# Local Primitive Histograms for Patent Binary Image Retrieval

Naeem A. Bhatti
*Institute of Computer Aided Automation*
*Vienna University of Technology*
*Favoritenstr. 9/183-2, 1040 Vienna, Austria*
*bhatti@caa.tuwien.ac.at*

Allan Hanbury
*Institute of Software Technology and Interactive Systems*
*Vienna University of Technology*
*Favoritenstr. 9/188, 1040 Vienna, Austria*
*hanbury@ifs.tuwien.ac.at*

***Abstract*—Local primitives are useful in the analysis, recognition and retrieval of document and patent images. In this paper, local primitives are classified in 4 and 8-directional spaces at optimally detected junction and end points by using a distance based approach. Local primitives are quantized by using a variant of Local Binary Patterns. Spatial relationships between local primitives are established by using a morphology based approach. Binary images are described by Local Primitive Histograms of the classified local primitives in 4 and 8-directional spaces capturing their occurrences and pair-wise co-occurrences. Performance evaluation of the proposed Local Primitive Histograms for patent binary image retrieval shows improvement in comparison with histograms obtained by SIFT description of Local Primitives.**

***Keywords*-local primitives; spatial relationships; granulometric curve; binary image retrieval;**

## I. Introduction

To deal with prior art and the infringement issues in patents, the retrieval of all the patent images similar to a query image is an important and challenging task. Patent images such as technical drawings, diagrams, flowcharts etc. are binary. Different geometric shapes and their spatial arrangements form the contents of these images. The local pattern formed by the composition of the intersecting or crossing lines at a crossing point (junction point) is called a local primitive. Local primitives are useful in the analysis, recognition and retrieval of document and patent images.

A number of approaches to capture the content in binary images [5], [11], [14], [1], [12] have been developed. Working on patent image retrieval, Huet et al. [5] employed the relational skeletons approach to capture the geometric structure of the image. The approach uses Voronoi Skeletonization to extract the line patterns and a polygonization technique to create line segments from the extracted line patterns. The angle and position of a line in relation to another line are used as relational features. Targeting general purpose CBIR, Mahmoudi et al. [8] capture the local correlation of neighboring edges in a 2-dimensional histogram called the edge orientation autocorrelogram (EOAC) where an element $(j, k)$ indicates the number of edges with the $j^{th}$ orientation lying at distance $k$ apart in pixels. EOAC are employed as the basic technique in [11] using Canny edge detection. The approach quantizes the gradient of the filtered edges to 36 bins with 5 degrees each. Yang et al. [13] exploit the invariant characteristics of the images at the local and global levels by partitioning the image in an adaptive hierarchical way and computing the geometric centroids vector of an image at each partition level. Vrochidis et al. [12] proposed the adaptive hierarchical density histogram (AHDH) which calculates the distribution of black pixels on a white plane. The density estimation is done at local as well as global level. At local level the image plane is divided into four regions based on the centroid of the image and the local density estimation is performed in each region. Sidiropoulos et al. [9] extended the work presented in [12] by introducing quantized relative density features. They obtain the hierarchical partitioning of an image following the technique presented in [13]. At the sub-regions, the approach computes the relative density feature (the number of black pixels of a sub-region compared to the ratio of sub-region area over the region area) and quantizes it to a defined lexicon of 16 combinations (which are obtained based on the 4 full or empty sub-regions of a region in respect to the number of black pixels). The relative density features are computed at the lower partition levels where the information to capture decreases significantly. The density features (the number of black pixels in a sub-region divided by the total number of black pixels in all regions) are computed at the higher partition levels which are less than an experimentally defined threshold for partition levels. AHDH are obtained by concatenating the density features and the quantized relative density features. The experiments for binary image retrieval reveals the superiority of the AHDH in comparison to EOAC [8] and geometric centroids [13].

In the above mentioned approaches, the local as well as the global characteristics of images exploited in [13], [9] suffer from rotation variance and failed to retrieve similar images with different orientations. The local information exploited by EOAC and edge based techniques suffers from its inherent dependence on drawing style and fails to capture accurate edge directions due to the very thin or very thick line drawings and filled-in figures [6], [9].

In view of the inherent deficiencies of the local and global information mentioned in the above approaches, the proposed approach in this paper focuses on the local information and attempts to overcome its sensitivity to the varying thickness of the lines in line drawings. A new representation based on "Local Primitives" is proposed. The main contributions are: 1) Local Primitive Histograms (LPHs) for the composition of Local Primitives in 4-directional and 8-directional spaces, 2) taking into account the spatial relationships between Local Primitives, construction of LPHs based on their occurrence, co-occurrence and combined occurrence and co-occurrence frequencies in each image, 3) performance evaluation of proposed LPHs and SIFT-based LPHs for similar image retrieval in a database of 2000 patent images. It is found that the proposed Local Primitive Histograms are robust in retrieving similar images with local variations and rotations in their drawing styles.

The proposed approach is explained in Section II. The experimental results for similar image retrieval and a performance comparison of Local Primitive Histograms in 4 and 8-directional spaces is presented in Section III and a conclusion is drawn in Section IV.

## II. The Proposed Approach

In this section, the steps involved in the proposed approach are explained.

**a. Detection of Junction Points and End Points.** The detection of junction points (JPs) and end points (EPs) in the homotopic skeleton using template based matching [10] gives rise to false detections due to spurious skeletal lines introduced during the thinning process of the image. A morphological spurring operation [10] removes parasitic skeletal lines in the skeleton image by the number of operations it is performed. The false detections can be eradicated by taking an intersection of the skeleton image containing detections with a skeleton image obtained by performing iterative morphological spurring operations. To determine an optimum number of iterations for the morphological spurring operation, the proposed approach takes into account the average thickness of lines $Th_{lines}$ obtained by taking a weighted average of the pattern spectrum [10] $PS_k\left(I\right)$ obtained from the granulometry of the original image, given as:

$$Th_{lines} = \frac{\sum_k PS_k\left(I\right) * k}{\sum_k PS_k\left(I\right)} \tag{1}$$

where $PS_k\left(I\right)$ is the value of bin $k$ of the $PS_k\left(I\right)$ and is obtained by taking the discrete derivative of the granulometric curve of the original image. To remove the noisy detections, an optimum number of spurring iterations for EPs are computed as $GI_{EP} = \lceil Th_{lines}/2 \rceil$ and for JPs as $GI_{JP} = \lfloor Th_{lines} \rfloor$ (the values for $Th_{lines}$ obtained are floating point) [3].

**b. Detection and Classification of Local Primitives.** The local primitives found at the optimally detected JPs and EPs are called Junction Point Primitives (JPPs) and End Point Primitives (EPPs) respectively. All the local primitives (EPPs and JPPs) are classified into primitive classes using a distance based approach [2]. The classification of primitives is performed by taking into account their composition in 4-directional and 8-directional spaces as shown in Figure 1 (a) and (b) respectively. Lines in regions around junction and end points are quantized to reference directions of 4 or 8-directional space. The quantized lines are represented in a binary vector similar to the approach used in the Local Binary Pattern (LBP) [4]. Regarding the 8-directional composition, there are 8 end point primitive classes and 244 junction point primitive classes (not 248 because the junction point primitives composed of lines having an angular difference of exactly $180^o$ do not fulfill the definition of a junction point primitive). Similarly, to quantize into 4 directions instead of 8 results in only 14 classes (4 for EPPs and 10 for JPPs). As a result of this process, each EPP and JPP has a class number associated with it.

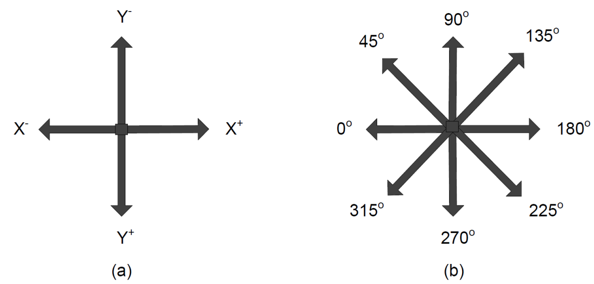
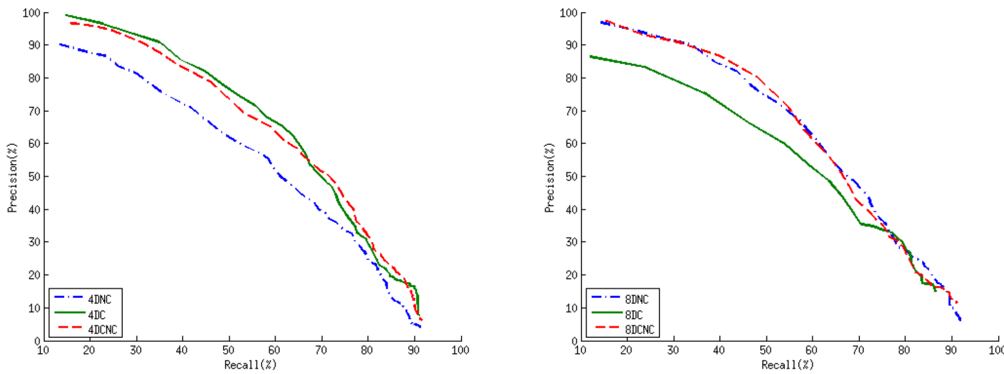



Figure 1. (a) X and Y space and reference line segments in 4-directional space (b) Reference line segments in 8-directional space.

**c. Morphology based Spatial Relationships between Local Primitives.** To establish the spatial relationships between the local primitives, a morphology based approach is adopted. By taking a marker image with unique gray values at the local primitives and the skeleton of the original image as the mask image, the proposed approach performs successive gray scale geodesic dilations [10]. The geodesic paths along the skeleton between the local primitives are traversed and their points of contact are protected by updating the mask image after each geodesic dilation iteration. By scanning the final marker image for the contact points of the traversed geodesic paths, connections between the local primitives are established which allows to capture the pairwise co-occurrence of local primitives. Two primitives are considered to co-occur if they are connected by a line.

**d. Local Primitive Histograms.** Each possible configuration of lines in the 4 or 8 directional space results in a number by encoded as a binary sequence in a clockwise order, as done for the LBP. A binary image is represented by a histogram of the occurrence frequency of local primitives "4DNC" (where "D" is for directional space and "NC" for non co-occurrence) of length $14$ in the 4-directional space and "8DNC" of length $252$ in the 8-directional space. The proposed approach also captures the co-occurrence of local primitives by constructing a histogram based on co-




(a) (b)

Figure 2. (a) Precision - Recall Curves for 4 directional LPHs (b) Precision - Recall Curves for 8-directional LPHs.

occurrence frequency "4DC", and "8DC" in the respective 4 and 8-directional spaces. For $n$ distinct local primitives, there are $m = n.(n+1)/2$ distinct pairs of local primitives. Each bin of the co-occurrence histogram of length $m$ represents the occurrence frequency of a pair of local primitives. A fusion of occurrence and co-occurrence histograms is obtained by their concatenation as occurrence co-occurrence histogram "4DCNC" in the 4-directional space and "8DCNC" in the 8-directional space. In addition to 4 and 8-directional classification of LPs, we perform SIFT [7] description of the local primitives and construct a vocabulary of 252 visual words (same as the 8-directional space) by using k-means clustering. Binary images are represented by "LPSiftNC", "LPSiftC" and "LPSiftCNC" histograms for occurrence, co-occurrence and combined occurrence co-occurrence of LPs respectively. The significance of all these developed Local Primitive Histograms is explored in the retrieval of similar images in section III.

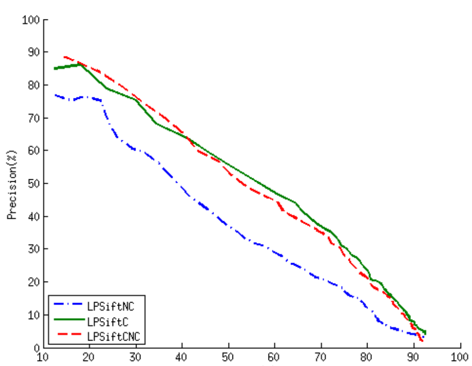



Figure 3. Precision - Recall Curves for SIFT-based LPHs

## III. Results and Discussion

For an evaluation of the proposed approach, similar image retrieval experiments for 2000 binary patent images[1] are conducted. A query base of 96 images is used where each image has 25 similar images at most. L1 distance is used as a similarity measure. To calculate the precision and recall values, the number of retrieved images are decided by using a distance interval from $D_{min} - D_{max}$. Where $D_{max} = D_{min} \times T_d$ and $D_{min}$ is the distance value of the closest image for the L1 distance obtained for a query image and $T_d = 1.1, 1.2, 1.3...10$. The precision and recall values are determined by using the number of retrieved results obtained within the distance interval. The retrieval results for the proposed Local Primitive Histograms (LPHs) based on the classification of local primitives in 4, 8-directional spaces and SIFT description are shown in the form of precision-recall (PR) curves in Figures 2 and 3 respectively. The retrieval results show that for the 4 and 8-directional local primitives the PR curves for "4DC" and "8DNC" are better respectively. In the 4-directional case, histogram "4DNC" of length 14 is too short to sufficiently capture the content of binary images. However the co-occurrence histogram "4DC" with length 105 of LPs pairs better captures the content of images exhibited by its better performance. Regarding the 8-directional space, the high number (31878) of the local primitive pairs in "8DC" makes it sparse as compared to the "8DNC" of length 252 affecting the retrieval performance which is clear from their respective PR curves. However these PR curves are better than all the PR curves of SIFT-based LPHs. The concatenation histograms "4DCNC", "8DCNC" and "LPSiftCNC" do not add a noticeable positive effect in the retrieval performance.

Two visual examples for query images of a technical graph and a technical diagram and their first retrieved results are shown in Figures 4 and 5 respectively. Figure 6 shows a retrieval example where LPHs fail to retrieve correct results. In this example the local information in the query image and retrieved results is comparable, however the images differ in global appearance. It shows that an inclusion of the global information with LPHs can potentially deal with this restraining factor of LPHs.

## IV. Conclusion and Future Work

In this paper, to capture the content of binary images at local level, a new representation based on Local Primitives

[1] http://mklab.iti.gr/content/patent-database

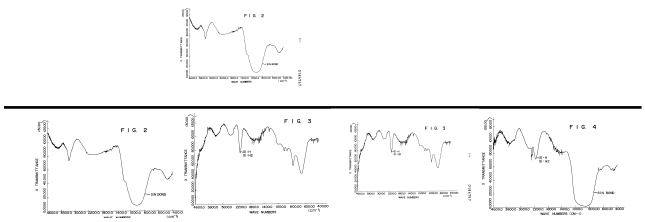


Figure 4. A technical graph query image and the first retrieved results.

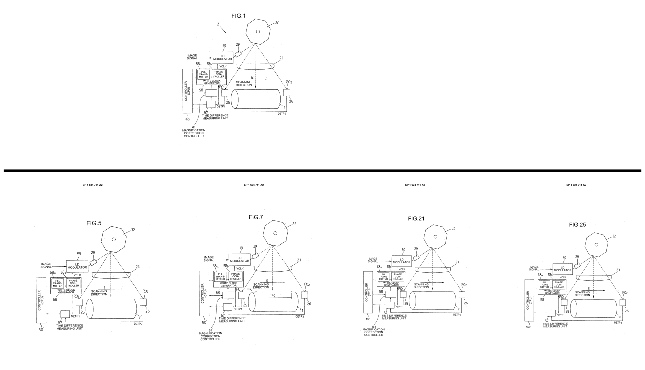


Figure 5. A technical diagram query image and the first retrieved results.

is presented. The granulometric information of the image is used to detect optimal junction and end points. To classify the local primitives found at the detected points in 4 and 8-directional spaces a distance based approach is used. Spatial relationships between local primitives are established by using a morphology based approach. Local Primitive Histograms are developed based on the occurrence and co-occurrence frequency of 4, 8-directional classification of Local Primitives and SIFT-based description of local primitives. Evaluation of the performance of the proposed histograms for binary image retrieval is performed and the comparison shows that histograms obtained by distance based classification of Local Primitives perform better than SIFT based Local Primitive Histograms where the histograms for both the cases have equal length. A future task is to incorporate the global information of the binary patent images which could potentially enhance its performance to deal with global appearance of images.

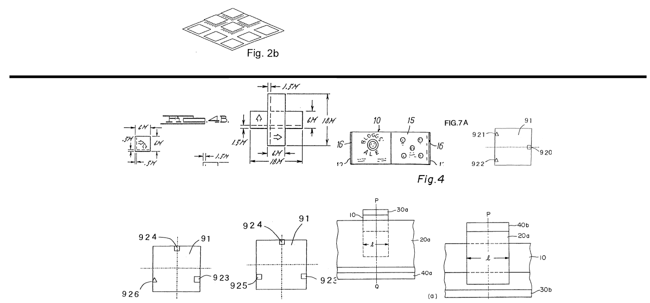


Figure 6. An example of patent image for which LPHs fail to retrieve quality results.

## REFERENCES


[1] Naif Alajlan, Mohamed S. Kamel, and George H. Freeman. Geometry-based image retrieval in binary image databases. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30:1003–1013, 2008.

[2] Naeem A. Bhatti and Allan Hanbury. Detection and classification of local primitives in line drawings. In *35th Annual Workshop of the Austrian Association for Pattern Recognition*, 2011.

[3] Naeem A. Bhatti and Allan Hanbury. Granulometry based detection of junction and end points in patent drawings. In *Proceedings of the 7th International Symposium on Image and Signal Processing and Analysis (ISPA)*, 2011.

[4] M. Heikkila, M. Pietikainen, and C. Schmid. Description of interest regions with local binary patterns. *Pattern Recognition*, 42(3):425–436, March 2009.

[5] Benoit Huet, Gennarino Guarascio, Nicholas J. Kern, and Bernard Mérialdo. Relational skeletons for retrieval in patent drawings. In *ICIP (2)*, pages 737–740, 2001.

[6] Anil Jain and Aditya Vailaya. Shape-based retrieval: A case study with trademark image databases. *Pattern Recognition*, 31:1369–1390, 1998.

[7] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110, 2004.

[8] Fariborz Mahmoudi, Jamshid Shanbehzadeh, Amir-Masoud Eftekhari-Moghadam, and Hamid Soltanian-Zadeh. Image retrieval based on shape similarity by edge orientation autocorrelogram. *Pattern Recognition*, pages 1725–1736, 2003.

[9] P. Sidiropoulos, S. Vrochidis, and I. Kompatsiaris. Adaptive hierarchical density histogram for complex binary image retrieval. In *Content-Based Multimedia Indexing (CBMI), 2010 International Workshop on*, page 16. IEEE, 2010.

[10] P. Soille. *Morphological Image Analysis: Principles and Applications*. Springer, 2nd edition, 2003.

[11] Avinash Tiwari and Veena Bansal. Patseek: Content based image retrieval system for patent database. In *ICEB*, pages 1167–1171, 2004.

[12] Stefanos Vrochidis, Symeon Papadopoulos, Anastasia Moumtzidou, Panagiotis Sidiropoulos, Emanuelle Pianta, and Ioannis Kompatsiaris. Towards content-based patent image retrieval: A framework perspective. *World Patent Information*, 32(2):94–106, 2010.

[13] Mai Yang, Guoping Qiu, Jiwu Huang, and Dave Elliman. Near-duplicate image recognition and content-based image retrieval using adaptive hierarchical geometric centroids. In *ICPR*, pages 958–961, 2006.

[14] Zeng Zhiyuan, Zhao Juan, and Xu Bin. An outward-appearance patent-image retrieval approach based on the contour-description matrix. In *FCST*, pages 86–89, Washington, DC, USA, 2007. IEEE Computer Society.