

Sampled Multi-scale Color Local Binary Patterns

Yu Zhang¹, Stephane Bres² and Liming Chen¹

¹Universite de Lyon, CNRS, Ecole Centrale de Lyon, LIRIS, UMR5205, F-69134, Lyon, France

²LIRIS-INSA de Lyon, 20 Avenue Albert Einstein, 69621 Villeurbanne Cedex, France

Keywords: Local Binary Pattern, Feature Extraction, Object Recognition, Patch Reorganization.

Abstract: In this paper, we propose a novel representation, called sampled multi-scale color Local Binary Pattern (SMC-LBP), and apply it to Visual Object Classes (VOC) Recognition. The Local Binary Pattern (LBP) has been proven to be effective for image representation, but it is too local to be robust. Meanwhile such a design cannot fully exploit the discriminative capacity of the features available and deal with various changes in lighting and viewing conditions in real-world scenes. In order to address these problems, we propose SMC-LBP, which randomly samples the neighboring pixels across different scale circles, instead of pixels from individual circular in the original LBP scheme. The proposed descriptor presents several advantages: (1) It encodes not only single scale but also multiple scales of image patterns, and hence provides a more complete image information than the original LBP descriptor; (2) It cooperates with color information, therefore its photometric invariance property and discriminative power is enhanced. The experimental results on the PASCAL VOC 2007 image benchmark show significant accuracy improvement by the proposed descriptor compared with both the original LBP and other popular texture descriptors.

1 INTRODUCTION

Texture, color and local gradients features play a major role in content-based image categorization task. Identifying patches with texture features is at the heart of many computer vision algorithms. It is widely applied in object category recognition and image retrieval application(Ozuysal et al., 2010). Identifying patches is difficult because of drastic surface appearance which depends on how the image texture information is captured. To address this problem, many texture descriptors have been proposed in the literature, such as Grey Co-occurrence Matrix(GLCM)(Tuceryan and Jain, 1998), Texture Auto Correlation(TAC)(Tuceryan and Jain, 1998), Gabor filter(Zhang et al., 2000), Brief(Calonder et al., 2010) and LBP(Ojala et al., 2002).

Among all these texture features, LBP is one of the most popular texture descriptors. It was introduced and used in texture classification based on local binary patterns and nonparametric discrimination of sample and prototype distributions(Ojala et al., 2002). It can be seen as a unified approach to statistical and structural texture analysis. Fig. 1 gives an example. The LBP descriptor encodes one pixel of an image by thresholding the neighborhood of each pixels with the center value. Then the threshold results are multiplied

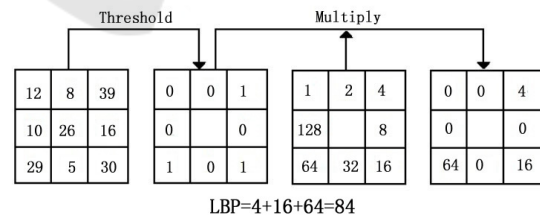


Figure 1: Calculation of the original LBP descriptor.

with weights given by powers of two. Finally the LBP code is obtained by summing up all the weighted results. This process is done for each pixel, and the image representation is obtained by counting the histogram based on these codes. The LBP descriptor is further extended to multi-scale using a circular neighborhood with variant radius and variant number of neighboring pixels, as shown in Fig 2.

Because of its descriptive power for analyzing both micro and macro texture structures, and computational simplicity, LBP has been widely applied for texture classification(Ojala et al., 2002) and object recognition(Zhu et al., 2010), and is demonstrated excellent results and robustness against global illumination changes. It has also been used successfully for texture segmentation(Blas et al., 2008), recognition of facial identity(Guo et al., 2010) and expression(Shan et al., 2009).

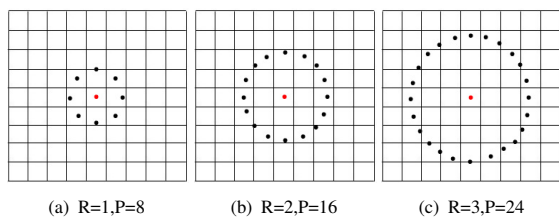


Figure 2: Single-scale LBP descriptor.

However, the original LBP descriptor also has several drawbacks in its application. It covers a small spatial support area, hence the bit-wise comparisons are made through single circular pixel values with the central of pixel. This means that the LBP codes are easily affected by noise(Liao et al., 2007). Moreover, features calculated in a single circular neighborhood cannot capture larger scale structure (macrostructure) that may be dominant features. Meanwhile, the original LBP descriptor ignores all color information (its calculation is based on gray image), while color plays an important role for distinction between objects, especially in natural scenes(Zhu et al., 2010). There can be various changes in lighting and viewing conditions in real-world scenes, leading to large variations of objects in surface illumination, scale, etc., which make the original LBP performance is not very good in VOC recognition tasks. In order to address these drawbacks, many improve method of LBP descriptors have been proposed, such as Multi-scale Block LBP(Liao et al., 2007), Hierarchical Multi-scale LBP(Guo et al., 2010), Multi-scale Color LBPs(Zhu et al., 2010) and so on.

Traditionally, in order to capture larger scale structure (macrostructure), there are the histogram fusion and extending radius approaches to be proposed. The histogram concatenation approach. Firstly, LBP features of different scale are extracted, and then the histograms are concatenated into a long feature. Vector joint distribution could contain more information, but it suffers from huge feature dimension. Meanwhile this approach can not represent the image with the histogram uniquely. Usually, considering bigger neighborhood (more neighboring pixels with bigger radius) could lead to better performance because more local information is obtained. However, the drawback lies in the high dimensional histogram produced by the LBP codes. According to the definition, if the length of binary bitstring is p , the resulting histogram will be of 2^p dimension. The dimensionality growth is exponential when the number of neighboring pixels is increasing, and it is impractical to feed the histograms with such huge dimension into the classifier for classification. Although many approaches reduced the dimension(e.g. ri, u2(Ojala et al., 2002)) were pro-

posed, the drawback are still not solved completely.

In this work, we propose a novel representation, called Sample Multi-scale Color Local Binary Pattern (SMC-LBP), to overcome the mentioned limitations of LBP and extend the LBP feature to patch. To validate the proposed feature, we apply it to VOC Recognition problem. In SMC-LBP, the computation is done based on randomly sampling the neighboring pixels from multi-scale circles. Furthermore, in order to enhance photometric invariance property and discriminative power, the proposed descriptor is computed in different color spaces. To summarize, the SMC-LBP descriptor presents several advantages:

- It encodes not only single scales but also multiple scale of image patterns, extends the LBP to the patch, and hence provides a more complete image representation than the original LBP descriptor.
- It incorporates with color information, therefore its photometric invariance property and discriminative power are enhanced.

In section 2, we introduce Sample Multi-scale Binary Pattern in detail. Section 3 presents Sample Multi-scale Color Local Binary Pattern. The Framework of the experiment is introduced in section 4. The experimental results are shown in section 5. Finally some conclusions and future work are given in section 6.

2 SAMPLE MULTI-SCALE LOCAL BINARY PATTERN

2.1 SM-LBP Approach

Our approach is inspired by earlier work(Ozuysal et al., 2010) that image patches could be effectively classified on the basis of a relatively small number of pairwise intensity comparisons(Calonder et al., 2010). Here we randomly sample across different scale circles, as shown in Fig 3 and is further extended to use the circular neighborhood with variant radius and variant number of neighboring pixels.

More specifically, the SM-LBP descriptor at pixel location $g_c(x_c, y_c)$ is defined as follows:

$$SM-LBP_N := \sum_{n=1}^N \tau(g_c, g_n) 2^n \quad (1)$$

where $g_n(g_1, g_2, \dots, g_N)$ $N = 2\pi/\theta$, g_n is the pixel gray value of the multi-scale circular neighborhood. N is the number of neighbor pixels which we randomly choose from different scale circles. How to generate the g_n is introduced in the next section.

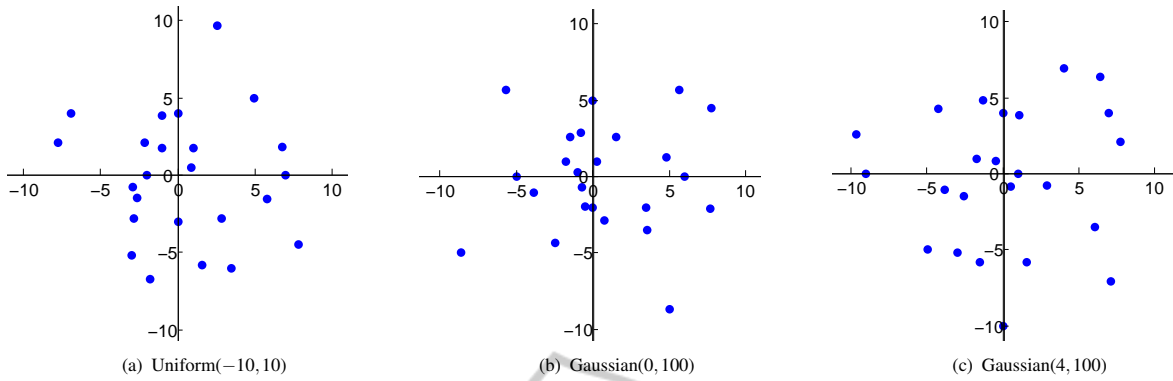


Figure 3: Different approaches to choosing the g_n location. All the radius R are selected by randomly sampling from different circles.

$$\tau(g_c, g_n) = \begin{cases} 1 & \text{if } g_c < g_n \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

Compared to the original LBP, the MS-LBP replaces comparisons between the central pixel and single circular pixels with comparisons between the central pixel and the pixels which are randomly chosen from multi-scale circles. In this way, the neighboring pixels randomly chosen could come from the different scales, this means that our new descriptor can capture more information from larger region. In this paper, the follow experiments we consider $N = 8, 16, 24$; $g_n \in \{R_1, R_2, R_3, R_4, R_5\}$.

2.2 Sample Arrangement of SM-LBP

There are many options for generating the radius R_n from different distributions. We experiment with three sampling approaches. Assuming the origin of the patch coordinate system to be located at the patch center. The patch size S is $\max(R_i)$. The center point $g_c(x_c, y_c)$ to be located at the patch center, $g_n(x_n, y_n)$ are given by $(-R_n \sin(2\pi n/N), R_n \cos(2\pi n/N))$, R_n can be described as follows.

- $R_n \sim$ i.i.d. Uniform($-R, R$): The g_n locations are evenly distributed over the patch.
- $R_n \sim$ i.i.d. Gaussian($0, S^2$), the radius R_n is sampled from a Gaussian distribution with mean parameter 0 and standard deviation parameter S^2 centered around the origin g_c . This forces the g_n to be more local. g_n locations outside the patch are clamped to the edge of the patch.
- $R_n \sim$ i.i.d. Gaussian(R_i, S^2), the radius R_n is sampled from a Gaussian distribution with mean parameter $R_i \neq 0$ and standard deviation parameter S^2 centered around the origin g_c . g_n locations outside the patch are clamped to the edge of the patch.

3 SAMPLE MULTI-SCALE COLOR LOCAL BINARY PATTERN

3.1 Model Analysis for Illumination Changes

The VOC task is important to access visual information on the level of objects and scene types(van de Sande et al., 2010). In order to enhance the descriptor's illumination invariance and discriminative power, we farther proposed color MS-LBP, called SMC-LBP. The diagonal model (3) and the diagonal-off model (4) can be used to model changes in the illumination(van de Sande et al., 2010).

$$\begin{pmatrix} R^c \\ G^c \\ B^c \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} R^u \\ G^u \\ B^u \end{pmatrix} \quad (3)$$

$$\begin{pmatrix} R^c \\ G^c \\ B^c \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \begin{pmatrix} R^u \\ G^u \\ B^u \end{pmatrix} + \begin{pmatrix} O_1 \\ O_2 \\ O_3 \end{pmatrix} \quad (4)$$

where u is an light source, and c is the canonical illumination. The eq. (4) presents that maps colors that are taken under an unknown light source to their corresponding colors under the canonical illumination(Ozuysal et al., 2010). In order to deal with a wider range of imaging conditions, Finlayson et al extend the diagonal model to the diagonal-off model with an offset $(O_1, O_2, O_3)^T$ (Finlayson et al., 2005).

Based on above two models, illumination change can be defined. If a constant factor in all channels ($a = b = c$) In eq. (3), it presents the light intensity change; If Image values change by an equal offset in all channels ($a = b = c = 1, O_1 = O_2 = O_3$) in eq. (4),

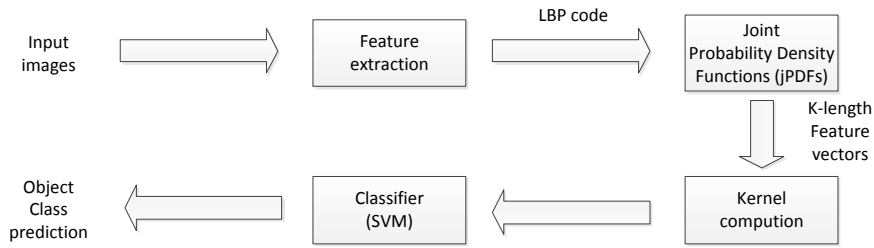


Figure 4: Flow chart of our system for VOC.

it presents light intensity shift. If $(a = b = c, O_1 = O_2 = O_3)$ in eq. (4), it means light intensity change and shift. light color change depend on all channels independently $(a \neq b \neq c)$, as eq. (3) and light color change depend on all channels independently with arbitrary offsets $(a \neq b \neq c, O_1 \neq O_2 \neq O_3)$, as eq. (4).

3.2 SMC-LBP Descriptors

In order to enhance SM-LBP's photometric invariance property and discriminative power, three color SMC-LBP descriptors are proposed. The main idea is to compute the SMC-LBP descriptor independently over all the channels of certain color space.

RGB-SM-LBP This descriptor is obtained by computing LBP over all three channels of the RGB color space independently, and then concatenating the results together. It is invariant to monotonic light intensity change due to the property of the original LBP, and has no additional invariance properties.

Opponent-SM-LBP This descriptor is obtained by computing LBP over all three channels of the opponent color space:

$$\begin{pmatrix} O_1 \\ O_2 \\ O_3 \end{pmatrix} = \begin{pmatrix} (R - G)/\sqrt{2} \\ (R + G - 2B)/\sqrt{6} \\ (R + G + B)/\sqrt{3} \end{pmatrix} \quad (5)$$

Due to the subtraction, O_1 and O_2 channels are invariant to light intensity shift. O_3 channel represents the intensity information, and has no invariance properties.

Hue-SM-LBP This descriptor is obtained by computing LBP for the Hue channel of the HSV color space:

$$\text{Hue} = \arctan\left(\frac{O_1}{O_2}\right) = \arctan\left(\frac{\sqrt{3}(R - G)}{R + G - 2B}\right) \quad (6)$$

Due to the subtraction and the division, Hue channel is scale-invariant and shift-invariant, therefore this descriptor is invariant to light intensity change and shift.

4 THE FRAMEWORK OF VOC

Our framework for VOC is depicted in Fig. 4

4.1 Feature Extraction

The SM-LBP descriptors extracted from input images at every pixel location as their features. With the radius R_n which is sampled from a Gaussian or Uniform, the neighboring pixels $g_n(x_n, y_n)$ is generated. By this way, the LBP descriptor is extended to use the multi-circular neighborhood with variant radius and variant number of neighboring pixels. It is more suitable for VOC task. Moreover, in order to increase photometric invariance property and discriminative power of the SM-LBP descriptors, The SMC-LBPs are used in this system.

4.2 Classification

Once all the jPDFs representations of the input images are obtained, they are then feed into certain classifier for classification. Here we apply the Support Vector Machine (SVM) for the final classification. The benefits of SVM for histogram-based classification have been clearly demonstrated in (Caputo et al., 2005).

In our experiments, the χ^2 distance is computed to measure the similarity between each pair of the feature vectors F and F' (n is the size of the feature vector):

$$\text{dist}_{\chi^2}(F, F') = \sum_{i=1}^n \frac{(F_i - F'_i)^2}{F_i + F'_i} \quad (7)$$

Then, the kernel function based on the χ^2 distance is used for SVM to train the classifier:

$$K_{\chi^2}(F, F') = e^{-\frac{1}{D} \text{dist}_{\chi^2}(F, F')} \quad (8)$$

where D is the parameter for normalizing the distances. Here D is set to the average distance of all the training data. Finally, for each test image, the output



Figure 5: Example images of the PASCAL VOC 2007 benchmark.

probabilities of SVM classifier are used to predict the object categories.

5 EXPERIMENT

We perform the VOC experiments on the standard PASCAL VOC 2007 benchmark. The dataset has 20 different object classes, such as sheep, train, boat, bus, sofa, table, etc. Some example images are shown in Fig. 5. The dataset is pre-defined into 50% for training/validation and 50% for testing. In total there are 9,963 images, where 2501 are for training, 2510 are for validation and 4952 are for test. For evaluation we use mean average precision (mAP)(Yue et al., 2007).

5.1 Experiment Results

In order to evaluate the performance of our descriptors, we compare SM-LBP and SMC-LBP descriptors with the other texture features. Meanwhile we also compare these descriptors with the SIFT(Lowe, 2004) which is one of the most powerful image descriptors.

5.1.1 Comparison with the Original LBP

The proposed SM-LBP descriptors are compared with the original LBP. In our experiment, we set the $N = 8, 16, 24$, and g_i is generated by the Gaussian distribution and the Uniform distribution. The final mAP value is obtained by the mean of 20 experimental results. Table 1 shows the comparison of proposed SM-LBP descriptors and the original LBP on PASCAL 2007. It can be seen that the SM-LBP gets the better performance of mAP. Compared with the original LBP, the SM-LBP obtains a better performance improvement (nearly 2%). Fig. 6 shows comparison of the proposed SMC-LBP descriptors and original color

LBP. It shows that the SMC-LBP all further outperform the original color LBP, with the improvements from 2% to 5.8%.

Table 1: Comparison of proposed SM-LBP descriptors and the original LBP on PASCAL 2007(original LBP: $N=8$, the circle of radius $R=1$; $N=16$, the circle of radius $R=2$; $N=24$, the circle of radius $R=3$. U,G: U is the Uniform distribution; G is the Gaussian distribution;).

mAP(%)	N=8	N=16	N=24
LBP(original)	28.40	31.64	29.78
SM-LBP(U(-5,5))	30.40	33.83	33.02
SM-LBP(G(0, 25))	30.61	33.42	33.20
SM-LBP(G(2, 25))	29.98	33.34	33.32

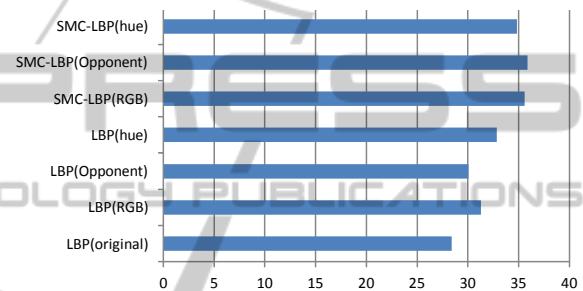


Figure 6: Comparison of the proposed SMC-LBP descriptors and original color LBP(For original color LBPs, $N=24$, $R=3$; For the SMC-LBPs, $N=24$, the distribution is Gaussian(0, 25)).

5.1.2 Comparison with other Texture Descriptors

As one kind of texture feature, SM-LBP and SMC-LBP are compared with three widely-used popular texture descriptors, including Gabor filter, Texture Auto Correlation (TAC), and Grey Level Co-occurrence Matrix (GLCM). We set 5 scales and 8 orientations for Gabor filter. For TAC, the rang of x and y directions is $[0,8]$ with interval of 2. For GLCM, 4 directions (horizontal, vertical and diagonal) with 1 offset between two pixels are considered.

From the results shown in Fig 7, it can be seen that the original LBP already outperforms other popular texture descriptors, proving that LBP is one of the best texture features available today. Our new descriptors further improve the performances to almost double of the other texture descriptors, demonstrating that the strong power of the proposed descriptors benefit from the properties of illumination-invariant and scale-invariant.

5.1.3 Comparison with SIFT Descriptor

Nowadays SIFT, a kind of local gradient descriptors

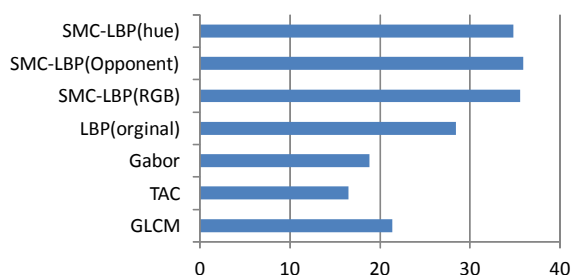


Figure 7: Comparison of the proposed SMC-LBP descriptors and other texture descriptors(SMC-LBPs, N=24, the distribution chosen Gaussian(0, 25)).

Table 2: Comparison of the proposed SMC-LBP and the SIFT(SMC-LBPs, N=24, the distribution chosen Gaussian(0, 25)).

	mAP(%)
LBP(original)	28.40
SMC-LBP(hue)	34.82
SMC-LBP(Opponent)	35.87
SMC-LBP(RGB)	35.59
SIFT	38.00

is one of the most powerful image descriptors in the literature. Comparison of the proposed SMC-LBP and the SIFT, the performance of our texture SMC-LBP descriptor is close to SIFT.

6 CONCLUSIONS

In this paper, we propose a novel SM-LBP which can obtain multi-scale patterns and provide a patch texture representation. Moreover, in order to deal with the deficiency of color information and sensitivity to non-monotonic lighting condition changes, SMC-LBP descriptor is proposed. The main contributions are that SM-LBP and SMC-LBP not only have more discriminative power by obtaining more local information, but also possess invariance properties to different lighting condition changes. In addition, they keep the advantage of computational simplicity from the original LBP descriptor. The proposed descriptors are validated by applying on the PASCAL VOC 2007 image benchmark. Compared with the original LBP, the experimental results exhibit better recognition accuracy.

REFERENCES

Blas, M. R., Agrawal, M., Sundaresan, A., and Konolige, K. (2008). Fast color/texture segmentation for outdoor robots. In *IROS*, pages 4078–4085.

- Calonder, M., Lepetit, V., Strecha, C., and Fua, P. (2010). Brief: binary robust independent elementary features. In *Proceedings of the 11th European conference on Computer vision: Part IV, ECCV'10*, pages 778–792, Berlin, Heidelberg. Springer-Verlag.
- Caputo, B., Hayman, E., and Mallikarjuna, P. (2005). Class-specific material categorisation. *IEEE International Conference on Computer Vision (ICCV)*, 2:1597–1604.
- Finlayson, G. D., Hordley, S. D., and Xu, R. (2005). Convex programming colour constancy with a diagonal-offset model. In *ICIP (3)'05*, pages 948–951.
- Guo, Z., 0006, L. Z., Zhang, D., and Mou, X. (2010). Hierarchical multiscale lbp for face and palmprint recognition. In *ICIP*, pages 4521–4524. IEEE.
- Liao, S., Zhu, X., Lei, Z., Zhang, L., and Li, S. Z. (2007). Learning multi-scale block local binary patterns for face recognition. In Lee, S.-W. and Li, S. Z., editors, *ICB*, volume 4642 of *Lecture Notes in Computer Science*, pages 828–837. Springer.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110.
- Ojala, T., Pietikäinen, M., and Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):971–987.
- Ozuysal, M., Calonder, M., Lepetit, V., and Fua, P. (2010). Fast keypoint recognition using random ferns. *IEEE Trans. Pattern Anal. Mach. Intell.*, 32(3):448–461.
- Shan, C., Gong, S., and McOwan, P. W. (2009). Facial expression recognition based on local binary patterns: A comprehensive study. *Image Vision Comput.*, 27(6):803–816.
- Tuceryan, M. and Jain, A. K. (1998). Texture analysis. In *Handbook of Pattern Recognition and Computer Vision*, pages 235–276.
- van de Sande, K. E. A., Gevers, T., and Snoek, C. G. M. (2010). Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1582–1596.
- Yue, Y., Finley, T., Radlinski, F., and Joachims, T. (2007). A support vector method for optimizing average precision. In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, SIGIR '07*, pages 271–278, New York, NY, USA. ACM.
- Zhang, D., Wong, A., Indrawan, M., and Lu, G. (2000). Content-based image retrieval using gabor texture features. In *IEEE Transactions PAMI*, pages 13–15.
- Zhu, C., Bichot, C.-E., and Chen, L. (2010). Multi-scale Color Local Binary Patterns for Visual Object Classes Recognition. In IEEE, editor, *International Conference on Pattern Recognition (ICPR)*, pages 3065–3068.