

CBIR of Brain MR Images using Histogram of Oriented Gradients and Local Binary Patterns: A Comparative Study

Athira T. R.

Computer Science and Engineering
Adi Shankara Institute of Engineering and Technology
Kalady, India

Abraham Varghese

Computer Science and Engineering
Adi Shankara Institute of Engineering and Technology
Kalady, India

Abstract— Retrieval of similar images from large dataset of brain images across patients would help the experts in the decision diagnosis process of diseases. Generally, visual features such as color, shape and texture are used for the retrieval of similar images in Content-Based Image Retrieval (CBIR) process. In this paper, Histogram of Oriented Gradients (HOGs) based feature extraction method is used to retrieve similar brain images from large image database. HOG, a shape feature extraction method is proven to be an effective descriptor for object recognition in general. It has been compared with the texture descriptor called Local binary pattern (LBP) and the results show that method outperforms the texture descriptor. The accuracy of the method is tested under different noise levels and intensity non-uniformity.

Keywords- CBIR, Histogram of Oriented Gradients, Local Binary Patterns.

I. INTRODUCTION

Interest in the potential of digital images has increased enormously over the last few years, fuelled at least in part by the rapid growth of imaging on the World-Wide Web. The enormous increase of digital information available locally or on the Internet makes it almost impossible to annotate the digital objects manually. Efficient image searching, browsing and retrieval tools are required by users from various domains, including remote sensing, fashion, crime prevention, publishing, medicine, architecture, etc. For this purpose, many general-purpose image retrieval systems have been developed.

There are two frameworks: text-based and content-based. The text-based approach can be traced back to 1970s. In such systems, the images are manually annotated by text descriptors, which are then used by a database management system (DBMS) to perform image retrieval. There are two disadvantages with this approach. The first is that a considerable level of human labour is required for manual annotation. The second is the annotation inaccuracy due to the subjectivity of human perception.

To overcome the above disadvantages in text-based retrieval system, content-based image retrieval (CBIR) was introduced in the early 1980s. Content-Based Image Retrieval (CBIR) systems aim to recognize and retrieve information

based on content of images instead of looking at metadata provided with the images. The major content in images used consists of color, shape and texture features.

A. Color [1]

In Image Retrieval, color is an important feature which is commonly applicable in stock photography (large, varied databases for being used by artists, advertisers and journalists). It is relatively robust to background noise, image size and orientation but can be altered by surface texture, lighting, shading effects and viewing conditions. These colour distortions are extremely difficult to handle for computers, because they extract the colour information from an image without context information. Additionally, in different contexts, people use various levels of colour specificity. Color Histogram is the commonly used method for color feature extraction in digital images. Color histograms have the advantages of speed and low memory space. Color histogram method is invariant to rotation but it is not invariant to scaling.

B. Texture [1]

Texture is an intuitive concept which allows us to distinguish regions of the same colour. Together with colour, texture is a powerful discriminating feature, present nearly everywhere in nature. Textures can be described according to their spatial frequency or perceptual properties (e.g., periodicity, coarseness, preferred direction, or degree of complexity). In particular, textures emphasize orientation and spatial depth between overlapping objects. In contrast to color, textures describe repeating visual patterns in homogeneous regions, i.e., texture is a region instead of a pixel property.

C. Shape [1]

Shape is an important feature for Content Based Image Retrieval, because it is invariant to translation, rotation and scaling. Unlike texture, shape is a fairly well-defined concept and there is considerable evidence that natural objects are primarily recognized by their shape. There are two types of shape representations: contour based and region based. Contour

based uses only outer boundary of the shape while the region based is uses the entire region.

In medical domain application especially for brain MR Images CBIR becomes more challenging. In this paper, we use HOG for feature extraction. HoG features are first proposed by Navneet Dalal and Bill Triggs[2]. The main idea behind the HoG descriptors is based on edge information. But HOG performs poorly when the background is cluttered with noisy edges. Local Binary Pattern[3][4] is complementary in this aspect which gives a textural information of the images. In this paper, an attempt has been made to evaluate the performance of the shape descriptor and texture descriptor with respect to the retrieval of similar brain MR images.

II. STATE OF ART

A. Shape Matching Techniques

There are mainly two types of shape matching techniques. They are:

1) Geometry Based Shape Matching

Geometry based Shape matching is a mathematical approach of matching the geometrical properties like points, lines, curve, circle etc. This geometric features are usually enough to represent any given shape of an image. So with the help of the geometric features, shape matching is performed to improve the accuracy.

a) Point Based Shape Matching [5][6][7]

Here, image matching is performed based on a single point or tuples of point matching. In a set of single point matching, it considers a point from the source image to match a point in the target images and these points could be an edge or a corner. Image matching done in this way may sometimes fail, if the images hold similar local appearances. So to improve the matching accuracy, a better solution is to include some additional constraints like distances, angles between the points, which has lead to the development of tuples of point matching techniques. Tuples of point matching may refer to point pair matching, triple point matching or higher point matching. In point pair matching, two points separated by a fixed distance and in triple point matching, a triangle of three points, having same total angle value or same distance value is used to compute image matching. Thus the accuracy is increased with the additional burden of computational cost.

b) Line, Curve, Polygon Based Matching [8][9][10]

Line, curve, polygon are the categories of geometrical properties. Lines calculated from the images can be used as a feature for image matching. Curve includes circle, ellipse, semi-circle etc which are helpful for extracting needed features for matching. Similarly different polygon shapes (like triangle) are also considered for image matching. Each method has its own strength and based on the requirement, a particular method can be chosen.

c) Blob and Graph Based Matching[11][12]

A blob is a region of a digital image in which some properties are constant or vary within a prescribed range of values. All the points in a blob can be considered in some sense to be similar to each other. Blob detection is used to obtain regions of interest for further processing which may not be possible with other edge or corner detectors. These regions could signal the presence of objects or parts of objects in the image domain which finds application in object recognition and/or object tracking.

Graphs are a flexible and powerful representation mechanism for complex scenes that have been successfully applied in computer vision, pattern recognition and related areas. When graphs are used to represent objects of a particular domain, the recognition problem turns into the task of graph matching. It can be formulated as an attributed graph matching problem, where the nodes of the graphs correspond to local features of the image and edges correspond to relational aspects between features. Graph matching then used to match the correspondence between nodes of the two graphs such that they look most similar.

2) Structural based Shape Matching [13][14]

The Structural shape matching includes features that are based on structures which may be parts of an object or symmetric resemblance of an object with various transformations. It provides a higher level of compositional shape description which could discriminate a level higher than the geometrical shape description by updating its similarity measure. Normally, human recognizes a shape not only by its local and global geometrical variations, but also by a high-level understanding of the shape structure. So to improve the retrieval accuracy of semantically related images, a combination of structural and geometrical features have been adopted.

B. Texture Matching Techniques

There are many methods for texture feature extraction. Based on the domain from which the texture feature is extracted, these methods are broadly classified in to two categories. They are Spacial texture feature extraction method and Spectral texture feature extraction methods.

In spatial approach, texture features are extracted by computing the pixel statistics or finding the local pixel structures in original image domain. The spatial texture feature extraction techniques can be further classified as structural, statistical and model based. In spectral texture feature extraction techniques, an image is transformed into frequency domain and then feature is calculated from the transformed image.

1) GCLM

Haralick [15] suggested the use of grey level co-occurrence matrices (GLCM) to extract second order statistics from an image. Statistical features of grey levels were one of the earliest methods used to classify textures.

GLCMs have been used very successfully for texture classification in evaluations. Haralick defined the GLCM as a matrix of frequencies at which two pixels, separated by a certain vector, occur in the image. The distribution in the matrix will depend on the angular and distance relationship between pixels. Varying the vector used allows the capturing of different texture characteristics. Once the GLCM has been created, various features can be computed from it. These have been classified into four groups: visual texture characteristics, statistics, information theory and information measures of correlation.

2) *Tamura*

Tamura et al. [16] have proposed a texture features corresponding to human visual perception. It is very useful for optimum feature selection and texture analyzer design. Here they approximated in computational form six basic textural features, namely, coarseness, contrast, directionality, line-likeness, regularity, and roughness. Computational measures have been developed and improved in order that they may correspond to the psychological measurements. With respect to coarseness, contrast, and directionality, we have attained very successful results. These three features are so significant in global descriptions of textures that they can be expected to be useful separately in cases where the texture differs only in one of them or in combinations for image classification and segmentation problems. Especially, coarseness is a highly essential factor in texture. In order to improve the other features, its results should be utilized.

3) *FD*

The FD method [17] is based on the theory of fractal geometry which characterises shapes or patterns of self-similarity. It attempts to find the smallest structure which replicates the whole pattern. In practice, FD method models a grey level image as a 3D terrain surface, and a differential box counting is conducted under the surface to measure how rough the surface is. As the logarithmic number of boxes and the logarithmic box size have a linear relationship, FD can be estimated from the least-square fit of the two variables. As FD only models the roughness feature, other features like directionality and contrast are missed from FD. Therefore, in Chaudhuri and Sarkar, six FDs have to be computed from a number of modified images. Furthermore, FD is not rotation invariant.

3) *Gabor*

One of the most popular signal processing based approaches for texture feature extraction has been the use of Gabor filters [18]. These enable filtering in the frequency and spatial domain. It has been proposed that Gabor filters can be used to model the responses of the human visual system. Turner first implemented this by using a bank of Gabor filters to analyse texture. A bank of filters at different scales and orientations allows multichannel filtering of an image to extract

frequency and orientation information. This can then be used to decompose the image into texture features.

4) *FT/DCT*

FT method is used in Hervé and Boujemaa [19]. Two histograms are computed from the FT spectra, one with the circular partition and the other with the wedged partition. Although FT is a powerful image analysis tool, it can only capture global features which are not sufficient for texture analysis. Ngo et al. arrange the low order DCT [20] coefficients into Mandala space which represents the partial derivatives of the original image. Texture features are computed from the variance of each of the derivative images. DCT method is also used by Lu et al., but they use mean and standard deviation as the texture features instead. Similar to FT, DCT method can only capture the global features.

5) *Wavelet*

In wavelet method [21], images are decomposed into different frequency components using filters at different scales. Texture features are then extracted from each of the frequency components. Both the pyramid-structured wavelet transform (PWT) and tree-structured wavelet transform (TWT) can be used. Although the TWT method gives a slight better result, the difference is not significant. The wavelet method shows significant advantage over FT method as it captures local spectral features at multiple resolutions. However, the problem with wavelet method is that wavelets are usually sensitive to point singularities instead of edge singularities which are crucial for extracting texture features. The other problem with wavelet method is that it does not adapt to directional textures.

6) *Curvelet*

Recently, a new wavelet like method called curvelet [10] has been introduced to overcome the limitation of wavelet. Curvelets are a special set of wavelets which are designed to adapt to curved edges in images. With curvelet, image edge information is captured at different orientations and scales. It appears to take the advantages of both wavelet and Gabor filters, and has been successfully used in image denoising and enhancement. There are also a few early applications in image classification and retrieval. However, none of them has considered rotation invariance and region based image retrieval. Joutel et al. have created an assistance tool for the identification of ancient handwritten manuscripts using ridgelet transform. However, only curvature and orientation features are extracted which are not sufficient for image retrieval. Arivazhagan et al. and Shekhar et al. used curvelet features for color image classification and retrieval. However, color images are not homogenous and they cannot be classified without segmentation. Semler and Dettori applied curvelet in medical domain and used curvelet texture features on cropped CT images for organ classification.

III. METHODOLOGY

Here we have done a comparison of two methods HOG and LBP.

A. Histogram of Oriented Gradient (HoG)

Histogram of oriented gradient (HoG) feature descriptors was proposed by Dalal and Triggs [2]. The idea behind HoG descriptor is based on edge information. The HoG descriptor technique counts occurrences of gradient orientation in localized portions of an image - detection window, or region of interest (ROI).

The first step is divide the image into small connected regions called cells, and for each cell computes a histogram of gradient directions or edge orientations for the pixels within the cell. Then discretize each cell into angular bins according to the gradient orientation. Gradient value is computed by apply the 1dimensional centered point discrete derivative mask in both the horizontal and vertical directions. Each cell's pixel contributes weighted gradient to its corresponding angular bin. Groups of adjacent cells are considered as spatial regions called blocks. The grouping of cells into a block is the basis for grouping and normalization of histograms. Normalized group of histograms represents the block histogram. The set of these block histograms represents the descriptor. An example input image and its corresponding Histogram of Oriented Gradients shown in figure 2.

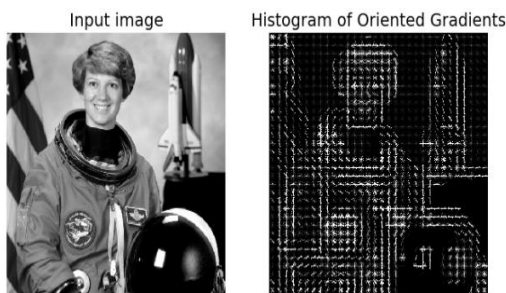


Figure 2: Input image and Histogram of Oriented Gradients

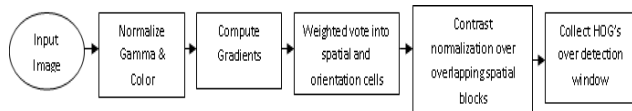


Figure 1: Block diagram of Histogram of Oriented Gradients

B. Local Binary Patterns

Local binary patterns were introduced by Ojala et al [3] as a fine scale texture descriptor. Local Binary Pattern (LBP) [4] is a simple yet very efficient texture operator which labels the pixels of an image by thresholding the neighborhood of each pixel and considers the result as a binary number. Due to its discriminative power and computational simplicity, LBP texture operator has become a popular approach in various applications. It can be seen as a unifying approach to the

traditionally divergent statistical and structural models of texture analysis. Perhaps the most important property of the LBP operator in real-world applications is its robustness to monotonic gray-scale changes caused, for example, by illumination variations. Another important property is its computational simplicity, which makes it possible to analyze images in challenging real-time settings.

LBP description of a pixel is created by thresholding. The values of the 3X3 neighborhood of the pixel against the central pixel and interpreting the result as a binary number. The process is illustrated in figure 4.

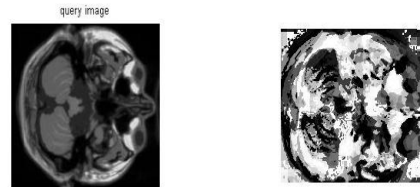


Figure 3: Query image and its corresponding LBP pattern

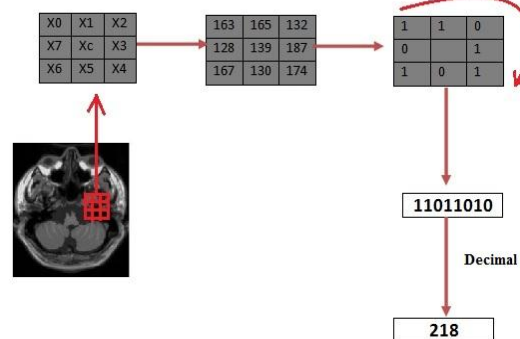


Figure 4: The standard LBP calculation principle

C. Similarity Measure

For comparing the similarity of two metric we use Euclidean distance. In Cartesian coordinates, if $\mathbf{p} = (p_1, p_2, \dots, p_n)$ and $\mathbf{q} = (q_1, q_2, \dots, q_n)$ are two points in Euclidean n -space, then the distance (d) from \mathbf{p} to \mathbf{q} , or from \mathbf{q} to \mathbf{p} is given by the Pythagorean formula:

$$d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (1)$$

IV. PERFORMANCE EVALUATION

A. Average Rank

Average rank is calculated using the images retrieved based on a set of random number of query images. In perfect case its value is 1 i.e., relevant images are in a succeeding order without the presence of irrelevant images in between them. Average rank is calculated using the formula:

$$Average\ rank = \frac{1}{N_R} \left(\sum_{i=1}^n R_i - \frac{N_R(N_R-1)}{2} \right) \quad (2)$$

V. RESULT

In our experiment, different types of axial view of brain image datasets are used, which is downloaded from publicly available BrainWeb dataset [22]. The data set consists of 150 T1 weighted images with 1 mm slice thickness, different noise levels(0%, 1%, 3%, 7%, 9%) and different intensity non-uniformity (0%, 20%, 40%).

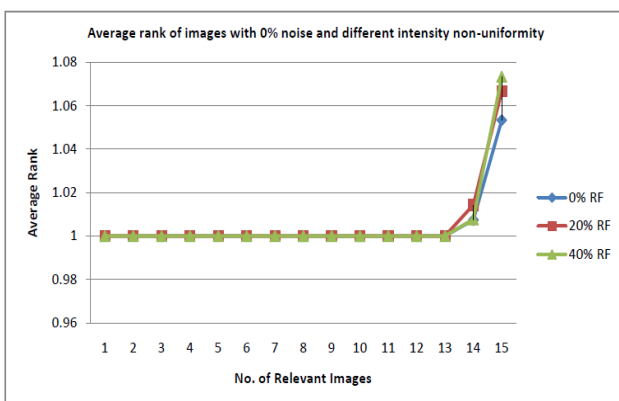
In figure 5, the average rank versus number of images retrieved in different noise levels are shown. Figure 5(a) shows average rank in retrieving first 15 relevant images with 0% noise and 0%, 20% and 40% intensity non-uniformity. Figure 5(b) shows average rank of images with 1% noise and 0%, 20% and 40% intensity non-uniformity. Figure 5(c) shows average rank of images with 3% noise and 0%, 20% and 40% intensity non-uniformity. Figure 5(d) shows average rank of images with 5% noise and 0%, 20% and 40% intensity non-uniformity.

Figure 5(e) shows average rank of images with 7% noise and 0%, 20% and 40% intensity non-uniformity And Figure 5(f) shows average rank of images with 9% noise and 0%, 20% and 40% intensity non-uniformity. By analyzing these results there can be concluded that as intensity non-uniformity increases, performance decreases slightly.

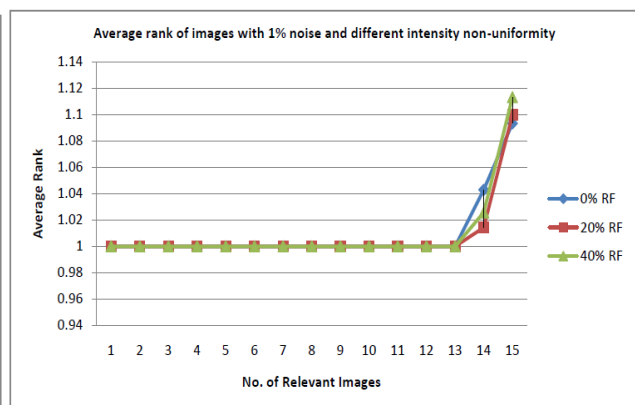
In Table 1 indicates the comparison of HOG and LBP for the retrieval of 15 relevant images. From this table we can analyze that the performance of HOG is much better than the texture feature extraction method LBP. Two datasets are used for the comparison. They are “t1_icbm_normal_1mm_pn0_rf0”, (i.e., t1 modality normal dataset with 1mm thickness, noise 0% and RF 0%) and “t1_icbm_normal_1mm_pn9_rf40”, (i.e., t1 modality normal dataset with 1mm thickness, noise 9% and RF 40%).

TABLE I. COMPARISON OF HOG AND LBP FOR THE RETRIEVAL OF 15 RELEVANT IMAGES

Noise	RF	Method	Number of relevant images retrieved														
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0%	0%	HOG	1	1	1	1	1	1	1	1	1	1	1	1	1	1.007	1.053
		LBP	1	1	1	1	1.08	1.15	1.28	1.47	1.65	1.91	2.17	2.43	2.75	3.55	4
9%	40%	HOG	1	1	1	1	1	1	1	1	1	1	1	1	1.0077	1.0429	1.1267
		LBP	1	1.2	1.46	2.05	2.84	3.58	4.38	5.26	6.03	6.88	7.97	9.46	10.96	13.47	16.72



(a)



(b)

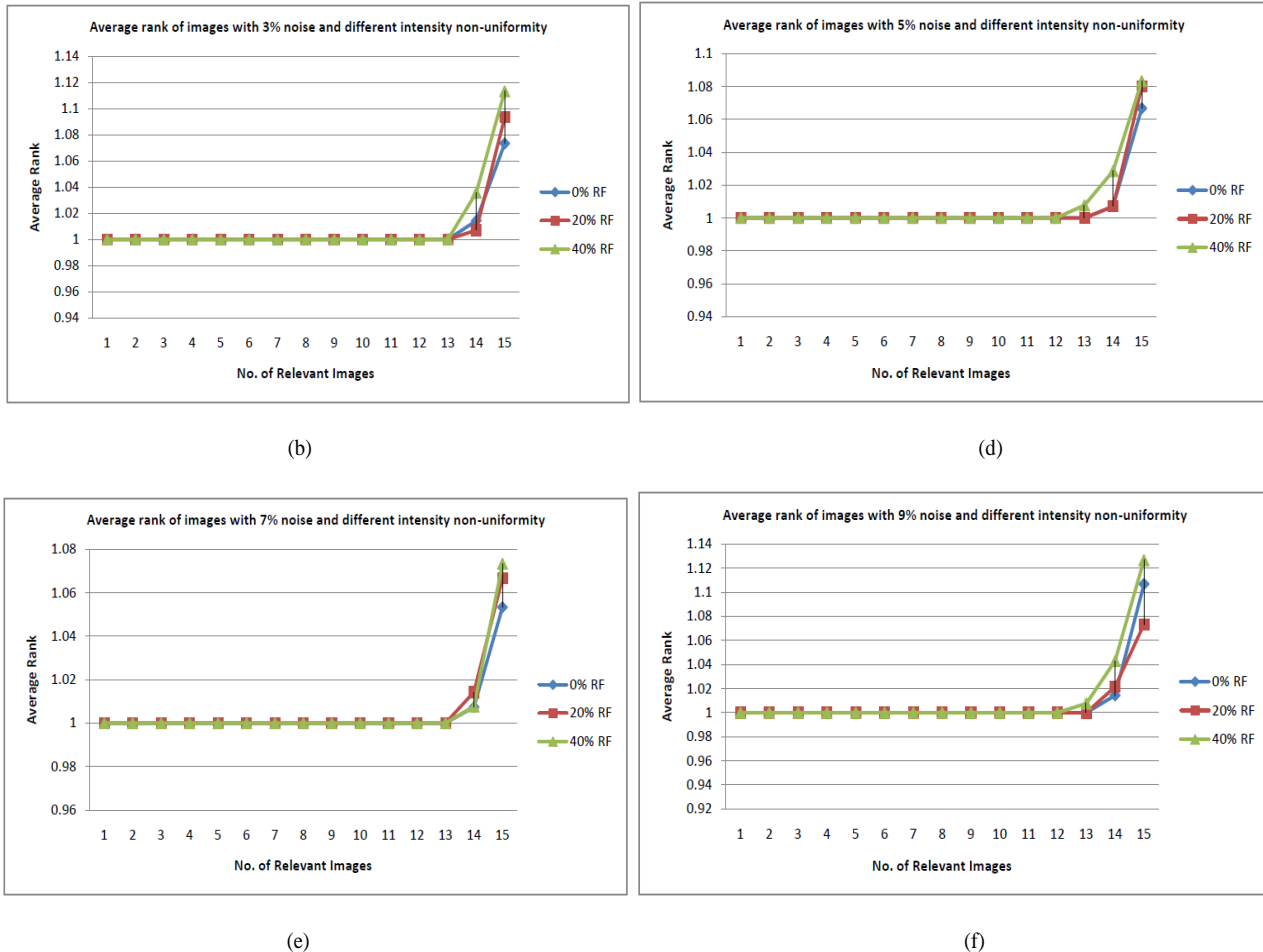


Figure 4 :Average Rank versus the number of relevant images retrieved with different noise levels and different intensity non-uniformity("RF").

VI. CONCLUSION

Medical imaging technologies, such as Magnetic Resonance Imaging (MRI), provide important insights for understanding the disease pathology and are essential for biomedical research and health care. However, the increasingly large medical collections pose great challenges in medical data management and retrieval. In this paper, a comparison of HOG and LBP feature is performed. Histogram of Orient Gradients (HOGs) is a shape feature extraction method and Local Binary pattern is a texture feature extraction method. Histograms of Oriented Gradients (HOGs) have proven to be an effective descriptor for object recognition in general. Local binary pattern (LBP) has been recently proved useful in describing medical images, which has a low computational complexity and a low sensitivity to changes in illumination. But the result shows that in this application LBP outperforms than HOG. So we can conclude that HOG is much accurate and gives better performance in retrieval of brain MR images.

REFERENCES

- [1] Sameer Antani, Rangachar Kasturi, Ramesh Jain, "A surveyon the use of pattern recognition methods for abstraction, indexing and retrieval of images and video", *Pattern Recognition*, Vol. 35, pp. 945–965, 2002.
- [2] Dalal N. and Triggs B., "Histograms of oriented gradients for human detection," *IEEE conference on Computer Vision and Pattern Recognition*, pp. 886–893, 2005.
- [3] Ojala T., Pietikäinen M., Mäenpää., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns", *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002).
- [4] Ojala T., Valkealahti K., Oja, E., Matti Pietikainen, "Texture discrimination with multidimensional distributions of signed gray-level differences", *Pattern Recognit.* **34**(3), 727–739 (2001).
- [5] Jianning Liang, Zhenmei Liao, SuYang, YuanyuanWang," Image matching based on orientation–magnitude histograms and global consistency", *Pattern Recognition* 45, pp. 3825–3833, 2012.
- [6] Glauco V. Pedrosa, Marcos A. Batista, Celia A. Z. Barcelos," Image feature descriptor based on shape salience points", *Neurocomputing* 120, pp. 156–163, 2013.

- [7] Zhi-Quan Cheng, Yin Chen, Ralph R. Martin, Yu-Kun Lai, and Aiping Wang, "SuperMatching: Feature Matching Using Supersymmetric Geometric Constraints", IEEE Transactions on Visualization and Computer graphics, Vol. 19, No. 11, pp. 1885-1894, November 2013.
- [8] Huijing Fu, Zheng Tian, Maohua Ran, Ming Fan, "Novel affine-invariant curve descriptor for curve matching and occluded object recognition" IET Computer Vision, Vol. 7, No. 4, pp. 279-292, 2013.
- [9] Jianfang Dou, Jianxun Li, "Image matching based local Delaunay triangulation and affine invariant geometric constraint", Optik 125, pp. 526- 531, 2014.
- [10] Min Chen, Zhenfeng Shao, Chong Liua, Jun Liu," Scale and rotation robust line-based matching for high resolution images", Optik 124, pp. 5318- 5322, 2013.
- [11] Chunhui Cui and King Ngi Ngan," Global Propagation of Affine Invariant Features for Robust Matching", IEEE Transactions on Image Processing, Vol. 22, No. 7, pp. 2876-2888, July 2013.
- [12] Amir Egozi, Yosi Keller and Hugo Guterman," Improving Shape Retrieval by Spectral Matching and Meta Similarity", IEEE Transactions on Image Processing, Vol. 19, No. 5, pp. 1319-1327, May 2010.
- [13] Yu Shi, Guoyou Wang, Ran Wang, Anna Zhu," Contour descriptor based on space symmetry and its matching technique", Institute for Pattern Recognition and Artificial Intelligence, Optik 124, pp. 6149- 6153, 2013.
- [14] Seungkyu Lee," Symmetry-driven shape description for image retrieval", Image and Vision Computing 31, pp. 357-363, 2013.
- [15] Haralick, R.: Statistical and structural approaches to texture. Proceedings of the IEEE 67, pp. 786-804, 1979
- [16] Tamura, H., Mori, S., Yamawaki, T.: Textural features corresponding to visual perception. IEEE Trans on Systems, Man and Cybernetics 8 (1978) 460-472.
- [17] Turner, M.: Texture discrimination by Gabor functions. Biological Cybernetics 55 (1986) 71-82.
- [18] Hervé, N., & Boujemaa, N. Image annotation: which approach for realistic databases? In Proc. of the 6th ACM international conf. on image and video retrieval , Amsterdam, Netherlands (pp. 70- 177), 2007.
- [19] Manjunath, B. S., & Ma, W. Y. Texture features for browsing and retrieval of large image data. IEEE Transactions on Pattern Analysis and Machine Intelligence , 18 (8), 837-842, 1996.
- [20] Lu Z., Li S, Burkhardt H., A content-based image retrieval scheme in jpeg compressed domain, International Journal of Innovative Computing, Information and Control 2(4), pp. 831-839, 2006.
- [21] Do, M. N., & Vetterli, M. . The finite ridgelet transform for image representation. IEEE Transactions on Image Processing , 12 (1), 16-28, 2003.
- [22] <http://brainweb.bic.mni.mcgill.ca/brainweb/>