A Fast Algorithm for Retrieval of Images in a Library of Masonmarks

V. Kiiko[1], V. Matsello[1], H. Masuch[2], G. Stanke[3]

[1]Int' l Research and Training Centre of Information Technologies and Systems of National Academy of Sciences of Ukraine, Prospect Academica Glushkova, 40,252022, Kiev; Tel.: +380 (0)44 266-3018, Fax: +380 (0)44 266-1570, e-mail: kiyko@image.kiev.ua

[2]Private, Husarenstraße 2, 30163 Hannover; Tel.: +49 (0)511 669060, Fax: +49 (0)511 3944859

[3]Gesellschaft zur Förderung angewandter Informatik (GFaI), Rudower Chaussee 5, D12484 Berlin; Tel.: +49 (0)30 6392-1601, Fax: +49 (0)30 6392-1602, e-mail: stanke@gfai.de

1. Introduction

Developing of systems for the archival, retrieval and distribution of electronic documents containing or consisting of images has becoming of great importance in many fields of human activities, and development of effective retrieval procedure is usually a key problem of such systems. If there exists already a symbolic indexing or describing character strings provided by OCR-systems are available, common data base techniques can be used. But, for documents where the characterising information is a pictorial one the indexing and retrieval problem is in general unsolved, even if several groups are developing systems suitable for defined classes of images. These procedures usually consist in characterising the input image by colour histograms, special derived features a.s.o. and search in database all documents that correspond to this image or contain it as a subimage. Because of large archives in the historical and cultural field the demand arises for content based retrieval in image databases.

Different approaches for retrieval of images from database either use geometric features of the input image or rely on the luminance and/or colour information of this image. The luminance/colour approaches [1-4] use initial information about grey values and colours of the pixels to estimate the similarities between images. One of the first ideas of such approaches is to use colour histograms [1]. Instead of using colours, grey value descriptors are also used for histograms, in particular for circular ones [2]. Another luminance approach based on local grey value invariance' s which are computed for automatically detected points of interest is described in [3]. The results of low-level retrieval of grey scale images using distance-based image dissimilarity measure are presented in [4]. An interesting approach by reduced contour coding using different resolution levels is described in [17].

But, the essential content of images is (often) determined by simple line like structure elements and their mutual relation (comp. fig. 1). That' s why an alternative "geometric" approach is to get a structural description of the image as a set of extracted geometric features (lines, vertices, arcs, etc.) and their relations and to match it to model structural descriptions or to structural descriptions of images in the database.

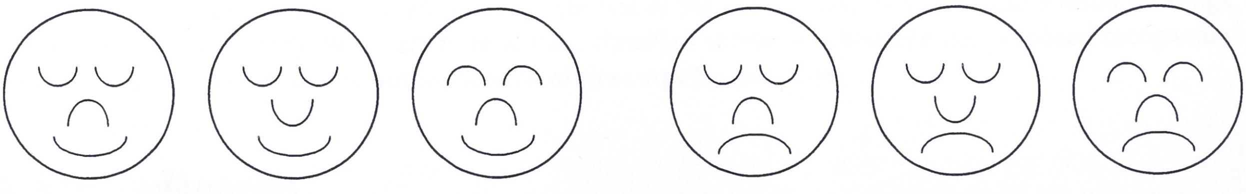

Fig. 1 Classes of pictures using line type symbolic elements, happy (left) and lousy (right)

One of the most powerful tools for describing structured objects are attributed relational graphs (ARGs). The nodes of ARGs represent primitives or subpatterns of structured objects, and branches between nodes represent relations between these primitives [5]. A number of algorithms [6-10] are developed for matching two graphs (graph isomorphism) or for matching a graph to sub-graph of another graph (subgraph isomorphism), in our context it means they are able to handle also partial image descriptions. The algorithms of syntactic structural approach [11-14] are also developed to match an input image or its description not to one but to a set of images defined by some formal tools. The names of these tools usually contain "grammar" as a key word because they have much in common with developed before grammars that generate chains of symbols. Instead of global luminance algorithms the structural approach can be used for local search of the objects and for correct retrieval in the case of partial visibility, extraneous features, it is invariant against some deformations of these objects. It must be mentioned also that difficulty of geometry based approaches is that they usually use human-made models of the objects and require special procedures to get CAD-like or vector representation of the images.

This paper addresses the problem of matching masonmarks to a large set of images. Examples of these marks are presented in fig. 2. Historians and historical architects need tools for retrieving the information about masonmarks to determine the data of creation the churches and other ancient buildings, and also for ancient paper identification containing these marks. But, this information is often spread over different institutions and the characterising information is available only pictorial. For example the "Frauenkirche" in Dresden, now under reconstruction, has about 3000 different masonmarks [15].

The algorithm for retrieval the images of masonmarks is described in Section 2. Experimental results are presented in Section 3, and finally, some concluding remarks are given in Section 4.

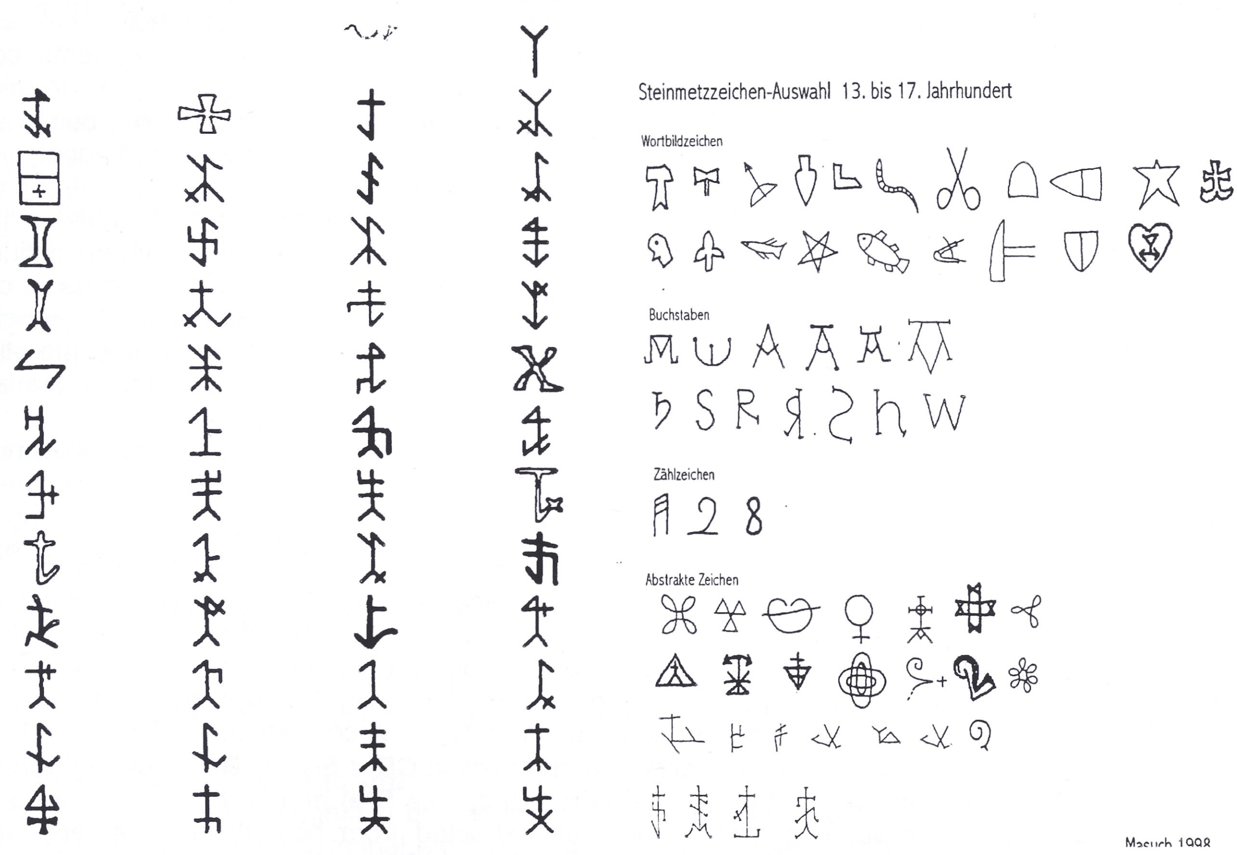

Fig. 2. Examples of masonmarks

2. Recognition of the masonmarks

The algorithm developed is intended for recognition of graphical objects including masonmarks in skeletonized images. Skeletonization is widely used as pre-processing procedure for obtaining of vector descriptions of images. There is a number of algorithms but we use a modified version of the algorithm given in [16] to get a skeleton graph of a binary image. This graph consists of arcs corresponding to skeleton line segments and nodes corresponding to ends of these lines.

The algorithm proposed in this paper for recognition of skeletonized masonmarks is based on "geometric" approach by the following main reasons:

1. The masonmarks are graphical objects consisting of geometric subpatterns (lines, arcs, circles etc.) and differ from each other mainly not by luminance features but by the shape and relative position of some of these subpatterns.

2. A widely-used way for acquisition of images of masonmarks consists in copying by hand these marks directly from stones. Some geometric features and the way of drawing can vary from one hand drawn object to another. Besides that, the skeletonization procedure leads usually to some distortions and changeable skeleton representations of not only hand drawn but standard objects also. That is why skeletonized images of masonmarks are changeable and the structural syntactic approach is the most preferable for recognition even of such non-standard objects. Under this approach some tools like formal grammars are used for defining a set of model or etalon descriptions. The proposed grammar for characterising masonmarks and the recognition algorithm using this grammar are described below.

2.1. Description of the grammar

Recognition of the objects in the skeletonized image results in extraction of all subgraphs of the skeleton graph defined by the grammar $\mathbf{G = \{ V_T, V_N, P, S \}}$, where $\mathbf{V_T}$ and $\mathbf{V_N}$ are sets of terminal and non-terminal symbols, **P** is the set of production rules and **S** is the starting symbol.

The set $\mathbf{V_T}$ of terminal symbols consists of horizontal, vertical and inclined line segments, concave and convex curved segments, initial and of empty terminal symbols also. Every terminal symbol can be assigned to some part (terminal element) of the skeleton graph if two nodes are defined as tail and head points of this part and its correspondent geometric and topological features are satisfied. Therewith the initial terminal element is the skeleton node in the picture which can be defined, e.g., by such features as the number and directions of skeleton line segments touching this node. The empty terminal element is the "white" line segment defining interconnection of not connected parts of the picture. Any other terminal element consists of one or more skeleton line segments and can be defined by the direction from tail to head point, lengths and directions of skeleton lines touching to head point of this terminal element and by some other features. The tail and head points coincide in the initial and closed curved segment terminal elements and are different points in other ones.

The set $\mathbf{V_N}$ of non-terminal symbols is the set of marks corresponding to end points of terminal elements. Production rules from the set **P** are of type

$$\mathbf{A \rightarrow aB} \text{ or } \mathbf{A \rightarrow b},$$

where **A** is the non-terminal symbol, **B** is a set of one or more non-terminal symbols and $\mathbf{a, b \in V_T}$. Each production rule results in replacing the non-terminal symbol **A** in the graph node by a terminal element (with tail point in this node) and 0, 1 or more non-terminal symbols in the head point **p** of this terminal element. In the result some subsets of non-terminal symbols can be assigned to nodes of the skeleton graph during processing the picture. Therewith the changing of a set **S(p)** of non-terminals in point **p** after substitution of **aB** for **A** consists in the following: If the set **S(p)** is not empty it is necessary to define the intersection $\mathbf{CR = S(p) \cap B}$, to subtract the set **CR** from the set **S(p)** and **B** and then to merge resulting sets **S(p)** and **B**. It means that if sets **S(p)** and **B** contain the same symbol it must be extracted from both these sets, and after extracting all such symbols the set **B** must be added to the set **S(p)**.

Main features of the grammar can be also described employing the following notions of a joint and of a primary fragment. The joint is the pair **J** = {end point, **s** $\in$ **S**$_N$}, where "end point" represents either a tail or a head point of the terminal element. The primary fragment **f(b, A, B)** is determined if the terminal element **b** and the subsets **A, B** $\in$ **V**$_N$ corresponding to its end points are defined. Two primary fragments can be joined by their common joint only. This operation is like to using one of the production rules and results in producing of a new (not primary) fragment and disappearance of the joint **J**. Generation of the picture by the grammar is the process of joining fragments. This process is considered to be finished if the certain fragment containing no joints is produced. It follows that a set of primary fragments defines images produced, and in order to form this set it is necessary to have the corresponding grammar **G** and also to determine terminal elements corresponding to terminal symbols of this grammar.

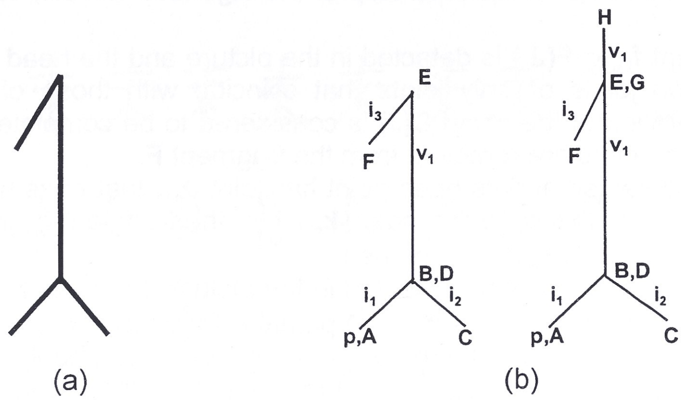

Fig. 3. The masonmark (a) and structural descriptions (b) of two corresponding skeletons

As an example let us consider the grammar ***G*** of the masonmark in fig 3a. This grammar is the set ***G*** = { V_T,V_N ,P,S }, where

V_T = { p, i_1, i_2, i_3, v_1, v_2 },

V_N = {A, B, C, D, E, F, G, H },

P = { S→ p A, A→ i_1 BD, B→ i_2 C, D→ v_1 E, D→ v_1 EG, E→ i_3 F, G→ v_2 H }.

In fig. 3b non-terminal symbols of the grammar are placed near the head points of corresponding terminal elements and terminal symbols are placed in the middle of these elements. The initial terminal element is the skeleton point such that only one skeleton segment begins in this point being left to right and bottom to top inclined one. The other terminal elements are either inclined or vertical line segments each defined by the direction from the tail to the head point and by some features of its head point also. For example, the terminal element **v_1** is the vertical line segment with the following features: 1) its tail point is lower than the head point, 2) three line segments start from the head point and 3) one of them is the right to left and top to bottom inclined line segment (**i_3**) and another is the from bottom to top short vertical one (**v_2**). Therewith this short line segment (**v_2**) can present in one skeleton of the masonmark and absent in another one depending on the thickness of the lines, inclination of the terminal element **i_3** and some parameters of the skeletonization algorithm.

Thus defined grammars can be used also for specifying a set of some masonmarks as well as for defining their individual fragments.

2.2. Description of the algorithm

The first step of the recognition procedure consists in detection of the skeleton points satisfying the features of the initial terminal element. The primary fragment containing this element has one or more joints and represents at this step the intermediate fragment **F** of the object detected. Next steps of the algorithm consist in tracing in the image so called permissible chains of primary fragments and joining these chains to the joints of the fragment **F**. Therewith the chain **C(J)** of joined primary fragments is considered as permissible one for the joint **J** of the fragment **F** if the tail point of this chain has the joint **J** and the head point has either no or only joints coinciding with those of the generated fragment **F**. The process of growing in such way of the fragment **F** results either in 1) detection of the object if the fragment **F** containing no joints is produced or 2) refusal of recognition the object that contains the current initial terminal element, if any permissible chain for the joint of the fragment **F** can not be detected.

Let us consider some features of the algorithm for detection permissible chains of primary fragments. Tracing of primary fragments in the image is carried out by special procedures each for tracing of the certain type of terminal elements (horizontal, vertical, inclined, curved and initial ones), and every of these procedures has some parameters defining the features of these elements. Let at some step of the algorithm the chain **C(J)** consisting of **n** primary fragments is traced in the image. Let $\mathbf{J_n}$ be the joint at the head point of this traced part and $\mathbf{F(J_n)}$ be the list of primary fragments $\mathbf{f_{i,n}}$ that can be joined to $\mathbf{J_n}$. At the next step of the algorithm we can obtain one of the following three results :

1) The primary fragment $\mathbf{f_{i,n} \in F(J_n)}$ is detected in the picture and the head point of the detected fragment has either no joints or only joints that coincide with those of the generated fragment **F**. In this case generation of the chain **C(J)** is considered to be completed, and all joints of this chain in its tail and head points are removed from the fragment **F**.

2) The fragment $\mathbf{f_{i,n}}$ is detected and its head point has joint $\mathbf{J_{n+1}}$ that does not coincide with any other joint of the fragment **F**. In this case the index $\mathbf{k_n = i}$ of the fragment $\mathbf{f_{i,n}}$ in the list $\mathbf{F(J_n)}$ is stored and the next fragment $\mathbf{f_{i,n+1} \in F\ (J_{n+1})}$ is searched.

3) Any primary fragment $\mathbf{f_{i,n}}$ can not be detected in the picture. In this case we must return to the previous joint $\mathbf{J_{n-1}}$ (i.e. to restore it) to search the primary fragment $\mathbf{f_{i,n-1} \in F(J_{n-1})}$ with the index $\mathbf{i > k_{n-1}}$. If such fragment is detected the index $\mathbf{k_{n-1} = i}$ of this fragment is stored and the next step of the algorithm is started. Otherwise we must return to $\mathbf{J_{n-2}}$ joint and so on. At last, if no other primary fragment can be detected for the first joint **J,** recognition of the object is cancelled and the search of the next initial terminal element is carried out.

There are two most widely used types of recognition algorithms using syntactic approach. Algorithms of the first type [11, 12] consist of two following steps. At the first step the picture processed is represented as a chain or a graph of connected terminal elements. At the second step syntactic analysis of such chains or graphs is carried out. Efficiency of these algorithms depends in a great extent on the quality of picture segmentation. The algorithms of the second type [13,14] overcome this deficiency by detecting at the first step all terminal elements in the picture what leads to much more memory and time expenses and to more complicated analysis procedures of such data too.

The proposed algorithm is based on the concept of tracing of terminal elements in the skeletonized image and takes up an intermediate position between the two types of described algorithms. Detection of terminal elements is fulfilled in the process of recognition under grammar control that usually leads to lesser (in comparison with [13, 14]) number of computations at rather high reliability of object recognition. A second essential peculiarity of the algorithm is that some features of the terminal elements and the application of special tracing procedures makes it possible to detect objects crossed by lines or touched by other objects.

3. Experiments

A program for recognition of graphical objects including masonmarks based on the developed algorithms has been implemented. The results of recognition consist of coordinates of masonmarks in the picture, it means in the document under treatment, and of their structural, i.e. contentional, descriptions, too. For introducing and recognition a new masonmark only its description in a predefined ASCII format must be formed as a separate file or it can be derived from a picture under question by algorithms to developed. This description defines the features of the primary fragments of the masonmark as well as such global features of the detected object as its size and proportions.

Different images were used for testing of the developed algorithm. An example of such image is shown in fig. 4. This image has the original size 150x120 mm^2 and contains 135 handdrawn masonmarks of 9 different types. The skeletonization of this image takes about 1 sec. (Pentium 233) and the storage volume of the skeleton representation after its slight linear approximation is about 19.2 kbyte. The masonmarks of all types in succession were searched in the image and the total time of this search takes 0.2 sec. Therewith a masonmark is considered as identified if all (not only some) its parts are detected in the image. The results of recognition process for this image are as follows: there are no errors but two masonmarks are not identified due to noise and some drawbacks of skeleton representation.

As the result of experiments has been found that recognition procedures based on the developed algorithm are fast and can be used for matching of input images of a large variety of masonmarks to a large set of images in database. The reliability of recognition depends mainly on the quality of the processed pictures and is about 97% for accurately hand drawn masonmarks.

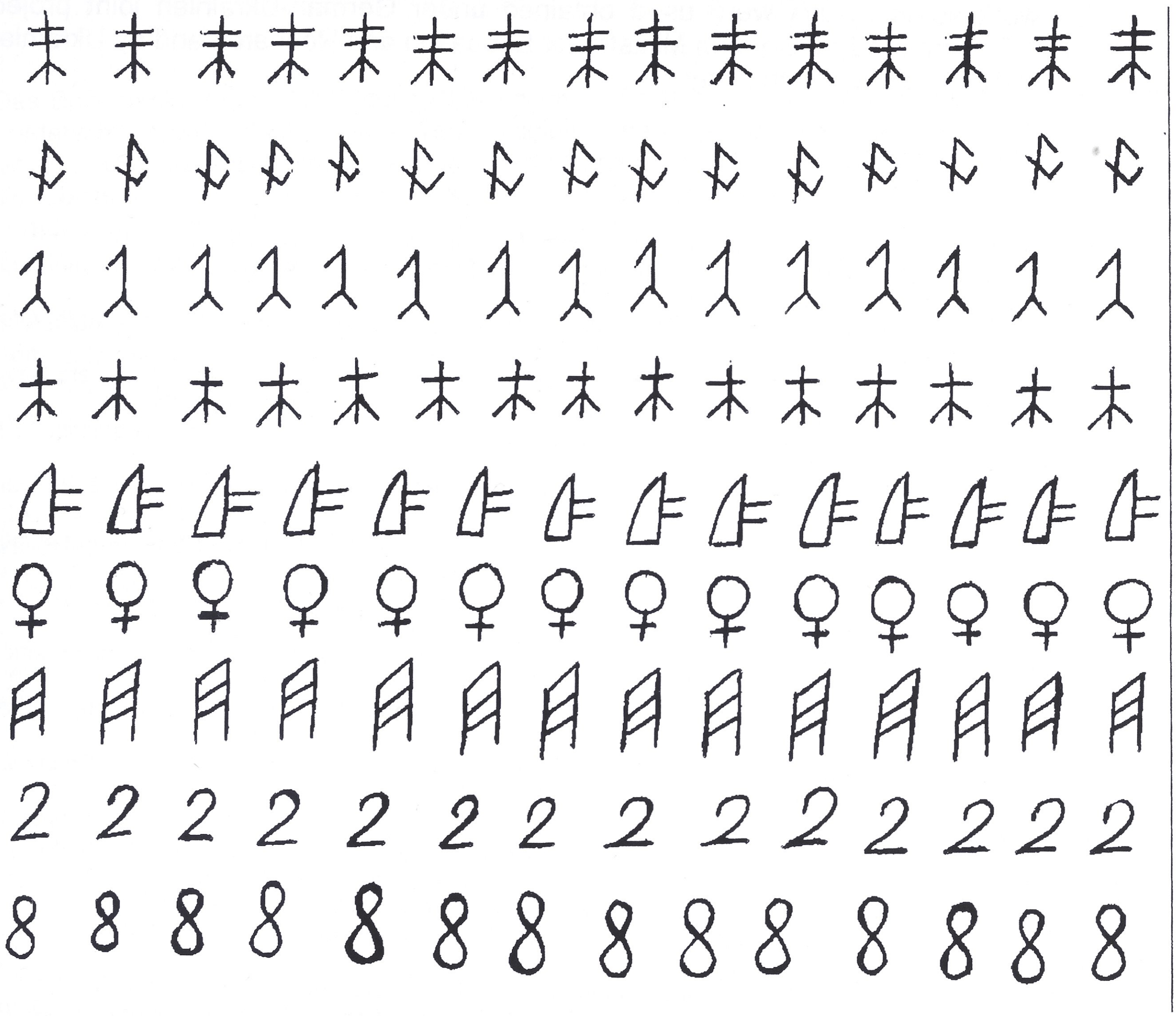

Fig. 4 sets of tested hand drawn masonmarks showing different characteristic shapes

4. Conclusion

A algorithm for recognition the images of masonmarks has been developed and implemented. This algorithm is based on a structural syntactic approach and can be used to identify and to describe input images or to retrieve images of masonmarks in large databases. The next goal of research is to develop grammars for more noise-stable fragment recognition (retrieval using only subpatterns) of masonmarks and to develop a friendly interface for supporting and editing these grammars.

Acknowledgement

For this work some results were used obtained under German-Ukrainian joint project UKR-007/97, supported by German Ministry for Education and Research and by Ukrainian Ministry for Science and Technology.

References

[1] M. J. Swain and D. H. Ballard, "Color indexing," *Int'l J. Computer Vision*, vol. 7, no. 1, pp11-32, 1991.

[2] C. Rauber, P. Tschudin and T. Pun, "Retrieval of images from a library of watermarks for ancient paper identification," *Proc. of the EVA*, Berlin, V14, 1997.

[3] C. Schmid and R. Morh, "Local grey value invariants for image retrieval," *Pattern Analysis Mach. Intell.*, 19(5), pp. 530-535, 1997.

[4] V. Starovoitov, "A new approach to low-lewel retrieving of grey scale images from image Data Base," *Proc. of 10th Scandinavian Conf. on Image Analysis*, Finland, June 1997, Vol. 2, pp. 781-788.

[5] A. Sanfeliu and K. S. Fu, "A distance between attributed relational graphs for pattern recognition," *IEEE Trans. Systems Man Cybernetics*, 13(3), 353-362, 1983.

[6] A. K. C. Wong, M. You and S. C. Chan, "An algorithm for graph optimal monomorpism," *IEEE Trans. Systems, Man Cybernetics* 20(3), pp. 628-636, 1990.

[7] M. A. Eshera and K. S. Fu, "An image understanding system using attibuted symbolic representation and inexact graph matching," *Pattern Analysis Mach. Intell.*, 8(5), pp. 604-617, 1986.

[8] S. Gold and A. Rangarajan, "A graduated assignment algorithm for graph matching," *IEEE Trans. Pattern Analysis Mach. Intell.*, 18(4), pp. 377-388, 1996.

[9] F. Depiero, M. Trivedi and S. Serbin, "Graph matching using a direct classification of node attendance," *Pattern Recognition*, 29(6), pp. 1031-1048, 1996.

[10] Y. El-sonbaty and M. A. Ismail, "A new algorithm for subgraph optimal isomorphism," *Pattern Recognition*, 31(2), pp. 205-218, 1998.

[11] Fu K. S., "Syntactic pattern recognition and stochastic languages", in: *Frontiers of Pattern Recognition (S.Watanabe, ed.),* Academic Press, New York, 1972.

[12] C.-S. Fahn, J.-F. Wang and J.-Y. Lee, "A topology-based component extractor for understanding electronic circuit diagrams," *Computer vision, graphics and image processing*, 44, pp. 119-138(1988).

[13] H. Yamada, "Contour UP matching method and its application to handprinted chineze character recognition," *Proc. IEEE seventh. Int. Conference on Pattern Recognition*, pp. 389-392(1984).

[14] A. Rosenfeld, R. Hummel, S. Zucker, "Scene labeling by relaxation operation", *IEEE Trans. Syst. Man and Cybernetics,* 6, 420-433(1976).

[15] H. Masuch, "Steinmetzzeichen – Eine Einführung zu einer systematischen Zusammenfassung", *in: Berichte über die Tätigkeit der Bau- und Kunstdenkmalpflege in den Jahren 1989-1990, Niedersächsische Denkmalpflege Band 14*, Institut für Denkmalpflege, Hannover, 1992

[16] V. M. Kiyko and M. I. Schlesinger, "Width-independent fast skeletonization algorithm for binary pictures," *Intern. Journal of Imaging Systems and Technology*, vol. 3, pp. 222-226, 1991.

[17] K. Hirato, T. Kato, "Query by visual Example", *Proceedings EDBT*, Wien, Inf.-Fachbereich, 1992