Commercial Exploitation of OCR and Document Analysis Systems Technology: Market and User Requirements

Stefan Knerr

A2iA
3 Rue de l'Arrivee
Tour CIT, BP 59
75749 Paris Cedex 15
France

Tin Kam Ho

Bell Laboratories
Lucent Technologies
700 Mountain Avenue, 2C-425
Murray Hill, NJ 07974
USA

Seven researchers from four countries participated at the discussion of this working group: Teruo Akiyama (NTT, Japan), Edward Cohen (OPEX, USA), Alexander Filatov (Parascript, Russia), Tin Kam Ho (Bell Labs, USA), Stefan Knerr (A2iA, France), George Nagy (RPI, USA), Yashiro Okada (Mitsubishi, Japan). Most participants are from the industrial world.

Several questions at the beginning opened the discussion:

- What are the potential markets and products?
- Why isn't there more money made with OCR and DAS?
- Has OCR/DAS missed out?
- What do the customers wait for and what do they really want?
- What is needed to build successful OCR/DAS systems?

1 Markets

Large volume document processing and data entry applications have considerable market opportunities. Among these applications are: postal address recognition for mail sorting, payment document processing (checks, giros, credit cars slips...), and forms processing (e.g. for insurances, tax forms, census). Here, imaging solutions are either already well established or currently making their way into the market. These solutions benefit from (i) small or limited vocabularies, possibly from (ii) constraints or standards on the layout of the documents, and demand for the recognition of machine print and handwriting. The large document volumes often justify the important investments in scanning and recognition technology customers have to make. Sometimes, the products even have to undergo some customization in order to provide the

best possible recognition performances. It is clear that in a relatively near future, OCR/DAS technology will have widely replaced manual data entry.

Another important market is the paperless office which demands for automated document input including OCRing the, mostly printed, text. Easy document access and processing (copy, cut and paste, etc.), better customer services, and, in the future, lower cost for archiving are some of the good reasons for going paperless. The fax market is also well established and has been one of the first markets attracting OCR technology providers. Along with the Internet gaining in popularity, digital libraries will also have a serious need for high quality OCR. Even one or two recognition errors per page will sum up to a huge amount of correction work when several million pages have to be scanned into the computers. Talking about the Internet: there is more and more interest in providing OCR/DAS services over the Net. Some see the future in this direction, making the acquisition of expensive and ever-changing recognition hardware and software obsolete. All it takes is a scanner and a Internet access.

In addition to off-line OCR/DAS systems, on-line handwriting recognition for pen-based input devices is slowly but surely emerging as a mature technology for certain applications such as mobile computing and data entry. This is particularly interesting for markets such as China which lack the standards for keyboard input.

Several applications already have useful products on the market. For others, it seems that the technology is there, but needs to be put together. Perhaps an "OCR developer's tool kit" would help to get the initial development cost down by providing modules easy to assemble for different solutions. One of the key points which helps OCR/DAS systems to become real products is the availability of relatively inexpensive CPU power. However, most commercial products are still targeted to large, industrial customers. Whether there is a household use of OCR/DAS systems other than for fax or for small markets, such as reading for the blind, still remains unclear.

2 What do the customers wait for and what do they really want?

OCR research is basically as old as computer engineering itself, and there were many early promises as to the performances which would be obtained by OCR/DAS systems. Most of these promises were not realistic. As a consequence, potential customers became somewhat suspicious regarding performance quotations and product announcements. Today, many potential customers show a wait-and-see attitude, and do not want to invest in first-generation solutions. This does not make things easier for the vendors of OCR/DAS systems. High user expectations in terms of read rates, error rates, and recognition speed remain one of the key issues of the domain. The general reference (and ultimate challenge) is still often enough the human accuracy or recognition performance.

Today, more and more recognition engines are offered on the market. While this increases the competition, which is usually good for the market, most users also want complete, integrated solutions, not only recognition engines. However, in order to meet the performance expectations of these users, the solution has to be customized, or adapted to the particular needs of the user. This of course increases the cost of the solution, and asks for user involvement during the development of the solution. Only large customers have the time, skills, and resources to test and understand the technology, and to co-develop working systems. Small customers tend to focus on their own particular application, eager to satisfy their goal with limited expenses.

3 What is needed to build successful OCR/DAS systems?

There is a clear need to build complete, working solutions. Don't just solve toy problems! It seems that most of the necessary technology is around. It needs to be pulled together, and the issue is possibly more one of large, complex systems than of algorithms.

All participants of this work group agreed on the importance of publicly available standard databases (images, dictionaries, ...) for the development of OCR/DAS systems and to conduct tests and performance comparisons. The latter are necessary in order to show potential customers what can be achieved in their domains and to help them find the best system on the market for their particular purposes. Here is also a need for concrete and realistic cost models to show potential savings for the customer. Along with the data bases, general guidelines for tests and performance comparisons should be defined. The work group participants expressed their hope that organizations such as the National Institute of Standards and Technology (NIST) and the Information Science Research Institute (ISRI) would take an active role in providing such databases and conduct or assist in conducting tests.

In several countries, the Postal Services have played an important role in pushing address recognition and therefore OCR and ICR as well as DAS. They provided data bases and financial support. Consequently, postal address recognition is among the most successful OCR/DAS applications. Similar situations in other domains would probably benefit researchers, developers and customers alike, for instance in the payment document domain.

There is a need for more interaction between researchers and developers on one side and the market on the other side. Some of the issues discussed in the OCR/DAS community may be false problems which could be eliminated or simplified by redesigning documents in the future (at least partly, and for large scale applications): more standardized layouts, no fancy background pictures, more redundancy in the text, etc. For instance, for Census 2000 the forms were designed to be easy to read.