Is Graphics Recognition an Unidentified Scientific Object?

Karl Tombre

LORIA-INRIA, B.P. 239, 54506 Vandœuvre-lès-Nancy CEDEX, France
`Karl.Tombre@loria.fr`

Abstract. This paper summarizes the presentation and discussions at the panel session held at the conclusion of the GREC'07 workshop. After making a short review of where the graphics recognition stands, we raise some questions (hopefully) of interest for the future of this community.

1 Introduction: A Short History of Graphics Recognition as an Identified Community

As the name of a scientific community and of a workshop, "Graphics Recognition" has not been used for such a long time. In 1988, the "Structural and Syntactical Pattern Recognition" workshop was organized in Pont-à-Mousson, France, by Prof. Roger Mohr. At that event, such a large number of papers were presented on document image analysis applications that it was decided that the next workshop, held in New Jersey in 1990 and organized by Dr. Henry Baird, would be completely focused on this field. Revised versions of the papers presented at the SSPR'90 workshop were published in what was probably one of the first scientific volumes focused on document image analysis [1]. At this workshop, it was discussed and decided to create a new series of international conferences, the ICDAR series whose first instance was held in 1991 in Saint-Malo, France; this opened the way to the visibility of the document analysis and recognition community as such—although work on these topics had of course existed for a long time before that, within the general pattern recognition community.

Within this general positioning of the community as such, Prof. Rangachar Kasturi moved in 1992 to give a new start to a Technical Committee of the International Association for Pattern Recognition, namely TC10 on "Line Drawing Interpretation", by focusing it on the subfield of document analysis and recognition devoted to graphics-rich documents and the specific problems raised by these documents (raster-to-graphics conversion, text-graphics separation, symbol recognition...) To emphasize the new focus, TC10 was renamed as the technical committee for *Graphics Recognition*. This gave birth in 1995 to the GREC series of workshops, which themselves have led to the publication of reference LNCS volumes for the community [2,3,4,5,6,7]. So one may say that it is only in the last 15 years that there has been a set of researchers identifying themselves as the "Graphics Recognition Community".

This paper, stemming from a panel discussion held at the end of the GREC'07 workshop in Curitiba, Brazil, tries to take a step back and look at the field, its achievements, its remaining challenges, and the homogeneity of its composition.

2   Looking Back: What Did We Conclude Two Years Ago?

GREC'05 marked the 10th anniversary of the workshop and was an opportunity for reviewing achievements, topics which we had drifted away from, others which were still relevant, and emerging new themes.

Our main conclusions at that time were the following [8]:

– There are still open problems in *building complete systems*, so scientists should not drift away too quickly from this question, despite the successes which have been achieved;
– *Symbol recognition* matures but there are remaining open problems, including the recognition of non-segmentable symbols, the issue of scalability, and the recognition of complex symbols made of the assembly of smaller symbols;
– *Cultural heritage documents* emerge as a new theme;
– The issue of *symbol spotting*, i.e. localizing a symbol without necessarily recognizing it explicitly, is getting more focus, including the need to take into account relevance feedback from the user;
– The community has failed to gather around a common base of *software*, largely due to the "not-invented-here" syndrome.

Some hot topics were identified:

– To achieve progress in close-to-optimal, automatic, and non-contextual *vectorization* on black-and-white images, progress will have to include the processing of gray-level images, sub-pixel precision in the segmentation tools, better curve segmentation algorithms, and a seamless integration of user input and of contextual knowledge;
– The problems to be solved for complete document analysis include low-level questions such as the digitizing resolution, and the analysis of digital documents with little or no structure;
– Performance analysis campaigns have been a success, but the contests' use of degradation models can be controversial, there are ususally too few participants, and we need to perpetuate the access to data and evaluation tools, also outside the contests.

It is interesting to look at the community's most recent achievement with these questions in mind, to see how fast and in which direction we are moving.

3   Some Topics Discussed during the GREC'07 Workshop

Let us now try to review a few topics addressed by our community these last years, and discuss to which extent they characterize graphics recognition as a field.

3.1 Features

A lot of work has dealt with defining and using appropriate features for recognition tasks, especially in graphics recognition. At GREC'07, work was presented on blurred shape models, ridgelets, graph representations, region-based signatures, to name a few. We also got a presentation of an interesting attempt at characterizing the performance and usability of various shape descriptors. In recognition as well as in information spotting applications, it is indeed necessary to work on the most appropriate features, and also on the right feature combination methods for optimal recognition.

But it becomes increasingly difficult to answer the question: which features distinguish graphics recognition from general pattern recognition problems. Comparing our contributions with those in content-based image retrieval, for instance, show a real convergence, where the fact that we are dealing with black-and-white, graphical information tends to become a detail.

One contribution our community can make, as it deals with a subset of all possible imaging applications, is to work on the characterization of various features for shape representation and recognition, thus contributing to building up a professionnal repository of features with their properties, and maybe avoiding the recurring appearance of "new" features which are only minor variations on old themes.

3.2 User Interaction

It is increasingly necessary to design analysis and recognition methods which do not work in stand-alone mode, but take into account the user's interaction, so as to be able to perform incremental learning, relevance feedback in recognition and retrieval applications, interactive recognition in sketching mode, etc. But little work has been done on modelling the user, who is mostly considered as some kind of ill-defined, external entity.

The fact is that there is nothing in common between a "vanilla plain" user who may be your uncle or grandma, browsing a collection of images and giving relevance feedback without really knowing anything about the application, and a highly specialized user able to input syntactical rules to represent the knowledge in a specific document analysis application.

If the purpose is indeed to build a highly specialized system, this may not be a problem, but when the application is potentially very general whereas the user interaction paradigm requires the user to have a PhD in pattern recognition or to have trained for months, there is a contradiction in the whole setup which limits the applicability of the method.

3.3 Building Large-Scale Systems

A number of large-scale applications are being dealt with, especially for sketching, and for the analysis, characterization and indexing of large databases of historical documents. However, in order to go beyond the proof-of-concept system, we need to put more emphasis on software, including reusable software

which has not been developed in our own group. The community still suffers of the "not invented here" syndrome, despite the availability of a significant number of larger image analysis platforms, some of them specialized in document analysis, others more generic. It will be difficult to have a significant impact without either building oneself production-quality code and putting in enough human resources for that purpose, or using others' code and pushing one's students to use existing code.

It was mentioned that graphics recognition was still looking for its "killer application" and some participants felt that a general sketching interface could be such an application. But in that case, efforts must be coordinated, within academia or in partnership with companies.

The dream of building completely automated systems for converting drawings, maps and diagrams into high-level representations seems to have vanished, as the methods we design reach their limits at a level where there is still a lot of user editing to be done. But as one participant mentioned, there is still a very interesting opportunity to build combined sketching/retrieval/recognition systems, making it possible to navigate in a large document base by sketching simple examples of what is being searched for.

Suprisingly, whereas there are a lot of high-interest applications in dealing with digital documents (e.g. PDF documents or web graphics, in which to search for and recognize various entities), a workshop like GREC seems not to be deemed the right place to present this, as we have not seen any work in this area.

3.4   Performance Evaluation and Contests

This is an area where the field of document analysis in general, and the subfield of graphics recognition in particular, have often been showing the path to the whole image analysis community. This may stem from the fact that the data and problems we work on are more easily circumscribed. But we have also been proactive in organizing contests, gathering ground-truthed data, making available performance evaluation metrics and tools.

But there are still concerns. We have few participants in the contests, despite the very hard work by the organizers. This is a little bit disappointing. Our community has also started referring to the data and evaluation tools continuously, in day-to-day work without any contest. This effort must be continued, but for that we need to have open-source, robust benchmarking tools available online, with a sufficient amount of ground-truthed data. How do we capitalize on contributions by various teams?

Another question is to assess whether we have benchmarking data covering all the needs. Finally, it is still necessary to ask ourselves what we want to evaluate, so that our benchmarks are not purely academic but that they model real-world challenges.

Some future directions in that area could be to challenge our existing methods by making them available as a web service and letting colleagues "attack" them to test their limits. Another idea would be to announce in advance a grand

challenge for the commmunity to work on, from one workshop to the other, for instance.

4   Is Graphics Recognition Still an Appropriate Scientific Area?

One final and somewhat controversial topic discussed at length in our closing panel was whether it still makes sense to gather a community around the theme of "Graphics Recognition". In the first GREC workshops, most people contributed with pure graphics recognition problems, such as vectorization or text-graphics segmentation. There are very few new contributions to these topics, as methods have reached maturity and are assimilated by people as state of the art.

Now, we see a variety of work related to topics such as visual languages, document image layout analysis, shape descriptors, information retrieval and information spotting, biometrics, etc. But in each case, we only have a subset of activities related to each of these areas. Is the subset representative of some specificities, or does it just happen to be the contributions to these topics of the groups used to define themselves as belonging to the "graphics recognition community" (and used to attend GREC)? We should have the courage to ask ourselves these questions...

5   Conclusion

The graphics recognition community has contributed significantly to the field of pattern recognition, setting the path for others in some areas, such as performance evaluation. It is now at crossroads where it has to ask itself whether its scientific interests need to be kept together or if it is time to reconsider the frontiers and join the most appropriate scientific communities.

Looking back at history, the concept of "Graphics Recognition" has been in use for around 15 years only, at least to define a scientific community. Science is a living phenomenon, where one of the biggest dangers is to bury oneself (and one's students) in a narrow pit where one looses the view of evolutions and new advances in neighboring areas. I have had the pleasure of being strongly involved in the adventure of this community for these 15 years, and am personally confident that the groups now active in graphics recognition have the ability to make the most appropriate moves. I would recommend to make them without any tabu on which structures to keep, which structures to throw away, and which structures to change.

References

1. Baird, H.S., Bunke, H., Yamamoto, K. (eds.): Structured Document Image Analysis. Springer, Heidelberg (1992)
2. Kasturi, R., Tombre, K. (eds.): Graphics Recognition 1995. LNCS, vol. 1072. Springer, Heidelberg (1996)

3. Chhabra, A.K., Tombre, K. (eds.): GREC 1997. LNCS, vol. 1389. Springer, Heidelberg (1998)
4. Chhabra, A.K., Dori, D. (eds.): GREC 1999. LNCS, vol. 1941. Springer, Heidelberg (2000)
5. Blostein, D., Kwon, Y.-B. (eds.): GREC 2001. LNCS, vol. 2390. Springer, Heidelberg (2002)
6. Lladós, J., Kwon, Y.-B. (eds.): GREC 2003. LNCS, vol. 3088. Springer, Heidelberg (2004)
7. Liu, W., Lladós, J. (eds.): GREC 2005. LNCS, vol. 3926. Springer, Heidelberg (2006)
8. Tombre, K.: Graphics Recognition: The Last Ten Years and the Next Ten Years. In: Liu, W., Lladós, J. (eds.) GREC 2005. LNCS, vol. 3926, pp. 422–426. Springer, Heidelberg (2006)