Open access • Book Chapter • DOI:10.1007/978-3-642-13728-0_25

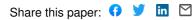# Graphics recognition: what else? — **Source link** ↗

Karl Tombre

Institutions: University of Lorraine

Published on: 22 Jul 2009 - Graphics Recognition

Topics: Grand Challenges and Graphics

Related papers:

- Symbol recognition using spatial relations

- Symbol Recognition with Kernel Density Matching

- The rise of the data translator

- The next twenty years in information retrieval: some goals and predictions

- Ubiquitous Computing: From 5G to the Edge and Beyond

Graphics Recognition – What Else?

Karl Tombre

Graphics Recognition—What Else?

Karl Tombre[1,2,3]

[1] INRIA, 615 rue du jardin botanique, 54600 Villers-lès-Nancy, France
[2] Université de Lorraine, INPL, LORIA, B.P. 239, 54506 Vandœuvre-lès-Nancy CEDEX, France
[3] CNRS, LORIA, B.P. 239, 54506 Vandœuvre-lès-Nancy CEDEX, France
Karl.Tombre@inria.fr

Abstract. This paper tries to sum up the discussions held during the sessions of GREC'09, as well as at the final panel session. As it is always good to know where you are coming from, the paper briefly takes a look back at the discussions held two years earlier, before looking ahead at the future challenges for our research community. A number of points raised two years ago remain very much valid, but we also try to identify some new grand challenges for the field of graphics recognition.

1 Introduction

In this paper, we will not repeat the brief summary of the history of graphics recognition, which was presented in the report from the GREC'07 panel discussion [1]. The historical view of the reasons for gathering a specific document analysis community, and within that a graphics recognition sub-group, remains valid and the reader should keep it in mind in the thoughts raised here.

At the end of the GREC'09 workshop, a panel session was organized to wrap up the numerous and lively discussions held throughout the sessions of the workshop, and try to come up with some more general conclusions. The panel members were Young-Bin Kwon, George Nagy, Sitaram Ramachandrula, and Karl Tombre. A number of workshop participants also contributed to the debate.

This article presents the main discussions and conclusions of the panel.

2 Looking Back: Some of our 2007 Conclusions

> *"The only person who likes change is a baby with a wet diaper."*
> *Mark Twain*

Traditionnally, we start by taking a look back at the "hot" topics discussed at the 2007 panel. At that time, we identified the following categories:

2.1 Features

It becomes increasingly difficult to answer the question: which features distinguish graphics recognition from general pattern recognition problems? This stems among other causes from the fact that we experience a deep convergence with the methods used in content-based image retrieval. The specificities of black-and-white images and of graphical information tend to become a detail in this broader context.

We pointed out that one interesting contribution would be to work on the characterization of various features for shape representation and recognition. In that sense, graphics recognition can be at the forefront of putting together large-scale repositories of features, so as to avoid the recurring appearance of "new" features which are actually minor variations on old ideas.

2.2 User Interaction

We pointed out that little work had been done, in the area of user interaction, on modeling the user, despite the fact that there are not many common features between a general, low-technicity user and a highly specialized technician, mastering the knowledge specific to a given application. This becomes a crucial problem when we produce applications aimed at the general public, but with complex user interaction which becomes accessible only after months of training.

2.3 Large-Scale Applications

A lot of discussions addressed the specific challenges of building large-scale applications, i.e. scaling from an academic problem to a really useful system. It was felt that this was not only an engineering problem but a cultural question and one of scaling our approaches. The need for a true policy on software development firmly rooted in the scientific achievements and tested on large datasets becomes crucial; we have to include the composition with reusable sofware and stress the building of production-quality code.

We also agreed that graphics recognition was still looking for its "killer application". One idea put forward was that a general sketching interface could be the answer, or maybe a combined sketching/retrieval/recognition system, making it possible to navigate in documentation by sketching simple examples of what is being searched for. It was pointed out that the GREC community does not seem to be very interested in dealing with digital documents such as PDF documents or web graphics. This may stem from a lack of good opportunities, or be an illustration of Mark Twain's quote cited above...

2.4 Performance Evaluation

We were disappointed by the low number of participants in our contests, even the more because this is an area where our community has often been showing

the path to the image analysis community at large. We agreed that beyond the contests, we need to have open-source, robust benchmarking tools available online, with a sufficient amount of ground-truthed data. The question remains to know whether we have access to benchmarking data covering all our needs.

We suggested to make the methods available as web services, so as to be able to test the limits of new methods as soon as they are developed. Another idea vented two years ago was to announce in advance a grand challenge for the commmunity to work on.

3 Topics Discussed During the GREC'09 Workshop

"All progress is precarious, and the solution of one problem
brings us face to face with another problem."
Martin Luther King Jr.

The workshop was a good opportunity to explore the variety of topics addressed nowadays by the graphics recognition community. The format of the workshop, the fact that it (still) gathers a significative part of the research teams active in the area, makes it a good observatory for the state of the art.

3.1 Technical / methodological aspects

Segmentation of graphics document is the historical theme of GREC, maybe also the only one for which you can be reasonably confident that you will see the most advanced results at GREC! We had several nice presentations at GREC'09; some of them dealt with *vectorization*, and especially the problem of robust arc detection, as well as specific ad hoc improvements, including domain knowledge, for various applications. Still, a very interesting question was raised during the workshop: why do we insist on vectorizing our raw data? Is it just something we have inherited from the "David Marr paradigm" or is it really necessary for the problems we have to deal with? Another question was whether it still makes sense to work mostly on binary images, whereas our raw data are often grey-level or color images.

Another typical segmentation topic of our community is *text-graphics separation*. We had several presentations dealing directly or indirectly with this question, including some nice applications such as the analysis of business cards on mobile phones. Somebody pointed out in a very interesting way that we improve our text-graphics separation methods because we have changed our focus: we now do text detection, not image segmentation!

Symbol recognition and spotting has also been a very active domain for many years already. Several papers addressed that area and questions asked included the way of dealing with complex symbols made of simpler symbols (we probably do not have robust approaches to that yet), the nature and the use of the available context (which is often difficult to capture), and the recurring search for genericity, which may be a kind of unreachable "Holy Grail" as it makes much more sense to build ad hoc systems for specific application areas.

Due to the very nature of the data we have to deal with, ***structural pattern recognition*** approaches are often a straightforward choice for many problems we deal with. This workshop was no exception and we saw a number of contributions using structural signatures, numeric signatures containing embedded graphs with the usual questions about graph edit distances, graph classification problems, adjacency grammars used in sketching applications, spatial relations represented in various ways. Using graphs necessarily leads to the general problems of graph matching, sub-graph matching, and the numerous embeddings and/or heuristics which aim at making these operations more efficient and applicable to large-scale problems. Note that these are not problems specific to graphics recognition.

It was pointed out that despite the known limitations of structural methods, they will remain useful for our problems for many reasons, including the structural and spatial nature of a lot of information we have to deal with, and the fact that we often do not have sufficient learning data to be able to perform statistical learning of our recognition methods.

Any recognition problem, be it solved by structural or statistical methods, has to deal with a representation of the shapes to be recognized through appropriate ***features***. With respect to our 2007 conclusions, we must acknowledge that at GREC'09, we saw some very nice new results on features, signatures and descriptors used in graphics recognition problems. But the question pointed out two years ago remains very much valid, maybe even more so than two years ago: which features distinguish graphics recognition from general pattern recognition problems?

This question is also at the core of work on *content-based indexing*, where we have some exciting results on our graphics-specific problems, but no real specific methodology, compared to the large area of information indexing and spotting in image databases.

Finally, our community remains active in ***performance evaluation and contests***, but we still have problems gathering participants, as pointed out two years ago. On the positive side, our GREC databases are used as reference in many papers. We still need to foster discussions about a broader policy of sharing data used in our publications.

3.2   Applications

We had the opportunity to see a lot of applications during these workshop days. Here is an overview of the variety encountered:

- Engineering drawings, architectural drawings, etc. do not seem to attract a lot of interest these days.
- Specific diagrams or notations such as chemical diagrams or music notation were more present; maybe this is a trend to move to very specialized areas, far from the main trend where a lot of effort has already been done (including with manual labor).
- There is a lot of interest in historical archives, i.e. legacy documents, often drawn and/or written by hand, with the purpose of archival, indexing and retrieval tasks.

– Ad hoc tasks such as identification and recognition of tables are addressed more easily than full-fledged, large-scale applications.
– The interest for sketching interfaces, already largely present two years ago, was confirmed at this workshop.

4   Some Hot Topics

"It's a long way to Tipperary, It's a long way to go.
It's a long way to Tipperary, To the sweetest girl I know!"
British music hall and marching song by Jack Judge

During the closing panel, we discussed several possible hot topics for the coming years.

A first item was the recurring wish for methods capable of efficiently combining structural and statistical methods. Of course, this is not specific to graphics recognition, but as said before, the very structural and spatial nature of the information we work with makes structural methods quite natural in the community. Their efficient integration into methods which also take full advantage of statistical learning and classification is certainly the right path to take.

As much as it was two years earlier, the need for the development of large-scale applications remains a strong incentive. We need toolboxes of robust document image processing algorithms. We need to make code and test data available. But we often end up with the dilemma of the cost we are ready to pay, as academic researchers, in order to develop professional-quality code.

Questions were also raised about the usefulness of the contest model for performance evaluation. Wouldn't it be desirable to stabilize noise models and evaluation metrics, make test databases available, maybe even have a consortium in charge of maintaining them and delivering the service of performance evaluation throughout the year.

Historical documents seem to become a major issue, but this is not specific to GREC and we even wondered whether it was the right place to deal with the issues which had been presented. The problems to be solved go all the way from image processing for restoration purposes to large-scale indexing and retrieval, based on the right features and descriptors, computed both from images, graphics and text.

5   Food for Thought

Will the last person to leave graphics recognition
please turn off the lights?

Let us conclude this report by mentioning some broader questions which were discussed at the panel:

– Is there life outside Google? Services such as those delivered by Google and other major industry players have drastically changed the way we deal with digitized information. How can we be reasonaly confident that we will not discover one day that we discuss at GREC "new" research trends which are already out there, available as a web service by one of these players? Said in other words, how can an academic community interested in graphics recognition have a real impact on large-scale applications for collections of historical documents, for instance?
– Can we define a Grand Challenge? In other fields (speech recognition, autonomous vehicles, etc.) the scientific community has worked towards solving a Grand Challenge, involving cooperation between different teams. This has been a driving force towards progress in these fields. How could this kind of higher goals be set in an area like graphics recognition? Would it make sense for it to be specific to graphics recognition, or should it be a broader problem?
– What is the value to end-users or to customers? More precisely, what are the difficult questions for today's users (i.e. those who use information), and what partial answers do we have to these questions? Are we willing and ready to consolidate these answers to bring real value, so as to have a real impact?

Of course, these questions look more like open and general problems than like a conclusion, but maybe the best way of concluding a panel is actually to leave the audience with some food for future thought...

References

1. Tombre, K.: Is Graphics Recognition an Unidentified Scientific Object? In Liu, W., Lladós, J., Ogier, J.M., eds.: Graphics Recognition: Recent Advances and New Opportunities. Selected papers from GREC'2007, 7th International Workshop on Graphics Recognition, Curitiba, Brazil, September 20-21, 2007. Volume 5046 of Lecture Notes in Computer Science. Springer-Verlag (2008) 329–334 Report from final panel discussion at GREC'07.