Complex and Composite Graphical Symbol Recognition and Retrieval: A Quick Review

K.C. Santosh$^{(\boxtimes)}$

Department of Computer Science, The University of South Dakota,
414 E Clark Street, Vermillion, SD 57069, USA
santosh.kc@usd.edu

Abstract. One of the key difficulties in graphics recognition domain is to work on complex and composite symbol recognition, retrieval and spotting. This paper covers a quick view on complex and composite symbol recognition, which is inspired by real-world industrial problem. Considering it as a pattern recognition problem, three different approaches: statistical, structural and syntactic are taken into account. It includes fundamental concepts or techniques and research standpoints or directions derived by a real-world application.

Keywords: Symbol recognition · Complex and composite symbols · Statistical · Structural and syntactic approaches · Document processing

1 Introduction

1.1 Overview

This paper substantially covers contributions made by the author (PhD thesis, 2011 at the INRIA Nancy Grand Est, Université de Lorraine [52]). Starting with the lineal and/or obvious isolated symbol, the paper reaches the research standpoints that were contributed by focussing on Fresh FP-6 Strep European project. Therefore, this view may not cover the whole literature, we have in graphics recognition community. At this moment, a thorough review can be found in [37,63,64].

1.2 What Are Graphical Symbols?

Graphical symbols are referred to as visual images or designs, interpreting information about the context. They are generally 2D-graphical shapes, including their composition in the highest level of conceptual information. Overall, it plays a crucial role in a variety of applications such as automatic interpretation and recognition of circuit diagrams [22,44], engineering drawings and architectural drawings [16,20,35,79], line drawings [81], musical notations [46], maps [51], mathematical expressions [6], and optical characters [28,59,60,84]. To avoid possible confusions, in this paper, other works – in the framework of graphics recognition – such as logo detection/spotting and music scores are not taken in this study.

1.3   Organization of the Paper

The remainder of the paper is organized as follows. Section 2 deals with the position of graphical symbol recognition in framework of document processing. It also includes regular contests (primarily organized together with the international workshops) and the real-world issues whether they have covered. Section 3 reviews research standpoints by categorizing them into three different approaches: statistical, structural and syntactic, and personal views based on more than a decade of experience. It also includes possible use of hybrid approaches (taking two or more approaches to develop) for graphics recognition. Section 4 concludes the paper by highlighting remarks.

2   Document Image Processing (DIP)

2.1   Graphical Symbol Recognition: Where Does It Lie in DIP?

Document analysis or processing is mainly related to texts and graphics. It includes text and or graphics separation, localization and recognition [24]. According to [42], document analysis is related to document image analysis (DIA) since both research works have been concerned with document image interpretation. In a similar manner, Kasturi et al. [31] categorize document image analysis into two domains:

(1) textual processing and
(2) graphics processing.

In both articles [31,42], the basic tasks are image segmentation, layout understanding and graphics recognition. Graphical symbol recognition, in particular, has a long history since the 70's, and it is considered as the core part of graphical document image analysis and recognition systems. In 1998 [70,71], a prominent researcher has made a statement: '*none of these methods works*' in general. Since then, it has been actively extended [17,36,71,73,74]. Very recently, the importance and the usage of graphics recognition have been reported [17,37,63,64].

Graphics are often combined with text, illustration, and color. Therefore, in document image processing, graphical symbols, for instance, convey crucial cues about the context in comparison to texts. Beside generic approaches, text recognition is distinct from symbol recognition, even though their boundaries are not obvious. Thus, their solutions complement each other [7,24].

The recognition of graphical symbols or any meaningful shapes has been the subject of numerous reviewed research articles [7,18,29,30,35,38,40]. Most of these proposed systems are roughly described using the following two major units: (1) data acquisition and preprocessing; and (2) data representation and recognition. The techniques used in data acquisition and preprocessing vary since they are problem dependent. Text/graphics separation aims at segmenting document into two layers so that one can focus on regions-of-interest, where graphical symbols lie. The usefulness of text/graphics separation can be found in [67]. Graphical symbols are represented either in the form of feature vectors

by estimating the overall shape or in more structured forms (i.e., graphs) by using meaningful primitives that are extracted from the whole symbol. Again, primitive selection tools are application dependent. As a consequence, matching techniques follow the way we represent symbols, to be used in the decision process. In general, a good data representation is assumed to be compact and discriminant, and minimizes the intra-class distance and maximizes the inter-class distance [36]. Existing approaches, specifically those based on feature based matching, can mainly be split into three different categories: statistical, structural and syntactic (see Sect. 3).

2.2 Do Regular Contests Hit Real-World Issues?

This section aims to include how far regular contests cover real-world issues (or problems). Since 1915, international association of pattern recognition (IAPR) sponsored graphics recognition (GREC) workshops, supported by technical committee 10 (TC-10: http://iapr-tc10.univ-lr.fr/) have been organized together with several contests: graphical symbol recognition, retrieval and spotting. The primary objective of the GREC contests is to evaluate the state-of-the-art of graphics recognition techniques and to generate performance evaluation tools and datasets for future research [13–15,48]. Figure 1 shows a few model symbols [26,78]. Beside several other contests, in recent years, researchers figured out the significance of 'end-to-end document analysis benchmarking' and 'open resource sharing repository' to advance as well as to facilitate fair comparison [33,34].

Considering a real-world problem, symbol recognition is not straightforward as shown in Fig. 1. In general, common applications are recognition and localization (in some cases, we call it - spotting) of graphical symbols in electronic documents, in architectural floor plans (see Fig. 2), wiring diagrams and network drawings [15,36,52]. More specifically, in this paper, a challenging problem has been addressed (see Fig. 3), where the dataset is composed of a variety of symbols such as lineal, complex and composite. These samples (entitled as FRESH dataset) are taken from [76]. The symbols may look either very similar in shape – and only differ by slight details – or completely different from a visual point of view [25,50,75]. Symbols may also be composed of other known and significant symbols and need
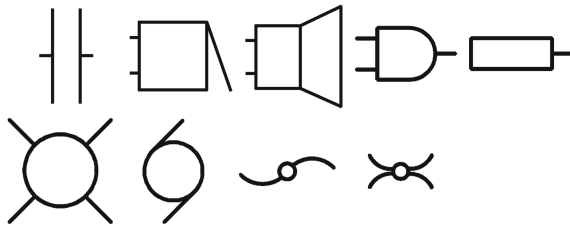
Fig. 1. Graphical symbols. An example illustrating lineal and fully isolated symbols [26].
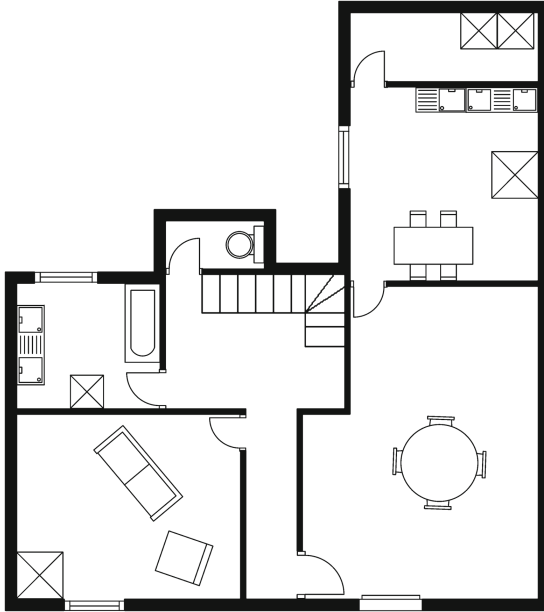
Fig. 2. Several different graphical symbols, appearing in the floor plans [15]. An example of how one can go for symbol spotting.

(a)   (b)   (c)   (d)
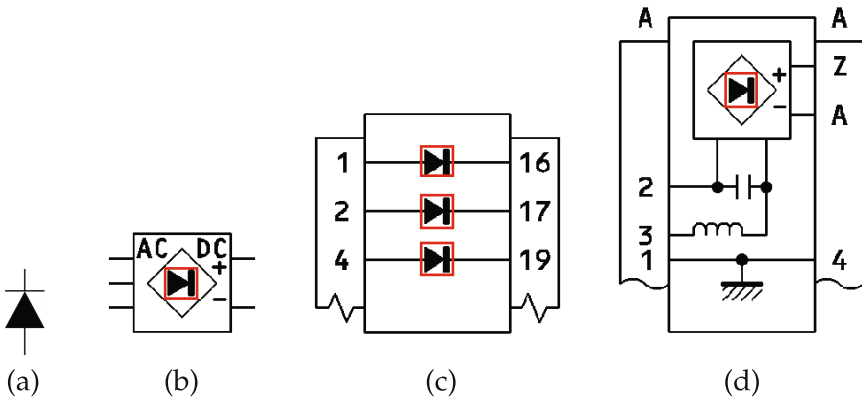
Fig. 3. An example of (a) a query and (b–d) graphical symbol/element spotting. This also illustrates the complexity of the FRESH dataset [52,76], starting with an isolated graphical symbol (moving from left to right). This example shows how the known and significant part can be spotted based on the applied query.

not necessarily be connected. For such complex and composite symbols, as before, an isolated query symbol is applied not only to recognize similar symbols from the dataset, but also to detect known and significant parts (graphical elements) which

are associated with the query symbol. Therefore, we are not just limited to symbol recognition problem. We also need to spot the visual elements (i.e., meaningful parts or regions). Besides, it is expected to see how two different symbols are similar to what extent. In the literature, the latter issue has been considered as the most challenging problem. On the whole, we refer to this task as either the parts or the whole graphical symbol recognition [15,36,39,43,49,52]. Such a problem requires strong knowledge about how we can represent graphical symbols and recognize, detect or spot them.

3   Research Standpoints

In general, the whole symbol recognition process is based on either

(1) matching features between a query and dataset symbols or
(2) comparing decomposed parts (or meaningful regions/primitives) such as lines and arcs as well as the relations between them.

These are commonly categorized into three different approaches: statistical, structural and syntactic. In most of the cases, methods are particularly suited for isolated line symbols, not for composite symbols connected to a complex environment [10,11,35,36].

3.1   Statistical Approaches

The techniques used in statistical approaches fairly computes the differences between two feature vectors [10,41]. An overview of the performance of the most commonly used shape descriptors (refer to the statistical approaches) for symbol representation is provided in [80]. Global shape representation is widely used because of its implementation simplicity since it does not require extra pre-processing and segmentation, in contrast to local pattern representation [1,85]. For more details about shape and symbol recognition, we refer to the works presented in [10,11,69], where usefulness of the shape descriptors for document analysis and a collection of techniques employed for graphical symbols recognition have been reported. Most methods are particularly suited for isolated line symbols, not for composed symbols connected to a complex environment [11,36]. In statistical approaches, global signal-based descriptors [4,32,66,84–86] are usually quite fault tolerant to image distortions, since they tend to filter out small change in details.

For a thorough shape analysis, in [82], authors computed a histogram for every pixel to figure out the distribution of constraints among the other pixels. These histograms are then statistically integrated to form a feature vector. In [87], authors proposed a similarity assessment of graphical symbols based on Kullback-Leibler divergence, where symbols are represented as 2D kernel densities. In a similar way, in [21], the authors deal with the changes in appearance (i.e., shape) from which these types of symbols differ. In [3], authors describe another framework to learn a model of shape variability in a set of patterns. Further, the Radon transform (RT)

has also been widely used to globally describe the shape of any pattern [12,66]. Motivated by this, the histogram of the RT has been used instead of compressing them (i.e., profiles) into a single vector [65], assuming that the studied patterns are equal in size. We remind the readers that the RTs are essentially a set of parametrized histograms. Therefore, in contrast to [65], in [54,57], authors used dynamic time warping (DTW) to align every histogram for each projecting angle to absorb varying histogram sizes resulting from image signal variations. In a recent PhD thesis [27], author developed the bridge between the literature of sparse representation and the visual vocabulary construction by apply the learned dictionary algorithm for learning a visual vocabulary based on local descriptors of symbols. These technique are, unfortunately, inappropriate in case where symbols are composed of other known and significant symbols (need not necessarily be connected) as well as texts (see Fig. 3).

3.2 Structural Approaches and Possible Integration with Statistical Features

Structural approaches are particularly used when symbols are decomposed into several meaningful graphical elements [20,47,52], for instance. Their interpretations, however, depend on the studied application as well as on their specific local context. Therefore, primitive extraction and symbol recognition, on the whole, are the key steps toward understanding and interpreting content within a document. They mainly include embedded graph based classification problems like attributed relational graphs (ARG) [39], region adjacency graphs (RAG) [35] and constraint networks [2], in general. Structural techniques are able to handle all types of symbols (e.g., isolated, composite), and have a powerful relational representation. However, they suffer from intense computational complexity due to the general NP-hard problem of sub-graph matching resulting from the variation of graph structure with the level of noise, occlusion, distortion *etc.*

In case of composite documents (wiring diagram, a real-world industrial problem, see Fig. 3) that contain textual and graphical elements, one needs to be able to extract and formalize the links that exist between the images and the surrounding text, in order to exploit the information embedded in those documents. Therefore, correct extraction, representation of both visual data, textual and graphical structures, and organization are the first steps towards further automated knowledge, information discovery and information retrieval or data mining on more complex data than just text. Within this context, we primarily focus on three main items:

(1) the extraction of visual elements (vocabulary) that compose an image [47];
(2) the expression of visual relations between the elements;
(3) knowledge discovery, formal learning techniques and classification using the vocabulary and relations mentioned above, including vocabulary shape analysis.

Having the aforementioned framework, we have tested the use spatial relational signatures between the possible pairs of labeled vocabulary types such as *circles*

and *corners.* These are basically used as a basis for building an attributed relational graph (ARG) that fully describes the symbol [56,61]. Thanks to our labeling of attribute types, corresponding relation alignments are possible between the two graphs while avoiding the general NP-hard graph matching problem. Further, very recently, we have introduced a new concept named bag-of-relations (BoRs) for symbol recognition, retrieval and spotting [62]. The key characteristic of the technique is to use topological relation information to categorize them in terms of bags and to guide directional relations. The method has been extended to a variety of datasets (symbols) in the domain. Further, usefulness of the method for symbol spotting and for user-friendly symbol retrieval applications has been attested.

Again, recognizing isolated symbols does not solve the real-world problem (wiring diagram, for instance), since symbols may appear either very similar in shape – and only differ by slight details – or completely different from a visual point of view [25,50,75]. In such a case, statistical signatures using shape descriptors, for instance, do not perform well because they take global appearance into account, and on the other side, usability of structural approaches may be limited due to intense computational complexity. In such a context, addressing the interest of integrating two different worlds would be a new scope of the work. Considering both approaches: structural and statistical, we have addressed the use of their best possible efficient combinations [52,58], which has been highlighted in the GREC-2010 workshop [74]:

> *'... the recurring wish for methods capable of efficiently combining structural and statistical methods'* and *'the very structural and spatial nature of the information we work with makes structural methods quite natural in the community.'*

In [56], the method is primarily based on the spatio-structural description of visual 'vocabulary'. But, it lacks the information about their shape features. Therefore, keeping the ARG based symbol description as reported in [56], shape signatures are integrated in two different ways to improve the performance. First, shape signatures are for labeling vertices [55]. Second, shape features are applied only to the vocabulary which show significant shape variations, and are grouped them via unsupervised clustering [58]. In both cases, major set of well-known state-of-the-shape descriptors are integrated with spatial relations. Overall, we bring an attention to the use of a hybrid approach in symbol recognition, and try to avoid the shortcomings of each of them: structural and statistical.

The problem can further be extended to symbol spotting, but one can view this as a kind of retrieval [15,45,49,62,68], which is guided by user queries. Additionally, the recognition/retrieval process can be made with the help of local descriptors like scale invariant feature transform (SIFT) and with the use of techniques like bag-of-features so that either primitive or region extraction (segmentation) can be avoided. The question always remains the same, 'what approach does what (i.e., performance) in which context?'.

3.3 Syntactic Approaches

The techniques based on graph grammar will be more suitable to search symbols in technical documents where information is close to a feature vector description that follows composition rules of primitives [8, 9, 19, 23, 53, 77, 83].

In the earlier framework (see Sect. 3.2 in second paragraph), very interestingly, we have presented the use of formal learning techniques to automatically learn non-trivial descriptions of symbols [53]. It means that we have transformed the vocabulary sets and their possible relations that exist between them into a first-order logic (FOL) description for a complete image. This representation is then used as an input to an inductive logic programming (ILP) solver, in order to deduce non obvious characteristics that may lead to a more semantic related recognition process. Considering the experience we have so far, the idea is appropriate for classifying a set of symbols (images) characterizing common behavior with respect to another set of symbols (images) or counter samples. But by definition, even though the reported concept in [53] can be extended to any image synthesis problem (for different application domain). It, however, is challenging to transform statistical signatures (quantified values) to the closer semantics.

4 Remarks

No doubt, graphics recognition has an extremely rich state-of-the-art literature in symbol recognition and localization [5, 11, 17, 36]. They are limited to solve specific problems, motivated by the limited posed by industrial partners. This means that major state-of-the-art methods for symbol recognition, do not conclude on the existence of a set of generic methods that can yield the best results, even though they are easy to implement (with fewer parameters) and are reproducible. This has been seen on three different approaches: statistical, structural and syntactic. Besides, very similar statement has been reported in [74], where the author pointed out: '*which features distinguish graphics recognition from general pattern recognition problems?*'. Therefore, there still exists growing interest in graphics recognition domain.

References

1. Adam, S., Ogier, J.M., Cariou, C., Mullot, R., Labiche, J., Gardes, J.: Symbol and character recognition: application to engineering drawings. Int. J. Doc. Anal. Recogn. **3**(2), 89–101 (2000)
2. Ah-Soon, C., Tombre, K.: Architectural symbol recognition using a network of constraints. Pattern Recogn. Lett. **22**(2), 231–248 (2001)
3. Almazán, J., Fornés, A., Valveny, E.: A non-rigid appearance model for shape description and recognition. Pattern Recogn. **45**(9), 3105–3113 (2012)
4. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. **24**(4), 509–522 (2002)
5. Bunke, H., Wang, P.S.P. (eds.): Handbook of Character Recognition and Document Image Analysis. World Scientific, Singapore (1997)

6. Chaudhuri, B.B., Garain, U.: An approach for recognition and interpretation of mathematical expressions in printed document. Pattern Anal. Appl. **3**(2), 120–131 (2000)

7. Chhabra, A.K.: Graphic symbol recognition: an overview. In: Proceedings of 2nd International Workshop on Graphics Recognition, Nancy, France, pp. 244–252, August 1997

8. Claus, V., Ehrig, H., Rozenberg, G. (eds.): Graph-Grammars and Their Applications to Computer Science and Biology. LNCS, vol. 73. Springer, Heidelberg (1979)

9. Cordella, L.P., Foggia, P., Genna, R., Vento, M.: Prototyping structural descriptions: an inductive learning approach. In: Amin, A., Dori, D., Pudil, P., Freeman, H. (eds.) SSPR /SPR 1998. LNCS, vol. 1451, pp. 339–348. Springer, Heidelberg (1998). doi:10.1007/BFb0033252

10. Cordella, L.P., Vento, M.: Symbol and shape recognition. In: Proceedings of 3rd International Workshop on Graphics Recognition, Jaipur, India, pp. 179–186, September 1999

11. Cordella, L.P., Vento, M.: Symbol recognition in documents: a collection of techniques? Int. J. Doc. Anal. Recogn. **3**(2), 73–88 (2000)

12. Deans, S.R.: Applications of the Radon Transform. Wiley Interscience Publications, New York (1983)

13. Delalandre, M., Ramel, J.-Y., Sidere, N.: A semi-automatic groundtruthing framework for performance evaluation of symbol recognition and spotting systems. In: Kwon, Y.-B., Ogier, J.-M. (eds.) GREC 2011. LNCS, vol. 7423, pp. 163–172. Springer, Heidelberg (2013). doi:10.1007/978-3-642-36824-0_16

14. Delalandre, M., Valveny, E., Lladós, J.: Performance evaluation of symbol recognition and spotting systems: an overview. In: Kise, K., Sako, H. (eds.) Proceedings of International Workshop on Document Analysis Systems, pp. 497–505. IEEE Computer Society (2008)

15. Delalandre, M., Valveny, E., Pridmore, T., Karatzas, D.: Generation of synthetic documents for performance evaluation of symbol recognition and spotting systems. Int. J. Doc. Anal. Recogn. **13**(3), 187–207 (2010)

16. Devaux, P.M., Lysak, D.B., Kasturi, R.: A complete system for the intelligent interpretation of engineering drawings. Int. J. Doc. Anal. Recogn. **2**(2/3), 120–131 (1999)

17. Doermann, D., Tombre, K.: Handbook of Document Image Processing and Recognition. Springer, London (2014)

18. Doermann, D.S.: An introduction to vectorization and segmentation. In: Tombre, K., Chhabra, A.K. (eds.) GREC 1997. LNCS, vol. 1389, pp. 1–8. Springer, Heidelberg (1998). doi:10.1007/3-540-64381-8_34

19. Dori, D., Pnueli, A.: The grammar of dimensions in machine drawings. Computer Vision, Graphics and Image Processing **42**, 1–18 (1988)

20. Dosch, P., Tombre, K., Ah-Soon, C., Masini, G.: A complete system for analysis of architectural drawings. Int. J. Doc. Anal. Recogn. **3**(2), 102–116 (2000)

21. Escalera, S., Fornés, A., Pujol, O., Lladós, J., Radeva, P.: Circular blurred shape model for multiclass symbol recognition. IEEE Trans. Syst. Man Cybern. Part B Cybern. **41**(2), 497–506 (2011)

22. Feng, G., Viard-Gaudin, C., Sun, Z.: On-line hand-drawn electric circuit diagram recognition using 2D dynamic programming. Pattern Recogn. **42**(12), 3215–3223 (2009)

23. Flasiński, M.: Characteristics of edNLC-graph grammar for syntactic pattern recognition. Comput. Vis. Graph. Image Process. **47**, 1–21 (1989)

24. Fletcher, L.A., Kasturi, R.: A robust algorithm for text string separation from mixed text/graphics images. IEEE Trans. Pattern Anal. Mach. Intell. **10**(6), 910–918 (1988)
25. FRESH: Final report on symbol recognition with evaluation of performances. http://www.aero-scratch.net/fresh.html. Deliverable 2.4.2.-FP6-516059
26. GREC: International symbol recognition contest at GREC2003 (2003). http://www.cvc.uab.es/grec2003/SymRecContest/
27. Ha, D.T.: Sparse representation over learned dictionary for document analysis. Ph.D. thesis, LORIA, Université de Lorraine, France (2014)
28. Heutte, L., Nosary, A., Paquet, T.: A multiple agent architecture for handwritten text recognition. Pattern Recogn. **37**(4), 665–674 (2004)
29. Jain, A.K., Duin, R.P.W., Mao, J.: Statistical pattern recognition: a review. IEEE Trans. Pattern Anal. Mach. Intell. **22**(1), 4–37 (2000)
30. Kasturi, R., Raman, R., Chennubhotla, C., O'Gorman, L.: document image analysis: an overview of techniques for graphics recognition. In: Pre-proceedings of IAPR Workshop on Syntactic and Structural Pattern Recognition, Murray Hill, NJ (USA), pp. 192–230 (1990)
31. Kasturi, R., O'Gorman, L., Govindaraju, V.: Document image analysis: a primer. Charact. Recogn. **27**(1), 3–22 (2002)
32. Kim, W.Y., Kim, Y.S.: A region-based shape descriptor using Zernike moments. Sig. Process. Image Commun. **16**(1–2), 95–102 (2000)
33. Lamiroy, B., Lopresti, D.P.: An open architecture for end-to-end document analysis benchmarking. In: International Conference on Document Analysis and Recognition, pp. 42–47 (2011)
34. Lamiroy, B., Lopresti, D.P., Korth, H.F., Heflin, J.: How carefully designed open resource sharing can help and expand document analysis research. In: Document Recognition and Retrieval XVIII, Part of the IS&T-SPIE Electronic Imaging Symposium
35. Lladós, J., Martí, E., Villanueva, J.J.: Symbol recognition by error-tolerant subgraph matching between region adjacency graphs. IEEE Trans. Pattern Anal. Mach. Intell. **23**(10), 1137–1143 (2001)
36. Lladós, J., Valveny, E., Sánchez, G., Martí, E.: Symbol recognition: current advances and perspectives. In: Blostein, D., Kwon, Y.-B. (eds.) GREC 2001. LNCS, vol. 2390, pp. 104–128. Springer, Heidelberg (2002). doi:10.1007/3-540-45868-9_9
37. Lladós, J., Rusiñol, M.: Graphics Recognition Techniques, pp. 489–521. Springer, London (2014)
38. Loncaric, S.: A survey of shape analysis techniques. Pattern Recogn. **31**(8), 983–1001 (1998)
39. Luqman, M.M.: fuzzy multilevel graph embedding for recognition, indexing and retrieval of graphic document images. Ph.D. thesis, Francois Rabelais University of Tours France and Autonoma University of Barcelona Spain (2012)
40. Marshall, S.: Review of shape coding techniques. Image Vision Comput. **7**(4), 281–294 (1989)
41. Müller, S., Rigoll, G.: Engineering drawing database retrieval using statistical pattern spotting techniques. In: Proceedings of 3rd International Workshop on Graphics Recognition, Jaipur, India, pp. 219–226, September 1999
42. Nagy, G.: Twenty years of document image analysis in PAMI. IEEE Trans. Pattern Anal. Mach. Intell. **22**(1), 38–62 (2000)
43. Nayef, N.: Geomatric-based symbol spotting and retrieval in technical line drawings. Ph.D. thesis, University of Kaiserslautern, Germany (2012)

44. Okazaki, A., Kondo, T., Mori, K., Tsunekawa, S., Kawamoto, E.: An automatic circuit diagram reader with loop-structure-based symbol recognition. IEEE Trans. Pattern Anal. Mach. Intell. **10**(3), 331–341 (1988)
45. Qureshi, R.J., Ramel, J.-Y., Barret, D., Cardot, H.: Spotting symbols in line drawing images using graph representations. In: Liu, W., Lladós, J., Ogier, J.-M. (eds.) GREC 2007. LNCS, vol. 5046, pp. 91–103. Springer, Heidelberg (2008). doi:10. 1007/978-3-540-88188-9_10
46. Rebelo, A., Capela, G., Cardoso, J.S.: Optical recognition of music symbols: a comparative study. Int. J. Doc. Anal. Recogn. **13**(1), 19–31 (2010)
47. Rendek, J., Masini, G., Dosch, P., Tombre, K.: The search for genericity in graphics recognition applications: design issues of the Qgar software system. In: Marinai, S., Dengel, A.R. (eds.) DAS 2004. LNCS, vol. 3163, pp. 366–377. Springer, Heidelberg (2004). doi:10.1007/978-3-540-28640-0_35
48. Rusiñol, M., Lladós, J.: A performance evaluation protocol for symbol spotting systems in terms of recognition and location indices. Int. J. Doc. Anal. Recogn. **12**(2), 83–96 (2009)
49. Rusiñol, M., Lladós, J.: Symbol Spotting in Digital Libraries: Focused Retrieval over Graphic-rich Document Collections. Springer, London (2010)
50. Salmon, J.P.: Reconnaissance de Symboles Complexes. Ph.D. thesis, Institut National Polytechnique de Lorraine (2008)
51. Samet, H., Soffer, A.: MARCO: map retrieval by content. IEEE Trans. Pattern Anal. Mach. Intell. **18**(8), 783–798 (1996)
52. Santosh, K.C.: Graphics recognition using spatial relations and shape analysis. Ph.D. thesis, INRIA - Institut National Polytechnique de Lorraine, Université de Lorraine, November 2011
53. Santosh, K.C., Lamiroy, B., Ropers, J.P.: Inductive logic programming for symbol recognition. In: Proceedings of International Conference on Document Analysis and Recognition, pp. 1330–1334. IEEE Computer Society (2009)
54. K.C., S., Lamiroy, B., Wendling, L.: DTW for matching radon features: a pattern recognition and retrieval method. In: Blanc-Talon, J., Kleihorst, R., Philips, W., Popescu, D., Scheunders, P. (eds.) ACIVS 2011. LNCS, vol. 6915, pp. 249–260. Springer, Heidelberg (2011). doi:10.1007/978-3-642-23687-7_23
55. Santosh, K.C., Lamiroy, B., Wendling, L.: Spatio-structural symbol description with statistical feature add-on. In: Kwon, Y.-B., Ogier, J.-M. (eds.) GREC 2011. LNCS, vol. 7423, pp. 228–237. Springer, Heidelberg (2013). doi:10.1007/ 978-3-642-36824-0_22
56. Santosh, K.C., Lamiroy, B., Wendling, L.: Symbol recognition using spatial relations. Pattern Recog. Lett. **33**(3), 331–341 (2012)
57. Santosh, K.C., Lamiroy, B., Wendling, L.: DTW-radon-based shape descriptor for pattern recognition. Int. J. Pattern Recogn. Artif. Intell. **27**(3), 1350008 (2013)
58. Santosh, K.C., Lamiroy, B., Wendling, L.: Integrating vocabulary clustering with spatial relations for symbol recognition. Int. J. Doc. Anal. Recogn. **17**(1), 61–78 (2014)
59. Santosh, K.C., Nattee, C., Lamiroy, B.: Relative positioning of stroke-based clustering: a new approach to online handwritten Devanagari character recognition. Int. J. Image Graph. **12**(2), 25 (2012)
60. Santosh, K.C., Wendling, L.: Character recognition based on non-linear multi-projection profiles measure. Front. Comput. Sci. **9**(5), 678–690 (2015)
61. Santosh, K.C., Wendling, L., Lamiroy, B.: Using spatial relations for graphical symbol description. In: Proceedings of the IAPR International Conference on Pattern Recognition, pp. 2041–2044. IEEE Computer Society (2010)

62. Santosh, K.C., Wendling, L., Lamiroy, B.: BoR: Bag-of-Relations for symbol retrieval. Int. J. Pattern Recogn. Artif. Intell. **28**(06), 1450017 (2014)
63. Santosh, K., Wendling, L.: Graphical Symbol Recognition, pp. 1–22. Wiley, New York (2015)
64. Tabbone, S., Terrades, O.R.: An Overview of Symbol Recognition, pp. 523–551. Springer, London (2014)
65. Tabbone, S., Terrades, O.R., Barrat, S.: Histogram of radon transform. A useful descriptor for shape retrieval. In: Proceedings of the IAPR International Conference on Pattern Recognition, pp. 1–4 (2008)
66. Tabbone, S., Wendling, L., Salmon, J.P.: A new shape descriptor defined on the radon transform. Comput. Vis. Image Underst. **102**(1), 42–51 (2006)
67. Tabbone, S., Wendling, L., Tombre, K.: Matching of graphical symbols in line-drawing images using angular signature information. Int. J. Doc. Anal. Recogn. **6**(2), 115–125 (2003)
68. Tabbone, S., Wendling, L., Zuwala, D.: A hybrid approach to detect graphical symbols in documents. In: Marinai, S., Dengel, A.R. (eds.) DAS 2004. LNCS, vol. 3163, pp. 342–353. Springer, Heidelberg (2004). doi:10.1007/978-3-540-28640-0_33
69. Terrades, O.R., Tabbone, S., Valveny, E.: A review of shape descriptors for document analysis. In: Proceedings of International Conference on Document Analysis and Recognition, pp. 227–231 (2007)
70. Tombre, K.: Analysis of engineering drawings: state of the art and challenges. In: Tombre, K., Chhabra, A.K. (eds.) GREC 1997. LNCS, vol. 1389, pp. 257–264. Springer, Heidelberg (1998). doi:10.1007/3-540-64381-8_54
71. Tombre, K.: Ten years of research in the analysis of graphics documents: achievements and open problems. In: Proceedings of 10th Portuguese Conference on Pattern Recognition, Lisbon, Portugal, pp. 11–17, March 1998
72. Tombre, K., Chhabra, A.K. (eds.): GREC 1997. LNCS, vol. 1389. Springer, Heidelberg (1998). doi:10.1007/3-540-64381-8
73. Tombre, K.: Graphics recognition: the last ten years and the next ten years. In: Proceedings of 6th IAPR International Workshop on Graphics Recognition, Hong Kong, pp. 422–426 (2005)
74. Tombre, K.: Graphics recognition - what else? In: Ogier, J.M., Liu, W., Lladós, J. (eds.) Graphics Recognition. Achievements, Challenges, and Evolution. LNCS, vol. 6020, pp. 272–277. Springer, Heidelberg (2010). doi:10.1007/978-3-642-13728-0_25
75. Tombre, K., Lamiroy, B.: Pattern recognition methods for querying and browsing technical documentation. In: Progress in Pattern Recognition, Image Analysis and Applications, 13th Iberoamerican Congress on Pattern Recognition, pp. 504–518 (2008)
76. Tooley, M., Wyatt, D.: Aircraft electrical and electronic systems: principles, operation and maintenance. Butterworth-Heinemann, Oxford (2008). Aircraft engineering principles and practice
77. Tsai, W.H., Fu, K.S.: Attributed grammar: a tool for combining syntactic and statistical approaches to pattern recognition. IEEE Trans. Syst. Man Cybern. **10**(12), 873–885 (1980)
78. Valveny, E., Dosch, P.: Symbol recognition contest: a synthesis. In: Lladós, J., Kwon, Y.-B. (eds.) GREC 2003. LNCS, vol. 3088, pp. 368–385. Springer, Heidelberg (2004). doi:10.1007/978-3-540-25977-0_34
79. Valveny, E., Martí, E.: A model for image generation and symbol recognition through the deformation of lineal shapes. Pattern Recogn. Lett. **24**(15), 2857–2867 (2003)

80. Valveny, E., Tabbone, S., Ramos, O., Philippot, E.: Performance characterization of shape descriptors for symbol representation. In: Liu, W., Lladós, J., Ogier, J.-M. (eds.) GREC 2007. LNCS, vol. 5046, pp. 278–287. Springer, Heidelberg (2008). doi:10.1007/978-3-540-88188-9_26
81. Wendling, L., Tabbone, S.: A new way to detect arrows in line drawings. IEEE Trans. Pattern Anal. Mach. Intell. **26**(7), 935–941 (2004)
82. Yang, S.: Symbol recognition via statistical integration of pixel-level constraint histograms: a new descriptor. IEEE Trans. Pattern Anal. Mach. Intell. **27**(2), 278–281 (2005)
83. You, K.C., Fu, K.S.: Distorted shape recognition using attributed grammars and error-correcting techniques. Comput. Vis. Graph. Image Process. **13**, 1–16 (1980)
84. Yuen, P.C., Feng, G.C., Tang, Y.Y.: Printed chinese character similarity measurement using ring projection and distance transform. Int. J. Pattern Recogn. Artif. Intell. **12**(2), 209–221 (1998)
85. Zhang, D., Lu, G.: Review of shape representation and description techniques. Pattern Recogn. **37**(1), 1–19 (2004)
86. Zhang, D., Lu, G.: Shape-based image retrieval using generic fourier descriptor. Sig. Process. Image Commun. **17**(10), 825–848 (2002)
87. Zhang, W., Wenyin, L., Zhang, K.: Symbol recognition with kernel density matching. IEEE Trans. Pattern Anal. Mach. Intell. **28**(12), 2020–2024 (2006)