



# Combining Local Binary Pattern and Speeded-Up Robust Feature for Content-Based Image Retrieval

Prashant Srivastava<sup>1</sup>(✉), Manish Khare<sup>2</sup>, and Ashish Khare<sup>3</sup>

<sup>1</sup> NIIT University, Neemrana, Rajasthan, India  
prashant.jk087@gmail.com

<sup>2</sup> Dhirubhai Ambani Institute of Information and Communication Technology, Gandhinagar, Gujarat, India

mkharejk@gmail.com

<sup>3</sup> Department of Electronics and Communication, University of Allahabad, Allahabad, Uttar Pradesh, India

ashishkhare@hotmail.com

**Abstract.** Large number of digital image libraries containing huge amount of images have made the task of searching and retrieval tedious. Content-Based Image Retrieval (CBIR) is a field which finds solution to this problem. This paper proposes CBIR a technique which extracts interest points from texture feature at multiple resolutions of image. Local Binary Pattern (LBP) has been used to perform texture feature extraction and interest points are gathered through Speeded-Up Robust Feature (SURF) descriptors. The multiresolution decomposition of image is done using Discrete Wavelet Transform (DWT). DWT coefficients of gray scale image are computed followed by computation of LBP codes of resulting DWT coefficients. The interest points from texture image are then gathered by computing SURF descriptors of resulting LBP codes. Finally, feature vector for retrieval is constructed through Gray-Level Co-occurrence Matrix (GLCM) which is used to retrieve visually similar images. The performance of the proposed method has been tested on Corel-1 K dataset and measured in terms of precision and recall. The experimental results demonstrate that the proposed method performs better than some of the other state-of-the-art CBIR techniques in terms of precision and recall.

**Keywords:** Content-Based Image Retrieval · Local Binary Pattern · Speeded-Up Robust Feature · Gray-Level Co-occurrence matrix

## 1 Introduction

The process of image acquisition is no longer a difficult task due to the invention of low-cost image capturing devices. As a result, there are huge amount of different types of images. In order to have an easy access to these images, it is important that the images are properly organized. Image retrieval is a field which attempts to solve this problem. There are two types of image retrieval systems-text-based retrieval systems and content-based

retrieval systems. Text-based retrieval systems perform searching and retrieval through keywords and text. Manual tagging of images is required in such systems and they fail to retrieve visually similar images. Content-Based Image Retrieval (CBIR) systems search and retrieve images based on the features present in the image. There is no need to manually annotate images and the images retrieved are visually similar [1].

Most of the early CBIR techniques exploited colour feature for constructing feature vector [2]. Later on, texture and shape features were also exploited for constructing feature vector for CBIR [3, 4]. The use of single feature, however proved to be insufficient because image consists of multiple objects and all objects may not be captured efficiently by single feature. These limitations are overcome by combining colour, texture, and shape features [5]. The combination of features not only extract more than one feature from an image but also overcome limitations of each other.

For extracting features from an image, most of the CBIR techniques exploit single resolution of image. However, single resolution processing of image may not construct efficient feature vector as an image consists of varying level of images. To overcome this limitation, multiresolution processing of images is performed. Extraction of features at multiple resolutions of image is advantageous as the features that are left undetected at one scale get detected at another scale. A number of CBIR techniques exploiting multiresolution techniques have been proposed [6, 7]. A novel CBIR technique has been proposed which extracts interest points from texture feature at multiple resolutions of image. Texture feature is computed using Local Binary Pattern (LBP) which has proved to be an efficient feature descriptor for object recognition. The interest points have been obtained through Speeded-Up Robust Feature (SURF) descriptor which efficiently extracts scale invariant features from an image and is faster to compute. Feature vector construction is done by combining LBP and SURF, and this combination is exploited at multiple resolutions of image. The multiresolution decomposition of image has been done through Discrete Wavelet Transform (DWT). The proposed method computes LBP codes of DWT coefficients followed by computation of SURF descriptors of resulting LBP codes. Finally, feature vector has been constructed using Gray-Level Co-occurrence Matrix (GLCM) which provides information about how frequently pixel pairs of specified value and in a specified direction occur in an image.

The rest of the paper is organized as follows- Sect. 2 discusses related work in the field of CBIR. Section 3 gives background concept of DWT, LBP and SURF. Section 4 discusses the proposed method. Section 5 discusses experiment and results and finally, Sect. 6 concludes the paper.

## 2 Related Work

The use of feature descriptors to construct feature vector for CBIR has been extensively used in the past few years. A number of feature descriptors exploiting multiple features of image have been proposed [8–10]. Liu et al. [11] proposed Multi-Texton Histogram (MTH) which exploits colour and texture feature for CBIR. Liu et al. [12] proposed Microstructure Descriptor (MSD) which combined colour, texture and shape features for CBIR. Most of the above descriptors perform single resolution processing of image for constructing feature vector. Single resolution processing of images may not be sufficient

for extracting complex details from an image as an image consists of varying level of details. To overcome these limitations, a number of multiresolution descriptors have been proposed [13–15]. These descriptors decompose image into multiple resolutions and construct feature vector by extracting features at multiple scales of image. A principle advantage of these multiresolution feature descriptors is that the features that are left undetected at one scale get detected at another scale. This paper proposes a combination of feature descriptors which are exploited at multiple resolutions of image.

### 3 Discrete Wavelet Transform, Local Binary Pattern and Speeded-Up Robust Feature

#### 3.1 Discrete Wavelet Transform

The wavelet series expansion of function  $f(x) \in L^2(R)$  relative to the wavelet  $\psi(x)$  and scaling function  $\varphi(x)$  is defined as [16],

$$f(x) = \sum_k c_{j_0}(k)\varphi_{j_0k}(x) + \sum_{j=j_0} d_j(k)\psi_{j,k}(x), \tag{1}$$

where  $j_0$  is an arbitrary starting scale,  $c_{j_0}(k)$  is approximation coefficients,  $d_j(k)$  is detail coefficients.

The one-dimensional transform when extended to two-dimensional images yields a two-dimensional scaling function  $\varphi(x, y)$  and three two-dimensional wavelets  $\psi^H(x, y)$ ,  $\psi^V(x, y)$ , and  $\psi^D(x, y)$ . These three wavelets perform measure of gray-level variations for images along different directions:  $\psi^H$  measures variations in horizontal direction,  $\psi^V$  measures variations in vertical direction and  $\psi^D$  measures variations in diagonal direction.

The DWT of function  $f(x, y)$  of size  $M \times N$  is given as

$$W_\varphi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)\varphi_{j_0,m,n}(x, y), \tag{2}$$

$$W_\psi^i(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y)\psi_{j,m,n}^i(x, y), \tag{3}$$

where  $j_0$  is an arbitrary starting scale and the  $W_\psi^i(j, m, n)$  coefficients define an approximation of  $f(x, y)$  at scale  $j_0$ . The  $W_\psi^i(j_0, m, n)$  coefficients add horizontal, vertical, and diagonal details for scales  $j \geq j_0$ .

#### 3.2 Local Binary Pattern

Ojala et al. [17] originally proposed the concept of Local Binary Pattern (LBP) for texture analysis in an image. The original LBP operator is computed for a  $3 \times 3$  pixel block of an image. The pixels in this block are thresholded by the centre pixel of the block. The LBP operator takes  $3 \times 3$  surrounding of a pixel and

- generates a binary 1 if the neighbor is greater than or equal to the value of centre pixel.
- generates a binary 0 if the neighbor is less than the value of centre pixel.

The values in the thresholded neighbourhood are multiplied by the weights provided to the corresponding pixels. Some of the important properties of LBP which are useful for image retrieval are-

- It performs encoding of the relationship between gray value of centre pixel and neighbourhood pixels.
- It is an efficient feature descriptor to extract local features in an image.
- The structural arrangement of pixels in an image are efficiently represented by LBP.

### 3.3 Speeded-Up Robust Feature

The concept of SURF was first introduced by Bay et al. [18] in the year 2006. SURF is a local feature descriptor which is used for extracting interest points from an image. The concept of SURF is based on the principles of Scale Invariant Feature Transform (SIFT). However, as compared to SIFT, SURF is faster to compute and is invariant to certain image transformations such as rotation, change in scale and illumination, and small change in viewpoint. SURF descriptors are extensively used in various Computer Vision applications such as object recognition, image classification, and image reconstruction. Following are important steps in SURF algorithm-

- Interest point detection using Hessian matrix.
- Orientation assignment to interest points.
- Feature matching between images to determine whether they have same contrast.

Following are some of the important properties of SURF which are useful for image retrieval-

1. It is an efficient feature descriptor to extract interest points in an image.
2. It covers the entire image for capturing interest points.
3. The distinct features in an image which are helpful in retrieving visually similar images are efficiently captured by SURF.
4. It is faster to compute than SIFT.

### 3.4 Gray-Level Co-occurrence Matrix

Haralick et al. [19] first proposed the concept of Gray Level Co-occurrence Matrix (GLCM). GLCM provides spatial information about intensity values. GLCM helps in determining how frequently adjacent pixel pairs of specified values and in specified directions occur in an image. This helps in getting information about spatial distribution of pixel values which other features such as histogram fail to provide. It also provides information about structural arrangement of pixels. GLCM helps in getting certain textural information of any surface such as smoothness, coarseness, and roughness etc.

## 4 The Proposed Method

The proposed method consists of the following steps-

1. Computation of DWT coefficients of gray scale image.
2. Computation of LBP codes of DWT coefficients.
3. Computation of SURF descriptors of resulting LBP codes.
4. Construction of GLCM for feature vector.
5. Similarity measurement.

### 4.1 Computation of DWT Coefficients

DWT coefficients of gray scale image are computed and stored in matrix. When DWT is applied on an image it produces three detail coefficient matrices which consists of coefficients computed in three directions- horizontal detail consisting of coefficients computed in horizontal direction, vertical detail consisting of coefficients computed in vertical direction, and diagonal detail consisting of coefficients computed in diagonal direction. These coefficients are stored in three separate matrices.

### 4.2 Computation of LBP Codes

LBP codes of resulting DWT coefficients are computed and stored in three separate matrices. LBP codes of directional DWT coefficients produce texture matrix from which the interest points are extracted using SURF.

### 4.3 Computation of SURF Descriptors

Interest points from texture image is obtained by computing SURF descriptors of resulting LBP codes. The resulting SURF descriptors are scale-invariant features which are used for constructing feature vector.

### 4.4 Construction of GLCM

The fourth step is computation of GLCM for constructing feature vector. GLCM determines co-occurrence of adjacent pixel pair values. This provides information about structural arrangement of pixel values which histogram fails to provide. GLCM of resulting three SURF descriptor matrices is computed and stored in separate matrices. Similarity measurement for each of these three matrices is done separately. GLCM is used as feature vector for retrieving visually similar images. In the proposed method, for the construction of feature vector, GLCM for  $0^\circ$  angle and distance 1 has been considered and rescaled to size  $8 \times 8$ .

#### 4.5 Similarity Measurement

Similarity measurement is performed to retrieve visually similar images. Let  $(f_{Q1}, f_{Q2}, \dots, f_{Qn})$  be the set of query images and let  $(f_{DB1}, f_{DB2}, \dots, f_{DBn})$  be the set of database images. Then the similarity between query image and database image is computed using the following formula-

$$Similarity(S) = \sum_{i=1}^n \left| \frac{f_{DBi} - f_{Qi}}{1 + f_{DBi} + f_{Qi}} \right|, \quad i = 1, 2, \dots, n \quad (4)$$

#### 4.6 Advantages of the Proposed Method

Following are the advantages of the proposed method-

- The proposed method attempts to extract interest points from texture image at multiple resolutions of image. Since texture represents structural arrangement of pixels, it can be used to extract scale invariant features if an efficient texture descriptor is used. LBP is an efficient local texture descriptor which effectively represents structural arrangement of pixels.
- LBP fails to extract directional information. However, its combination with DWT proves to be efficient as DWT computes coefficients in multiple directions and its combination with LBP extracts local information in multiple directions.
- The use of SURF descriptor for extracting interest points in an image proves to be efficient as SURF is an effective scale invariant descriptor which works faster than other scale invariant descriptors such as SIFT and extracts more distinct features than SIFT.
- The combination of LBP and SURF is exploited at multiple resolutions of image. The principle advantage of using multiresolution analysis of image is that features that are left undetected at one scale get detected at another scale.
- The use of GLCM for constructing feature vector proves to be advantageous as GLCM provides information about spatial distribution of intensity values which other features such as histogram fails to provide.

### 5 Experiment and Results

For performing experiment using the proposed method, images from Corel-1 K dataset [20] have been used. Corel-1 K dataset consists of 1000 natural images divided into ten different categories namely, Africans, Beaches, Buildings, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountains, and Food. Each category consists of 100 images and each image is of size either  $256 \times 384$  or  $384 \times 256$ .

Each image of the dataset is taken as query image. If the retrieved images belong to the same category as that of the query image, the retrieval is considered to be successful. Otherwise, the retrieval fails.

## 5.1 Performance Evaluation

The performance of the proposed method has been evaluated in terms of precision and recall. Precision refers to the ratio of total number of relevant images retrieved to the total number of images retrieved. Mathematically, precision can be represented as

$$P = \frac{I_R}{T_R} \quad (5)$$

where  $I_R$  denotes total number of relevant images retrieved and  $T_R$  denotes total number of images retrieved. Recall refers to the ratio of total number of relevant images retrieved to the total number of relevant images in the database. Mathematically, recall can be represented as

$$R = \frac{I_R}{C_R} \quad (6)$$

where  $I_R$  denotes total number of relevant images retrieved and  $C_R$  denotes total number of relevant images in the database. In this experiment,  $T_R = 10$  and  $C_R = 100$ .

## 5.2 Retrieval Results

To perform the experiment, two-level DWT decomposition of 2-D gray scale image is done. DWT decomposition of image at each level produces three detail coefficient matrices- horizontal detail, vertical detail, and diagonal detail. LBP codes of each of these detail coefficient matrices are obtained separately and stored in separate matrices. SURF descriptors of resulting LBP codes are obtained to extract interest points from texture image. Finally, construction of feature vector is done by computing GLCM of SURF descriptor matrices. In the proposed method, 2-D gray scale image has been decomposed through DWT for two levels only because DWT decomposition of an image causes down sampling by a factor of 2. This reduces the size of image. For very small size image, SURF descriptors do not get computed. For computing SURF descriptors, the size of the image should be considerably large. Hence, only two levels of DWT decomposition are chosen for constructing feature vector.

When DWT is applied on 2-D gray scale image, it results in three detail coefficient matrices- horizontal detail, vertical detail, and diagonal detail. For performing similarity measurement, the feature vector for each of these matrices is constructed separately. This results in three sets of similar images. Final image set of similar images is obtained by taking union of these sets. Computation of recall is performed by counting total number of relevant image sets in the final image set. Computation of precision values is done by counting top  $n$  matches for each set. This results in final image set. The top  $n$  matches in the final set are considered for evaluating precision. Mathematically, this can be stated as follows. Let  $f_H$  be set of similar images obtained from horizontal detail feature vector,  $f_V$  be set of similar images obtained from vertical detail feature vector, and  $f_D$  be set of similar images obtained from diagonal detail feature vector. Then, the final set of similar images denoted by  $f_{RS}$  is given as

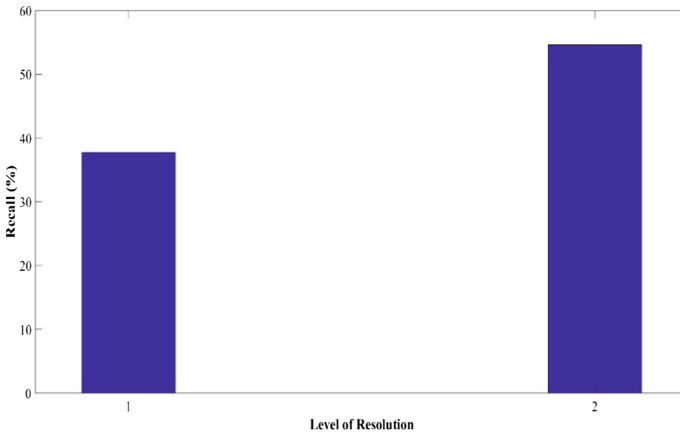
$$f_{RS} = f_H \cup f_V \cup f_D \quad (7)$$

**Table 1.** Average recall and precision values for two levels of resolution

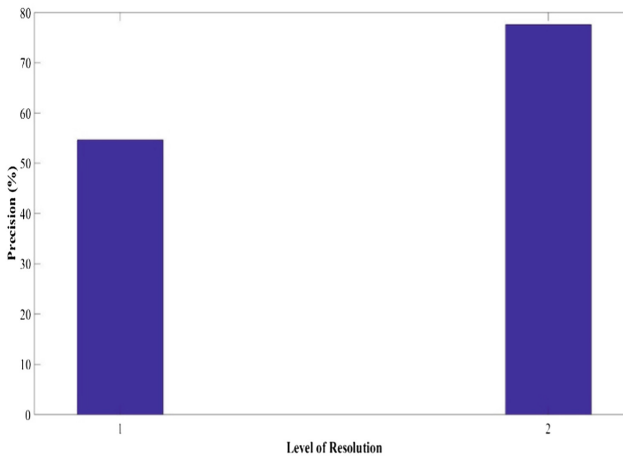
Levels of Resolution	Recall (%)	Precision (%)
Level 1	37.71	54.68
Level 2	56.59	77.61

Similarly, let  $f_H^n$  be set of top  $n$  images obtained from horizontal detail feature vector,  $f_V^n$  be set of top  $n$  images obtained from vertical detail feature vector, and  $f_D^n$  be set of top  $n$  images obtained from diagonal detail feature vector. Then the final set of top  $n$  images denoted by  $f_{PS}^n$  is given as

$$f_{PS}^n = f_H^n \cup f_V^n \cup f_D^n \quad (8)$$



(a)



(b)

**Fig. 1.** (a) Recall vs level of resolution (b) Precision vs level of resolution



The above procedure is repeated for two levels of resolution. In each level the relevant image set of the previous level is also considered and is combined with current level to produce relevant image set for that level. If the values of precision and recall are high, the retrieval is considered to be good. The average values of precision and recall obtained for two levels of resolution of image are shown in Table 1. Figure 1b shows plot between average values of precision and level of resolution and Fig. 1a shows plot between average values of recall and level of resolution. From Table 1, Fig. 1a and b, it can be observed that the average values of precision and recall increase with the level of resolution.

### 5.3 Performance Comparison

The performance of the proposed descriptor is compared with other state-of-the-art CBIR techniques such as Structure Element Descriptor (SED) [10, 13], Multi-Texton Histogram (MTH) [11, 13], Microstructure Descriptor (MSD) [12], Colour Difference Histogram (CDH) [8, 13], Xia et al. [14] and Srivastava et al. [15]. These techniques combine multiple features for constructing feature vector and produce high retrieval accuracy. However, most of these techniques exploit single resolution of image for constructing feature vector. Single resolution of image proves to be insufficient to extract varying level of details. The proposed method extracts interest points from texture feature by computing SURF descriptor of LBP codes at multiple resolutions of image. Hence it produces high retrieval accuracy than other CBIR techniques. Table 2 shows the performance comparison of the proposed method with other CBIR techniques Structure Element Descriptor (SED) [10, 13], Multi-Texton Histogram (MTH) [11, 13], Microstructure Descriptor (MSD) [12], Colour Difference Histogram (CDH) [8, 13], Xia et al. [14] and Srivastava et al. [15] in terms of precision and recall. The proposed method performs better than other state-of-the-art CBIR methods in terms of precision and recall as can be observed from Table 2.

**Table 2.** Performance comparison of the proposed method with other state-of-the-art methods in terms of Recall (%) and Precision (%)

Method	Recall (%)	Precision (%)
SED [10, 13]	9.39	61.58
CDH [8, 13]	7.89	65.75
MSD [12]	9.07	75.67
MTH [11, 13]	10.40	69.32
Xia et al. [14]	45.81	75.39
Srivastava et al. [15]	46.98	76.46
Proposed method	<b>56.59</b>	<b>77.61</b>

## 6 Conclusion

The method proposed in this paper extracts interest points from texture feature at multiple resolutions of image. The texture feature was extracted using LBP and interest points were obtained through SURF. The multiresolution decomposition of image was done through DWT. The advantages of the proposed method can be summarized as follows-

1. Texture represents structural arrangement of pixels which can be used to extract interest points. LBP proves to be an efficient texture descriptor which supplies sufficient texture information for extracting interest points.
2. SURF is an efficient feature descriptor which efficiently extracts scale invariant interest points from texture image produced by LBP.
3. The use of multiresolution processing of image for feature extraction proves to be advantageous as features that are left undetected at one level get detected at another level.

Performance of the proposed method was measured in terms of precision and recall. The experimental results demonstrate that the proposed method performs better than other state-of-the-art CBIR techniques. The proposed method can be further improved by exploiting other local patterns such as Local Derivative Pattern and using other multiresolution techniques which is going to be our future work.

## References

1. Dutta, R., Joshi, D., Li, J., Wang, J.Z.: Image retrieval: ideas influences and trends of the new age. *ACM Comput. Surv.* **40**(2), 1–60 (2008)
2. Smith, J.R., Chang, S.F.: Tools and Techniques for Color Image Retrieval. *Electronic Imaging, Science and Technology. International Society for Optics and Photonics* **2670**, 426–437 (1996)
3. Wang, X., Yu, Y., Yang, H.: An effective image retrieval scheme using color, texture and shape features. *Comput. Stand. Interfaces* **33**(1), 59–68 (2011)
4. Srivastava, P., Binh, N.T., Khare, A.: Content-based image retrieval using moments. In: *Proceedings of Context-Aware Systems and Applications, Phu Quoc, Vietnam*, pp. 228–237 (2013)
5. Srivastava, P., Binh, N.T., Khare, A.: Content-based image retrieval using moments of local ternary pattern. *Mob. Netw. Appl.* **19**, 618–625 (2014)
6. Youssef, S.M.: ICTEDCT-CBIR: integrating curvelet transform with enhanced dominant colors extraction and texture analysis for efficient content-based image retrieval. *Comput. Electr. Eng.* **38**, 1358–1376 (2012)
7. Srivastava, P., Khare, A.: Integration of wavelet transform, local binary pattern, and moments for content-based image retrieval. *J. Vis. Commun. Image Represent.* **42**(1), 78–103 (2017)
8. Liu, G.H., Yang, J.Y.: Content-based image retrieval using color difference histogram. *Pattern Recogn.* **46**(1), 188–198 (2013)
9. Zhang, M., Zhang, K., Feng, Q., Wang, J., Jun, K., Lu, Y.: A novel image retrieval method based on hybrid information descriptors. *J. Vis. Commun. Image Represent.* **25**(7), 1574–1587 (2014)
10. Wang, X., Wang, Z.: A novel method for image retrieval based on structure elements descriptor. *J. Vis. Commun. Image Represent.* **24**(1), 63–74 (2013)

11. Liu, G., Zhang, L., Hou, Y., Yang, J.: Image retrieval based on multi-texton histogram. *Pattern Recogn.* **43**(7), 2380–2389 (2008)
12. Liu, G., Li, Z., Zhang, L., Xu, Y.: Image retrieval based on microstructure descriptor. *Pattern Recogn.* **44**(9), 2123–2133 (2011)
13. Feng, L., Wu, J., Liu, S., Zhang, H.: Global correlation descriptor: a novel image representation for image retrieval. *J. Vis. Commun. Image Represent.* **33**, 104–114 (2015)
14. Xia, Yu., Wan, S., Jin, P., Yue, L.: Multi-scale local spatial binary patterns for content-based image retrieval. In: Yoshida, T., Kou, G., Skowron, A., Cao, J., Hacid, H., Zhong, N. (eds.) *AMT 2013. LNCS*, vol. 8210, pp. 423–432. Springer, Cham (2013). [https://doi.org/10.1007/978-3-319-02750-0\\_45](https://doi.org/10.1007/978-3-319-02750-0_45)
15. Srivastava, P., Khare, A.: Content-based image retrieval using multiscale local spatial binary Gaussian co-occurrence pattern. In: Hu, Y.-C., Tiwari, S., Mishra, Krishn K., Trivedi, Munesh C. (eds.) *Intelligent Communication and Computational Technologies. LNNS*, vol. 19, pp. 85–95. Springer, Singapore (2018). [https://doi.org/10.1007/978-981-10-5523-2\\_9](https://doi.org/10.1007/978-981-10-5523-2_9)
16. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*, 2nd edn. Prentice Hall Press, Upper Saddle River (2002)
17. Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. *Pattern Recogn.* **29**(1), 51–59 (1996)
18. Bay, H., Tuytelaars, T., Van Gool, L.: SURF: Speeded Up Robust Features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006. LNCS*, vol. 3951, pp. 404–417. Springer, Heidelberg (2006). [https://doi.org/10.1007/11744023\\_32](https://doi.org/10.1007/11744023_32)
19. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features of image classification. *IEEE Trans. Syst. Man Cybern. B Cybern.* **3**, 610–621 (1973)
20. <http://wang.ist.psu.edu/docs/related/>. Accessed Oct 2017