93.30 ± 0.01</b>
<b>DRLBP</b>	<b>87.10 ± 0.01</b>
WLD [7]	97.50 ± 0.60
Gabor [32]	95.90 ± 0.90
Urbach [54]	96.50 ± 0.60
LBP [39]	91.20 ± 0.70

TABLE VII  
RECOGNITION ACCURACY ON KTH-TIPS2-A TEXTURES

Approach	DRLTP	DRLBP	WLD [7]	LBP [39]	SIFT [30]
<b>Performance (%)</b>	<b>62.6</b>	<b>59.0</b>	56.4	49.9	52.7

We perform experiments with 10-fold cross validation [7]. In each round, the samples in each class are randomly divided into 2 subsets of the same size. One is used for training and the other for testing. The results are reported as the average accuracy value and standard deviation over 10 runs. Table VI shows the performance of the proposed features against some other approaches. The performance of both proposed descriptors is poorer than other state-of-the-art approaches. There are 2 reasons for this. Firstly, the problems addressed in this paper such as the lighting reversal and the relative reversal of gray levels between object and background do not exist in this data set. Hence, the advantages of the proposed approaches are not shown in this data set. Secondly, the textures have huge rotation and scaling variances. Since our features are neither rotation-invariant nor adept to handle scale variations, the performance is adversely affected.

#### H. Comparisons With Other Approaches on KTH-TIPS2-a

The KTH-TIPS2-a database contains four physical, planar samples of each of 11 materials under varying illumination, pose, and scale. The images are  $200 \times 200$  pixels in size. Images not of this size were removed from the training and testing [7]. The database contains images at nine scales, under four different illumination directions, and three different poses. There are a total of 4395 images in the data set.

Similar to [6], [7], only three samples are used for training of each material, while testing is performed on all images of the remaining samples. The experiment is repeated four times by randomly selecting three different samples for training. The results are reported as the average value over the four runs. Table VII shows the performance of the proposed features against some other approaches. DRLBP and DRLTP outperforms WLD and LBP in comparison to Brodatz because there are neither extreme rotations nor scaling variations in this data set. Furthermore, this data set contains textures set against backgrounds and the lighting reversal and the relative reversal of gray levels between object and background exists in this data set. Hence, the performance of the descriptors are much better than the state-of-the-art approaches.

## IV. CONCLUSION

This paper proposes 2 sets of novel edge-texture features, Discriminative Robust Local Binary Pattern (DRLBP) and Ternary Pattern (DRLTP), for object recognition. The limitations of existing texture features, Local Binary Pattern (LBP), Local Ternary Pattern (LTP) and Robust LBP (RLBP), for object recognition are analyzed. LBP and LTP differentiate a bright object against a dark background and vice-versa. This differentiation makes the object intra-class variations larger. RLBP solves the LBP problem by choosing the minimum of a LBP code and its complement. However, RLBP maps LBP codes and their complement in the same block to the same value. This causes some structures to be misrepresented. Furthermore, LBP, LTP and RLBP discard contrast information. This is not desired as object texture and contour both contain discriminative information. By capturing only the texture information, the contour is not effectively represented. The new features, DRLBP and DRLTP, are proposed by analyzing the weaknesses of LBP, LTP and RLBP. They alleviate the problems of LBP, LTP and RLBP by considering both the weighted sum and absolute difference of the bins of the LBP and LTP codes with their respective complement/inverted codes. The new features are robust to image variations caused by the intensity inversion and are discriminative to the image structures within the histogram block.

We present results of the proposed features on 7 data sets and compare them with several methods for object recognition. Results demonstrate that the proposed features outperform the compared recognition approaches on most data sets.

## REFERENCES

- [1] S. Agarwal, A. Awan, and D. Roth, "Learning to detect objects in images via a sparse, part-based representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 11, pp. 1475–1490, Nov. 2004.
- [2] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
- [3] H. Bay, A. Ess, T. Tuytelaars, and L. J. V. Gool, "Speeded-up robust features (surf)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, 2008.
- [4] O. Boiman, E. Shechtman, and M. Irani, "In defense of nearest-neighbor based image classification," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.
- [5] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. New York, NY, USA: Dover Publications, Aug. 1999.
- [6] B. Caputo, E. Hayman, and P. Mallikarjuna, "Class-specific material categorisation," in *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2, Oct. 2005, pp. 1597–1604.
- [7] J. Chen *et al.*, "WLD: A robust local image descriptor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1705–1720, Sep. 2010.
- [8] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, pp. 886–893.
- [9] H. Deng, W. Zhang, E. Mortensen, T. Dietterich, and L. Shapiro, "Principal curvature-based region detector for object recognition," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
- [10] P. Dollar, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 743–761, Apr. 2012.
- [11] L. Fei-fei, R. Fergus, and P. Perona, "One-shot learning of object categories," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 4, pp. 594–611, Apr. 2006.
- [12] R. Fergus, P. Perona, and A. Zisserman, "Object class recognition by unsupervised scale-invariant learning," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2003, pp. 264–271.
- [13] A. Fernández, M. Álvarez, and F. Bianconi, "Texture description through histograms of equivalent patterns," *J. Math. Imag. Vis.*, vol. 45, no. 1, pp. 1–27, 2012.
- [14] J. Gall and V. Lempitsky, "Class-specific hough forests for object detection," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1022–1029.
- [15] P. Gehler and S. Nowozin, "On feature combination for multiclass object classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2009, pp. 221–228.
- [16] C. Geng and X. Jiang, "Face recognition based on the multi-scale local image structures," *Pattern Recognit.*, vol. 44, nos. 10–11, pp. 2565–2575, 2011.
- [17] G. Griffin, A. Holub, and P. Perona, "Caltech-256 object category dataset," California Inst. Technol., Pasadena, CA, USA: Tech. Rep. 7694, 2007.
- [18] Z. Guo, L. Zhang, and D. Zhang, "A completed modeling of local binary pattern operator for texture classification," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1657–1663, Jun. 2010.
- [19] A. Hadid, "The local binary pattern approach and its applications to face analysis," in *Proc. 1st Workshops Image Process. Theory, Tools Appl.*, Nov. 2008, pp. 1–9.
- [20] C. He, T. Ahonen, and M. Pietikainen, "A Bayesian local binary pattern texture descriptor," in *Proc. IEEE Int. Conf. Pattern Recognit.*, Dec. 2008, pp. 1–4.
- [21] M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern Recognit.*, vol. 42, no. 3, pp. 425–436, 2009.
- [22] C. K. Heng, S. Yokomitsu, Y. Matsumoto, and H. Tamura, "Shrink boost for selecting multi-LBP histogram features in object detection," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 3250–3257.
- [23] S. Hussain and B. Triggs, "Visual recognition using local quantized patterns," in *Proc. Eur. Conf. Comput. Vis.* (Lecture Notes in Computer Science). Berlin, Germany: Springer-Verlag, 2012, pp. 716–729.
- [24] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2006, pp. 2169–2178.
- [25] K. Levi and Y. Weiss, "Learning object detection from a small number of examples: The importance of good features," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2004, pp. 53–60.
- [26] S. Li, R. Chu, S. Liao, and L. Zhang, "Illumination invariant face recognition using near-infrared images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 627–639, Apr. 2007.
- [27] S. Liao, M. Law, and A. Chung, "Dominant local binary patterns for texture classification," *IEEE Trans. Image Process.*, vol. 18, no. 5, pp. 1107–1118, May 2009.
- [28] S. Liao, G. Zhao, V. Kellokumpu, M. Pietikainen, and S. Li, "Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 1301–1306.
- [29] Z. Lin and L. Davis, "Shape-based human detection and segmentation via hierarchical part-template matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 4, pp. 604–618, Apr. 2010.
- [30] D. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
- [31] S. Maji, A. C. Berg, and J. Malik, "Efficient classification for additive kernel svms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 66–77, Jan. 2013.
- [32] B. Manjunath and W. Ma, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 8, pp. 837–842, Aug. 1996.
- [33] J. Maver, "Self-similarity and points of interest," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1211–1226, Jul. 2010.
- [34] Z. Miao and X. Jiang, "Interest point detection using rank order log filter," *Pattern Recognit.*, vol. 46, no. 11, pp. 2890–2901, 2013.
- [35] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005.
- [36] J. Mutch and D. Lowe, "Multiclass object recognition with sparse, localized features," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, Jun. 2006, pp. 11–18.
- [37] D. T. Nguyen, Z. Zong, P. Ogunbona, and W. Li, "Object detection using non-redundant local binary patterns," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 4609–4612.

- [38] T. Ojala, M. Pietikainen, and T. Maenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.
- [39] T. Ojala, K. Valkealahti, E. Oja, and M. Pietikäinen, "Texture discrimination with multidimensional distributions of signed gray-level differences," *Pattern Recognit.*, vol. 34, no. 3, pp. 727–739, 2001.
- [40] C. Papageorgiou and T. Poggio, "A trainable system for object detection," *Int. J. Comput. Vis.*, vol. 38, no. 1, pp. 15–33, Jun. 2000.
- [41] A. Porebski, N. Vandenbroucke, and L. Macaire, "Haralick feature extraction from LBP images for color texture classification," in *Proc. 1st Workshops Image Process. Theory, Tools Appl.*, Nov. 2008, pp. 1–8.
- [42] J. Ren, X. Jiang, and J. Yuan, "Noise-resistant local binary pattern with an embedded error-correction mechanism," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 4049–4060, Oct. 2013.
- [43] P. Sabzmeydani and G. Mori, "Detecting pedestrians by learning shapelet features," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
- [44] A. Satpathy, X. Jiang, and H.-L. Eng, "Human detection by quadratic classification on subspace of extended histogram of gradients," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 287–297, Jan. 2014.
- [45] A. Satpathy, X. Jiang, and H.-L. Eng, "Extended histogram of gradients feature for human detection," in *Proc. IEEE Int. Conf. Image. Process.*, Sep. 2010, pp. 3473–3476.
- [46] A. Satpathy, X. Jiang, and H.-L. Eng, "Extended histogram of gradients with asymmetric principal component and discriminant analysis for human detection," in *Proc. IEEE Can. Conf. Comput. Robot. Vis.*, May 2011, pp. 64–71.
- [47] A. Satpathy, X. Jiang, and H.-L. Eng, "Human detection using discriminative and robust local binary pattern," in *Proc. IEEE Int. Conf. Acoust. Speech. Signal Process.*, May 2013, pp. 2376–2380.
- [48] A. Satpathy, X. Jiang, and H.-L. Eng, "Visual object detection by parts-based modeling using extended histogram of gradients," in *Proc. IEEE Int. Conf. Image. Process.*, Sep. 2013, pp. 2738–2742.
- [49] H. J. Seo and P. Milanfar, "Training-free, generic object detection using locally adaptive regression kernels," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1688–1704, Sep. 2010.
- [50] E. Shechtman and M. Irani, "Matching local self-similarities across images and videos," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2007, pp. 1–8.
- [51] J. Stottinger, A. Hanbury, N. Sebe, and T. Gevers, "Sparse color interest points for image retrieval and object categorization," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2681–2692, May 2012.
- [52] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1635–1650, Jun. 2010.
- [53] T. Tuytelaars and C. Schmid, "Vector quantizing feature space with a regular lattice," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.
- [54] E. Urbach, J. B. T. M. Roerdink, and M. H. F. Wilkinson, "Connected shape-size pattern spectra for rotation and scale-invariant classification of gray-scale images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 272–285, Feb. 2007.
- [55] P. Viola, M. J. Jones, and D. Snow, "Detecting pedestrians using patterns of motion and appearance," *Int. J. Comput. Vis.*, vol. 63, no. 2, pp. 153–161, 2005.
- [56] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality-constrained linear coding for image classification," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 3360–3367.
- [57] X. Wang, X. Bai, W. Liu, and L. Latecki, "Feature context for image classification and object detection," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2011, pp. 961–968.
- [58] J. Yang, K. Yu, Y. Gong, and T. Huang, "Linear spatial pyramid matching using sparse coding for image classification," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 1794–1801.
- [59] H. Zhang, A. Berg, M. Maire, and J. Malik, "SVM-KNN: Discriminative nearest neighbor classification for visual category recognition," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, vol. 2, Jun. 2006, pp. 2126–2136.
- [60] J. Zhang, J. Liang, and H. Zhao, "Local energy pattern for texture classification using self-adaptive quantization thresholds," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 31–42, Jan. 2012.

- [61] W. Zhang, S. Shan, X. Chen, and W. Gao, "Local gabor binary patterns based on Kullback-Leibler divergence for partially occluded face recognition," *IEEE Signal Process. Lett.*, vol. 14, no. 11, pp. 875–878, Nov. 2007.
- [62] G. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 915–928, Jun. 2007.



Science, Technology, and Research, Singapore, as a Scientist.

**Amit Satpathy** received the B.Eng. and Ph.D. degrees in Electrical and Electronic Engineering from Nanyang Technological University, Singapore, in 2007 and 2014, respectively. He is the recipient of the Agency for Science, Technology, and Research Graduate Scholarship. His current research focuses on feature development and extraction for object detection and recognition, image/video processing, pattern recognition, computer vision, and machine learning. He is currently with the Program Director's Office, Institute for Infocomm Research, Agency for Science, Technology, and Research, Singapore, as a Scientist.



**Xudong Jiang** (M'02–SM'06) received the B.Eng. and M.Eng. degrees in Electrical Engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, and the Ph.D. degree in Electrical Engineering from Helmut Schmidt University, Hamburg, Germany, in 1983, 1986, and 1997, respectively. From 1986 to 1993, he was a Lecturer at UESTC, where he received two Science and Technology Awards from the Ministry for Electronic Industry of China. From 1993 to 1997, he was a Scientific Assistant with Helmut Schmidt University. From 1998 to 2004, he was with the Institute for Infocomm Research, Agency for Science, Technology, and Research, Singapore, as a Lead Scientist and the Head of the Biometrics Laboratory, where he developed a system that achieved the most efficiency and the second most accuracy at the International Fingerprint Verification Competition in 2000. He was with the Nanyang Technological University (NTU), Singapore, as a Faculty Member, in 2004, and served as the Director with the Centre for Information Security from 2005 to 2011. He is currently a Tenured Associate Professor with the School of Electrical and Electronic Engineering, NTU. He has authored more than 100 papers, with 17 papers in IEEE Journals: the IEEE TRANSACTIONS ON IMAGE PROCESSING (5), the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE (4), the IEEE TRANSACTIONS ON SIGNAL PROCESSING (3), the IEEE SIGNAL PROCESSING MAGAZINE, the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS-II, and the IEEE SIGNAL PROCESSING LETTERS. He holds seven patents. His current research interests include signal/image processing, pattern recognition, computer vision, machine learning, and biometrics.



**How-Lung Eng** (M'03) received the B.Eng. and Ph.D. degrees in Electrical and Electronic Engineering from Nanyang Technological University, Singapore, in 1998 and 2002, respectively. He is currently a Chief Technology Officer with Zweek Analytics, a Singapore homegrown technology company spun out from the Institute for Infocomm Research (I2R). Prior to this, he was at I2R as a Research Scientist and Programme Manager for Video Behavioral Analytics Programme. His research interests include real-time vision, pattern classification, and machine learning for abnormal event detection. He has made several PCT filings related to video surveillance applications and has actively authored works in the above areas of interest. He was a recipient of the 2000 Tan Kah Kee Young Inventors' Award (Silver, Open Section) for his Ph.D. study, the 2002 TEC Innovator Award, the 2006 and 2008 IES Prestigious Engineering Awards, and the 2012 IWA PIA Asia Pacific Regional Award for his works in the areas of video surveillance and video analytics to ensure safe drinking water.