A Novel Integration of Intensity Order and Texture for Effective Feature Description

Thao-Ngoc NGUYEN†(a), Bac LE††, Nonmembers, and Kazunori MIYATA†, Member

SUMMARY This paper introduces a novel approach of feature description by integrating the intensity order and textures in different support regions into a compact vector. We first propose the Intensity Order Local Binary Pattern (IO-LBP) operator, which simultaneously encodes the gradient and texture information in the local neighborhood of a pixel. We divide each region of interest into segments according to the order of pixel intensities, build one histogram of IO-LBP patterns for each segment, and then concatenate all histograms to obtain a feature descriptor. Furthermore, multi support regions are adopted to enhance the distinctiveness. The proposed descriptor effectively describes a region at both local and global levels, and thus high performance is expected. Experimental results on the Oxford benchmark and images of cast shadows show that our approach is invariant to common photometric and geometric transformations, such as illumination change and image rotation, and robust to complex lighting effects caused by shadows. It achieves a comparable accuracy to that of state-of-art methods while performs considerably faster.

key words: Center-Symmetric Local Binary Pattern, feature description, image matching, texture, intensity order

1. Introduction

Local features have been widely used in computer vision, such as texture classification [1], object recognition [2], [3], pedestrian detection [4] and human action recognition [5]. Furthermore, their effectiveness in wide baseline matching [6], [7] and panorama image stitching [8] provides a strong foundation for the in-between view and 3D scene construction in virtual/augmented reality. A general pipeline of feature extraction includes finding a set of interest points, defining a region around each point in an affine-invariant manner, and computing a descriptor for the region. Correspondences between features in two images are established under some similarity measure (e.g. Euclidean distance).

There are many detectors for obtaining interest regions. Maximally Stable Extremal Regions (MSER) [9], Harris-Affine/Hessian-Affine [10], Intensity and Edge-based Regions (IBR/EBR) [11] are most preferred since they are invariant with affine transformations. Readers who are interested in this topic should refer to the survey in [12].

A robust descriptor should be highly discriminative under common image transformations (e.g. illumination change, image rotation and image blur) and simple enough for implementation in real-time systems. This paper suggests a solution that well compromises those criteria, which was put forth in [13], providing four major contributions:

1. The proposed IO-LBP operator successfully improves the Center-Symmetric Local Binary Pattern so that high performance and speed are attained. It simultaneously captures gradients and textures in a local neighborhood and halves the number of necessary neighbors. Two variants that focus on either compactness or accuracy are designed to facilitate different purposes.

2. IO-LBP patterns are aggregated into intensity orders using the region division method in [14]. Accordingly, the distinctiveness is enhanced and patterns are rotation invariant without dominant orientation estimation.

3. An associated weighting scheme is introduced to accumulate intensity changes from neighboring point pairs. It is able to improve the performance better than that of Gaussian and uniform weighting schemes while requires no additional complex computation.

4. Multi support regions are employed to further boost the performance. The final IO-LBP descriptor is built from histograms locally computed on two support regions of different sizes. While MROGH is accurate but computationally heavy, ours keeps a good trade-off of these two factors and thus is more practical.

The proposed descriptor is tested on the Oxford benchmark [15] and images of cast shadows [16]. These datasets cover challenging transformations, such as illumination changes, image rotation, image blur and cast shadows, and thus are widely used for evaluation. Experimental results show that our method performs significantly better than baseline methods and comparably to state-of-the-art ones.

This paper contains a brief review of local descriptors in Sect. 2. The introduction to IO-LBP operator and the feature construction are described in Sect. 3 and Sect. 4 respectively. Experiment setup and results are presented in Sect. 5. Finally, Sect. 6 concludes our work.

2. Related Work

There are many types of descriptors using different features, such as pixel intensity, color, gradient and edge. Gradient-based methods are popular due to their effectiveness in several image transformations. SIFT [2] distributes gradient orientations into eight bins for each cell of a Cartesian grid and outperforms conventional methods such as shape...
context, steerable filters, spin images and moment invariants [15]. SIFT is improved by using a log-polar grid [15] or applying PCA to the gradient patch of keypoints [17]. Besides, RIFT [1] adopts concentric rings of equal width, DAISY [7] uses circular regions with different Gaussian kernels, and MROGH [14] aggregates gradients into intensity orders.

Gradients are robust to several image transformations, yet unstable under complex illumination changes. Intensity order is more consistent with this challenge, and thus are being preferred. Gupta and Mittal [18] calculate a weighted sum of order flips between point-pairs selected from extremal regions. Tang et al. [19] create a 2D histogram where intensities are binned both ordinally and spatially. LIOP [20] and MROGH [14] perform region segmentation based on the intensity order, and thus avoid orientation estimation errors and significantly improve the performance. However, intensity order may be vulnerable in the presence of Gaussian noise, especially when nearby pixel values are close.

Texture is able to overcome the above issue because it does not use intensities directly but abstracts the relationship of pixel pairs [6], [21]. However, original descriptors were almost neglected due to high dimensionality until the invention of Center Symmetric Local Binary Pattern (CS-LBP) [3], which is shown to be more robust and simpler than the popular SIFT. Chan et al. [22] compute LOCP from pairwise ordinal information in an adjacent circular neighborhood. Gupta et al. [23] introduce HRI-CSLTP, a combination of relative intensities and ternary patterns. These methods generally obtain good results in various applications.

Furthermore, a single support region is usually not enough to distinguish some incorrect matches from correct ones, causing a critical issue with performance [12]. Mortensen et al. [24] augment SIFT with a global context vector that adds curvilinear shape information from a much larger neighborhood to reduce mismatches. Fan et al. [14] aggregate gradient distributions into intensity orders in multiple support regions and excellently handle various transformations, yet heavy computation arises as a side effect.

Our work is fundamentally different from above approaches. The IO-LBP operator well preserves the gradient property with four neighboring points, instead of eight in CS-LBP, and thus is more efficient. The aggregation of patterns into histograms is done by adopting the intensity order-based region division method in [14] and our novel weighting scheme. Therefore, the descriptor is rotational invariant without dominant orientation estimation [2], which is a burden to SIFT, GLOH and CS-LBP. We experimentally select two support regions to improve the performance while keeping a good balance between accuracy and computational cost. Evaluations on challenging datasets show that our method is very promising for an image matching task.

3. Intensity Order Local Binary Pattern

3.1 Center-Symmetric Local Binary Pattern (CS-LBP)

CS-LBP [3] is a robust texture operator proposed by Heikkila et al. For a pixel, it compares pairs of center-symmetric neighboring points:

\[
CS - LBP_R = \sum_{i=0}^{(N/2)-1} s(p_i - p_i(N/2)2^j),
\]

\[
s(x) = \begin{cases} 
1 & x > T, \\ 0 & \text{otherwise}
\end{cases}
\]

where \( N \) points are equally spaced on a circle of radius \( R \) and \( p_i \) is the intensity of the \( i \)-th point (cf. Fig. 1 (a)). Among variants of different \( N \) and \( R \), CS-LBP_{2,8} performs best [3].

The advantage of CS-LBP over original LBP [21] lies in three folds. First, both texture and gradient are captured. Second, it is more robust on flat regions thanks to the threshold \( T \). Third, the feature vector is compact due to fewer comparisons. CS-LBP are widely adopted in object recognition [3], human detection [4], and action recognition [5]. However, it is difficult to find an appropriate threshold \( T \) and using a small offset may lose the invariance to intensity scale transform with a multiplying constant [25].

3.2 Intensity Order Local Binary Pattern (IO-LBP)

Motivated by CS-LBP_{2,8}, we propose the Intensity Order Local Binary Pattern operator (cf. Fig. 1 (b)) as follows:

\[
IO - LBP_R = \sum_{i=0}^{N/2} s(p_i - p_i(2^1)2^j + f(p_i - p_i(2^2)2^{i+2})
\]

\[
f(x) = f(p_a - p_b) = \begin{cases} 
1 & x > \min(p_a, p_b)\tau, \\ 0 & \text{otherwise}
\end{cases}
\]

The scaling factor \( \tau \) is empirically selected in [0, 0.05], providing a good starting point for parameter tuning. The higher \( \tau \) is, the larger intensity changes are allowed without affecting the results. \( \tau = 0 \) leads to the original threshold in [21], while \( \tau > 0.05 \) begins dropping the performance. Prior works use a constant \( T \) (e.g. 0.01 [26] or [0, 0.02] [3]) or associate a similar \( \tau = 0.05 \) with center pixels [25]. Meanwhile, we associate \( \tau \) with neighboring pairs and set a stricter range so that patterns are more adaptive and tolerant to heavy transformations. Expected noise and
3.3 Symmetric Intensity Order Local Binary Pattern

The advances of IO-LBP allow us to attain excellent results under challenging transformations, as shown in Sect. 5. However, this operator is asymmetric that may cause some instability if neighboring pairs are formed from another combination. We therefore develop a symmetric variant of IO-LBP to address this issue (cf. Fig. 1(c)).

\[ IO-LBP_{sym}^R = \sum_{j=0}^{3} f(p_j - p_{(j+1)mod4})^2 + \sum_{j=0}^{3} f(p_j - p_{j+2})^2 + 4 \]  

The new variant originally has 64 binary codes, yet the number of codes can be reduced to only 24 codes by applying the transitivity property again. Readers are referred to the previous section for the validity of this reduction. IO-LBP_{sym} integrates more textures to a pattern, and hence it is expected to improve the matching quality to some degree.

4. The IO-LBP Descriptor

We first obtain detected regions of \( N \) support sizes from an interest point, normalize them to local patches of uniform size, and segment each patch according to the overall intensity order. IO-LBP patterns are computed and pooled into corresponding segments. The descriptor is finally built by concatenating histograms in all segments of \( N \) support regions. The process of feature construction is outlined in Fig. 4. Note that the term “IO-LBP” in this section refers to both IO-LBP and IO-LBP_{sym} variants.

4.1 Interest Region Detection

The input image is first smoothed by a Gaussian filter \( \sigma_{\eta} \). Interest regions are then detected using an affine-covariant detector (e.g., Harris-Affine, Hessian-Affine). Since these regions usually vary in sizes and shapes, they are normalized to circular regions of uniform diameter [15]. Finally, another Gaussian filter \( \sigma_{\eta} \) eliminates noise that may occur during the bilinear interpolation. Smoothing steps are essential because the intensity order is usually sensitive to noise.

4.2 Region Division

To improve the distinctiveness, the descriptor usually divides a local patch into segments and collects features computed locally in each segment. Spatial-based division approaches are most common, which use a Cartesian grid (e.g. 4 \( \times \) 4 grid in SIFT [2], CS-LBP [3] and HRI-CSLTP [23]) or a log-polar grid (e.g. GLOH [15] with three radial bins and eight angular directions). The major drawback is that dominant orientation estimation must be adopted to achieve rotation invariance. However, experiments on both synthetic [27] and real [14] data show that the process is error-prone and computationally heavy, and thus it greatly affects the descriptor performance. RIFT [1] addresses the issue by computing features separately within each of five concentric rings. Nevertheless, this ring-shaped segmentation is less discriminative than the grid-shaped one.

In this paper, we adopt the intensity order-based region division method from Fan et al. [14]. Pixels in the local intensity transform may be good to find \( \tau \) theoretically, yet involve advanced image analysis, and thus are left for future work.

IO-LBP has three properties that support its effectiveness. First, only four neighbors are used to reveal the gradient property, two center-symmetric pairs depicting horizontal and vertical and two non-center-symmetric ones for diagonals, and thus IO-LBP is computationally simpler than CS-LBP. Second, the threshold is locally set for each neighboring pair to maintain the invariance under intensity scale transform by multiplying a constant. Figure 2 shows that both CS-LBP and IO-LBP are robust to noise but IO-LBP is more discriminative than CS-LBP. Third, the number of dimensions is reduced from 16 to 12. Because four binary codes, 0011, 1011, 0100 and 1100, appear with very low rates (0.005%), they are replaced by the nearest valid codes. We experimentally check the validity of this assumption by counting the number of patterns on images obtain from [16] (cf. Sect. 5.2) and the result is shown in Fig. 3. It can also be done by reasoning on the transitivity of \((p_0, p_1, p_2)\) using Eq. (2).
patch are first sorted by intensity and arranged into K segments according to the intensity order. Figure 4 describes segments in different colors. Therefore, our descriptors are rotation invariant without dominant orientation estimation and stable to monotonic intensity changes. Moreover, it captures more spatial information. We experimentally compare the performance of IO-LBP descriptors with that of descriptors using either a spatial-based or an intensity-based method, and hence the effect of each division method is evaluated.

4.3 Feature Construction and Multiple Support Regions

We compute IO-LBP histograms separately in K segments:

\[ H_k = \text{hist}_{p \in \text{segment}_i} (IO - LBP(p)) \]  

(4)

where \( \text{hist} \) denotes the histogram creation. Because gradient magnitudes are excluded in binary codes, IO-LBP cannot distinguish textures having same codes but different magnitudes. We associate each pattern with a value \( w \) denoting the sum of gradient magnitudes, and therefore it has a specific level of contribution to the histogram \( H \).

\[ w(p) = \sum_{i=1}^{P} |p_i^a - p_i^b| \]  

(5)

where \( P \) is the number of neighboring pairs (\( P = 4 \) in IO-LBP and \( P = 6 \) in IO-LBP\(^{sym}\)), \( p^a \) and \( p^b \) correspond to the intensities of two points in a neighboring pair. We then improve the component inside the \( \text{hist} \) function (cf. Eq. (4)) from \( IO - LBP(p) \) to \( w(p)IO - LBP(p) \). According to Fig. 5 (c), the proposed weighting scheme gives better performances than that of uniform and Gaussian weighting schemes.

As keeping a pattern rotational invariant, for a pixel \( p \), the first neighbor is located along the radial direction from the center of the local patch to \( p \). Other three neighbors are anticlockwise sampled. Our method requires no interpolation because it is not spatial-based and IO-LBP is quantized by its nature. Therefore, it achieves greater efficiency than SIFT using tri-linear interpolation.

The above steps are carried out on \( N \) support regions of different sizes, because a single support region is usually not enough to distinguish incorrect matches from correct ones. Two non-corresponding points may be coincidentally similar to each other or two corresponding ones may be ignored due to localization error. Multi regions can resolve this problem, and thus suffice to boost the performance of a region detector [12] or a feature descriptor [14].

The IO-LBP descriptor is finally defined as follows:

\[ IO - LBP\text{Descriptor} = \bigoplus_{n=1}^{N} \bigoplus_{k=1}^{K} H_{n_k} \]  

\[ H_{n_k} = \text{hist}_{p \in \text{segment}_n} (w(p)IO - LBP(p)) \]  

(6)

where \( \bigoplus \) denotes the concatenation operator, \( N \) and \( K \) is number of support regions and segments respectively. Experimental results in Sect. 5 shows that our descriptors use a
smaller $N$ than that of MROGH ($N = 2$ vs. $N = 4$) and well approximate MROGH. Since the core feature of MROGH is the gradient, this observation indicates the advantages of LBP-based texture over gradient in terms of distinctiveness and computation cost. While the precursor CS-LBP outperforms the gradient in SIFT [3], we again observe the superiority of IO-LBP over the gradient in MROGH.

5. Experiments

5.1 Datasets and Evaluation Criterion

The evaluation data includes the Oxford benchmark [15] (cf. Fig. 9) and images of cast shadows [16] (cf. Fig. 10). The former covers various geometric and photometric transformations, such as viewpoint change, zoom and rotation, image blur, illumination change and JPEG compression. The latter illustrates shadow casting effects when the light direction is moving, and hence is used to examine the descriptor performance under complex illumination changes.

We adopt the evaluation criterion from Mikolajczyk and Schmid [15], which is based on the number of correct matches and false matches for a pair of images. The nearest neighbor matching strategy on the Euclidean distance decides whether two regions are matched. The number of correct matches and correspondences are determined by the overlap error $\epsilon$ [12]. We assume a correct match if $\epsilon < 0.5$. The result is presented in terms of recall versus 1-precision:

Fig. 6 Experimental results on the Oxford benchmark [15] (Part I). Note that scales are different through figures to improve the clarity of the plots.
Fig. 7 Experimental results on the Oxford benchmark [15] (Part II). Note that scales are different through figures to improve the clarity of the plots.

\[
\text{recall} = \frac{\text{# correct matches}}{\text{# correspondences}} \tag{7}
\]

\[
1 - \text{precision} = \frac{\text{# false matches}}{\text{# correct matches} + \text{# false matches}} \tag{8}
\]

where \text{# correspondences} stands for the ground truth number of matching regions in a pair of images. The curves are obtained by varying the distance threshold. A good descriptor should have recall approaching to 1 at any precision.

5.2 Parameter Setting Selection

There are six parameters: 1) the smoothing \(\sigma_a\) on the input, 2) the smoothing \(\sigma_b\) after normalization, 3) the number of segments \(K\), 4) the number of support regions \(N\), 5) the sampling radius \(R\), and 6) the threshold \(\tau\). We examine the effectiveness of different parameter settings and weighting schemes on the data obtained from [16], which mainly contains zoom and rotation transformations. This dataset is separated from evaluation data to avoid bias.

According to Fig. 5 (a) and Fig. 5 (b), IO-LBP and IO-LBP \(^s\) perform better when the number of segments or support regions increases. \(K = 6\) and \(K = 8\) are comparable to each other and they are much better than \(K = 4\). Meanwhile, \(N = 2\) outdistances \(N = 1\), which confirms the advantage of multi support regions. However, the improvement becomes minor when \(N\) increases to 3 while computational cost and dimensionality raise dramatically. While IO-
LBP aims to balance the computational cost and accuracy. IO-LBP\textsuperscript{sym} places greater emphasis on accuracy. Therefore, two different parameter settings are selected for these descriptors (cf. Table 1) and are used in all subsequent experiments. The effects of weighting schemes are also investigated. Figure 5(c) shows that, compared to Gaussian and

| Parameter | Patterns | $\sigma_a$ | $\sigma_b$ | $K$ | $N$ | $R$ | $\tau$ |
|-----------|----------|-----------|-----------|-----|-----|-----|-------|
| IO-LBP    | 12       | 1.0       | 1.2       | 6   | 2   | 4   | 0.01  |
| IO-LBP\textsuperscript{sym} | 24       | 1.0       | 1.2       | 4   | 2   | 4   | 0.01  |
uniform schemes, ours (Eq. (5)) is best suited for IO-LBP descriptors.

5.3 Evaluation Results and Discussions

We adopt Harris-Affine (HarAff) and Hessian-Affine (HesAff) to detect interest regions because of their wide applications. Other affine-covariant detectors are also possible. All detected regions of different sizes are normalized to circular regions of uniform diameter (41 pixels) for a fair evaluation. We select five descriptors that are closely related to ours for comparison (see Table 2). SIFT [2] and DAISY [7] are state-of-the-art gradient-based descriptors, HRI-CSLTP [23] is texture-based, LIOP [20] and MROGH [14] share a similar region division method. Binaries and evaluation codes are downloaded from [14], [15], [20]. All experiments are conducted on an Intel Core i7 2.8GHz CPU desktop.

Results on the Oxford dataset are shown in Fig. 6 and Fig. 7. The evaluation is done on 1st vs. 2nd and 1st vs. 4th image pairs, which depict small and large transformations respectively. The descriptors are more unstable in 1st vs. 4th due to heavy distortions. IO-LBP and IO-LBP\textsuperscript{sym} perform consistently on both HarAff and HesAff, and hence are independent of detectors. IO-LBP\textsuperscript{sym} keeps pace with MROGH in most cases and they alternatively occupy the first place. The largest loss of IO-LBP\textsuperscript{sym} to MROGH is 5% in trees, in which heavy blur makes textures very similar to each other. This strongly confirms the advantage of a texture-gradient combination over gradient only. Besides, it is worth noting that both IO-LBP descriptors are over two times faster than MROGH. IO-LBP ranks third in 30/32 cases, except in haraff-wall 1-4 and haraff-trees 1-4. In zoom-rotation and viewpoint change, IO-LBP is higher than the fourth place with 5–12% gain, while the ranking of other methods varies through scenarios. In illumination change and JPEG compression, IO-LBP performs better than LIOP (0.5–4%) and outdistances other methods. IO-LBP and LIOP well detect large intensity changes since they are intensity order-based. However, the texture-gradient combination is more informative than permutations of neighboring points. trees is the most challenging scenario, in which all methods deteriorate severely. IO-LBP maintains good performance with a 10% gain compared to that of other methods, yet it fails in haraff-trees 1-4. Since HRI-CSLBP follows IO-LBP closely, the key of success in this scenario must be texture.

Figure 8 shows results on images of cast shadows. The overall performance is averaged over all image pairs of a sequence. IO-LBP, IO-LBP\textsuperscript{sym}, and MROGH again perform best and keep pace with each other. Our descriptors achieve performance gains of 3 – 5% and 15% compared to that of the fourth and the fifth places respectively. IO-LBP has only one failure in haraff-toy, yet IO-LBP\textsuperscript{sym} can compensate for this weakness. Shadows tend to make the affected region darker than other regions with a scale factor but preserves the texture information [25]. Since our descriptors adopt IO-LBP textures and an adaptive threshold, they are best suited for this challenge. Therefore, they are promising for tasks suffering heavy cast shadows, such as background subtraction.

In conclusion, an appropriate combination of segmentation strategy and feature is essential to the performance. Furthermore, spatial-based region division is still competitive, yet it is difficult to find a good dominant orientation. Our descriptors use the robust IO-LBP textures and the region division method in [14] with two support regions to effectively handle challenging transformations. They achieve high recalls from the outset, and thus are promising for image retrieval. Nonetheless, failures in haraff-wall 1-4 and haraff-trees 1-4 indicate their weakness in handling repetitive patterns and heavy blur. Resolving the sensitivity of intensity order in near-uniform regions is our future work.

6. Conclusion

This paper introduces an approach of feature description...
by combining the novel IO-LBP texture with current techniques of region division and multi support regions. IO-LBP and IO-LBP∗ operators creatively reduce the number of neighboring points while preserving the gradient property, hence they are robust and computationally simple. The proposed IO-LBP descriptors performs stably under several image transformations and cast shadows. They are designed for either accuracy or compactness, providing flexibility to different purposes of usage. Experiments on reliable datasets confirm their advantages over modern methods.

References

[1] S. Laezebnik, C. Schmid, and J. Ponce, “A sparse texture representation using local affine regions,” IEEE Trans. Pattern Anal. Mach. Intell., vol.27, no.8, pp.1265–1278, 2005.
[2] D.G. Lowe, “Distinctive image features from scale-invariant keypoints,” IJCV, vol.60, no.2, pp.99–110, 2004.
[3] 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.
[4] J. Xu, Q. Wu, J. Zhang, and Z. Tang, “Fast and accurate human detection using a cascade of boosted ms-lbp features,” IEEE Signal Process. Lett., vol.19, no.10, pp.676–679, 2012.
[5] R. Mattivi and L. Shao, “Spatio-temporal dynamic texture descriptors for human motion recognition,” in Intelligent Video Event Analysis and Understanding, pp.69–91, 2011.
[6] F. Schaaffalitzky and A. Zisserman, “Viewpoint invariant texture matching and wide baseline stereo,” Proc. ICCV, vol.2, pp.636–643, 2001.
[7] E. Tola, V. Lepetit, and P. Fua, “Daisy: An efficient dense descriptor applied to wide-baseline stereo,” IEEE Trans. Pattern Anal. Mach. Intell., vol.32, no.5, pp.815–830, 2010.
[8] M. Brown and D.G. Lowe, “Automatic panoramic image stitching using invariant features,” IJCV, vol.74, no.1, pp.59–73, 2007.
[9] J. Matas, O. Chum, M. Urban, and T. Pajdla, “Robust wide baseline stereo from maximally stable extremal regions,” Proc. BMVC, pp.384–393, 2002.
[10] K. Mikolajczyk and C. Schmid, “Scale and affine invariant interest point detectors,” IJCV, vol.60, no.1, pp.53–86, 2004.
[11] T. Tuytelaars and L.V. Gool, “Matching widely separated views based on affine invariant regions,” IJCV, vol.59, no.1, pp.61–85, 2004.
[12] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaaffalitzky, T. Kadir, and L.V. Gool, “A comparison of affine region detectors,” IJCV, vol.65, no.1-2, pp.43–72, 2007.
[13] T.N. Nguyen, B. Le, and K. Miyata, “Effective feature description using intensity order local binary pattern,” Proc. Cyberworlds, pp.175–182, 2013.
[14] B. Fan, F.C. Wu, and Z.Y. Hu, “Aggregating gradient distributions into intensity orders: A novel local image descriptor,” Proc. CVPR, pp.2377–2384, 2011.
[15] K. Mikolajczyk and C. Schmid, “A performance evaluation of local descriptors,” IEEE Trans. Pattern Anal. Mach. Intell., vol.27, no.10, pp.1615–1630, 2005.
[16] http://lear.inrialpes.fr/people/mikolajczyk/Database/
[17] Y. Ke and R. Sukthankar, “Pca-sift: a more distinctive representation for local image descriptors,” Proc. CVPR, pp.506–513, 2004.
[18] R. Gupta and A. Mittal, “Smd: A locally stable monotonic change invariant feature descriptor,” Proc. ECCV, pp.265–277, 2008.
[19] F. Tang, S. Lim, N. Chang, and H. Tao, “A novel feature descriptor invariant to complex brightness changes,” Proc. CVPR, pp.2631–2638, 2009.
[20] Z. Wang, B. Fan, and F. Wu, “Local intensity order pattern for feature description,” Proc. ICCV, pp.603–610, 2011.
[21] T. Ojala, M. Pietikäinen, and T. Mäenpää, “Multire solution gray-scale and rotation invariant texture classification with local binary patterns,” IEEE Trans. Pattern Anal. Mach. Intell., vol.24, no.7, pp.971–987, July 2002.
[22] C.H. Chan, B. Goswami, J. Kittler, and W. Christmas, “Local ordinal contrast pattern histograms for spatiotemporal, lip-based speaker authentication,” IEEE Trans. Inf. Forensics Security, vol.7, no.2, pp.602–612, 2012.
[23] R. Gupta, H. Paill, and A. Mittal, “Robust order-based methods for feature description,” Proc. CVPR, pp.334–341, 2010.
[24] E.N. Mortensen, H. Deng, and L. Shapiro, “A sift descriptor with global context,” Proc. CVPR, pp.184–190, 2005.
[25] S. Liao, G. Zhao, V. Kellokumpu, M. Pietikäinen, and S.Z. Li, “Modeling pixel process with scale invariant local patterns for background subtraction in complex scenes,” Proc. CVPR, pp.1301–1306, 2010.
[26] M. Heikkila and M. Pietikainen, “A texture-based method for modeling the background and detecting moving objects,” IEEE Trans. Pattern Anal. Mach. Intell., vol.28, no.4, pp.657–662, April 2006.
[27] S. Winder and M. Brown, “Learning local image descriptors,” Proc. CVPR, pp.1–8, 2007.

Thao- Ngoc Nguyen received her B.S. degree in Information Technology and M.S. degree in Computer Science from University of Science, Ho Chi Minh Vietnam National University, Vietnam, in 2006 and 2009. She is currently a Ph.D. candidate in School of Knowledge Science, Japan Advanced Institute of Science and Technology (JAIST). Her research interests include image feature extraction, pedestrian detection and pattern recognition.

Bac Le received his B.S. degree in Mathematics in 1984 and the M.S. degree in Computer Sciences in 1990. He received his Ph.D. in Mathematics for Computers and Computing Systems in 2000. He is currently the Head of Department of Computer Science, Vice Dean of Faculty of IT, University of Science, VNU - HCMC Vietnam. His research interests include artificial intelligent, soft computing, knowledge discovery and data mining, and data hiding.

Kazunori Miyata has been a Professor in School of Knowledge Science (2012-) and the Center for Knowledge Science (2002-2012) at the Japan Advanced Institute of Science and Technology (JAIST) since 2002. Prior to joining JAIST, he was an Associate Professor in the Department of Imaging Art at the Tokyo Institute of Polytechnics. He received a B.S. degree from Tohoku University in 1984, and a M.S. and Ph.D. from the Tokyo Institute of Technology in 1986 and 1997. His research mainly focuses on procedural modeling, texture generation, and multimedia applications. He is a member of ACM and IEEE.