

Graphics Recognition: The Last Ten years and the Next Ten Years

Karl Tombre

► **To cite this version:**

Karl Tombre. Graphics Recognition: The Last Ten years and the Next Ten Years. Liu Wenyin and Josep Lladós. Graphics Recognition - Ten Years Review and Future Perspectives, 3926 (3926), Springer Verlag, pp.422-426, 2006, Lecture Notes in Computer Science. <inria-00103723>

HAL Id: inria-00103723

<https://hal.inria.fr/inria-00103723>

Submitted on 5 Oct 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Graphics Recognition: The Last Ten Years and the Next Ten Years

Karl Tombre

LORIA-INPL, École des Mines de Nancy, Parc de Saurupt,
CS 14234, 54042 Nancy CEDEX, France
`Karl.Tombre@loria.fr`

Abstract. GREC'05 marked the 10th anniversary of the series of international workshops on graphics recognition, for which the first edition was held in Penn State in 1995. At the end of the workshop, a panel discussion was held to take a broader view of our field, to examine the shift in issues addressed by research teams, and define some challenges for the coming years. This paper tries to summarize the results of the panel discussion.

1 Traditional Issues, Emerging Themes and Failures

1.1 Topics for which Interest Seems to be Drifting Away

Looking at the research themes of the groups active in graphics recognition, it appears that the interest for some themes has more or less faded away. This is especially the case for document image analysis methods for pre-processing and text/graphics separation. The question naturally arises whether these issues are considered to be solved. Although we definitely have state-of-the-art methods for these problems, it probably cannot be said that these methods can be considered as definitive solutions. However, the problem has shifted in some way:

- Some applications aim at full reconstruction of the document in a high-level CAD or SIG system. In that case, users need excellent low-level results and the existing methods are not good enough. Hence, pragmatically, the low-level segmentation tools are performed manually or semi-automatically.
- Other applications do not aim at full reconstruction but rather at providing tools for browsing and indexing large documentation databases. In this case, the available methods are usually deemed to be sufficient and no strong need is present for more work on segmentation.

Several panelists noted that there seems to be less interest in building complete systems, at least from a scientific point of view. This is not due to an applicative lack of interest for such systems, but is rather an indication of the maturity to which we have come in several areas. However, there is a belief that there are still open problems in building complete, robust systems and researchers are encouraged not to forget this aspect.

1.2 Topics Which Have Been Along for Some Time and are Still Very much Relevant

A number of issues have been constantly present in our research field and remain so. *Vectorization*, i.e. raster-to-vector conversion, is still addressed and there is room for improvement, as the arc detection contest held at the workshop made clear. While traditional problems like the vectorization of archival material may be less present, new issues arise with hand-made sketches, for instance. Among the directions mentioned for further improvement, let us cite the direct processing of the grayscale image, as it seems that we are slowly coming to the limit of the accuracy we can get out of the black and white pixels of a binary image. Have we really made progress? At least, we have certainly now a better understanding of what the good techniques are and what their limitations are.

Symbol recognition is a topic receiving a lot of attention. While we are reasonably good at recognizing fully segmented, simple symbols, as illustrated by the symbol recognition contest held at the workshop, there are a number of open issues with handwritten symbols, complex symbols made by the combination of simple symbols and textual annotations, and symbols which can not easily be segmented out of their context. The scalability of symbol recognition methods, i.e. their ability to discriminate between several hundred different symbols, also remains an open problem, where we will probably need hybrid approaches incorporating both structural representations and classification techniques.

Interesting questions also arise with *sketching* and *online graphics recognition* tools. These can be used for querying existing documents or as interactive tools for a designer. The design of such tools leads to specific challenges due to the interactive process involved in the recognition and to the large variability of handwriting. On the other hand, online data provide more information than plain raster images.

1.3 Emerging Themes

There are also new issues, or issues which have received much more interest recently, mostly because of the applications they are related to. This includes taking into account *new media*, such as documents available in electronic format but with little or no structure or semantics (typical example is a PDF document), online sketching, paper ink and e-paper. It is felt that the problem of handling legacy documents will stay with us for a long time, at least with PDF and HTML documents. In that context, there is a special interest in *cultural heritage documents*, with all the specificities they represent.

Let us also mention *information retrieval and spotting* applications where there is an increasing interest in adding graphics features to the indexes and keys for navigating large information databases. We could speak of “graphics search” within document sets. Especially for legacy documents, the focus is shifting from recognition to search. *Symbol spotting* should be mentioned explicitly in this context; the idea is to be able to quickly localize instances of a possible symbol, even without having a library of known models to match against. In an ideal

case, the user should be able to delineate a graphics area of interest and do a query for similar areas within the document or in the whole document database. This could also be qualified as unsupervised or dynamic recognition of symbols, and relies on the ability to compute quickly and efficiently a number of general signatures which can be used for indexing and querying. It also necessitates taking into account relevance feedback from the user.

In these areas, there is an overlap between our community and other research and technical communities, interested in content-based image retrieval, trademark logo recognition, layout-based retrieval, etc. We should be eager to build bridges with these communities to take advantage of each other's progress.

1.4 [Putative] Failures

Our community has also had some relative failures, with respect to the hopes and plans generated and discussed at previous workshops. One of the most visible is our inability to gather around a common base of software. There is a lot of knowledge in the community, but various groups often prefer developing their own versions of various state of the art algorithms, instead of "plugging" their own work into some standard software environment. Thus, the knowledge remains partly fragmented and everybody spends a lot of time reprogramming existing methods. This is not because of lack of open software environments, but rather a common syndrome in many research groups leading to think that "home-made is better". There was no clear consensus at the panel discussion on how to avoid this or converge towards some more satisfactory solution.

2 Some "Hot" Topics

We spent some time during the panel discussion debating some of the topics perceived as being "hottest" in our field. The following lines try to summarize the (sometimes heated, but always constructive) discussions we had.

2.1 Vectorization

A first question to address with vectorization is its definition. What is a good raster-to-vector conversion? If we say that it means looking for the central lines of the raster image, we end up with the problem of having a clear definition of what the central lines are. Do we look for line fragments or for arcs? How do we discriminate between these two without contextual knowledge.

Some panelists stressed that a vectorization system has to be universal to be of any real interest, but there are actually two main choices. The first is to aim at a universal, non-contextual system which has to adopt some compromises between arc and line segment hypotheses, between simplicity and precision, etc. The second is to have an application-driven method; in that case there may be contextual knowledge about the presence and nature of arcs, the precision or the

speed needed, the possibility or not to have some kind of user interactions, the presence of free curves, etc.

Therefore, despite the contests organized at the workshop, there seems not to be any universally approved way of defining what a good vectorization is supposed to be.

Still, it is felt that with respect to automatic, universal, non-contextual vectorization, we are close to optimal methods when dealing with black-and-white images. Further progress will either include working on the gray-level image to achieve subpixel precision and better curve segmentation, or progress in the seamless integration of user input and contextual knowledge into application-specific methods.

2.2 Analysis of Complete Documents

In many cases, graphics recognition is just a part of a broader picture where the aim is to analyze complete documents, also containing text, logos, illustrations, etc. The analysis itself can be for document image understanding purposes, but also (and actually more and more often) for indexing purposes, to let a user browse through a large document set and quickly retrieve or spot relevant information.

One area with increasing focus is that of heritage documents which have often been scanned in large digitization campaigns, the need appearing afterwards for tools to organize these scanned documents and for browsing through them. In some cases, there is a real problem with the image quality, as the digitization had been performed solely with the purpose of having document images readable by a human, not necessarily resolutions good enough for document image processing and analysis. When document analysis people are involved in such projects from the very beginning, an important recommendation for them is to see to that the digitizing aspects are not neglected and that the resolution with which the information is scanned and stored is good enough. On historical documents, a resolution of at least 600 dpi should be requested.

Another application domain with a large potential in the future is that of electronic documents available with little or no structure, such as PDF or HTML documents. Specific challenges arise for large-scale processing of such documents.

2.3 Performance Evaluation: The Contests

Organizing contests has been one of the strong points of our workshops since the first edition. We have also seen lately that the very fact that these contests have been organized, has driven research groups to publish their methods with reference to the contest data and evaluation methods, also at other conferences or in journals.

But there are some drawbacks and pitfalls, which were discussed at length during the panel debate. One of the controversial issues is the use of *noise models*. They are felt to be necessary to model real problems. But they have also led

to participant methods which try to reduce the noise by “guessing” more or less which parameters of the noise model were applied to the data. Then the question arises: Do we actually test the quality of the de-noising method or the recognition capabilities of the method? Should we limit ourselves to real data and not use synthetic data obtained through noise models applied to perfect data? One of the problems is the extreme cost of building groundtruth on real data...

It was also felt that there are too few participants in these contests. Many people do not take the extra step to set up everything and compare their results in an objective way with that of others. Also, there were no commercial tools this time. Several solutions were explored for getting more people into participating, including offering rewards or letting people compete anonymously.

Still, besides the contest which is more or less a “one-shot” event, the work on performance evaluation also allows us to make reference data and objective evaluation tools available to the community. The aim is to have regular benchmarking campaigns where we can really get beyond the point of having a winner of a contest, to get a better understanding of the strengths and weaknesses of various approaches taken for recognition tasks.

3 Conclusion

At the end of the panel, workshop attendants were asked to cite topics which would be discussed at GREC’2015. Here are some of the answers, without any further comments:

- Vectorization
- Same program as GREC’95
- Geometry-based or shape-based recognition
- Knowledge-based recognition
- Hardware and software technology evolution

Acknowledgments

Special thanks to the panelists, in alphabetical order Thomas Breuel, Alex Gribov, Josep Lladós, Gerd Maderlechner, Jean-Marc Ogier, and Liu Wenxin, and to all other attendees who contributed to this enriching panel discussion and hence to the conclusions drawn up here.