PARASCRIPT®

Strategic White Paper

Automating Handwriting Recognition – Is NOW the Time?

Takeaway

- Learn key differences between OCR (machine print) and ICR (handprint / handwriting) technologies and their adoption
- Understand that ICR is under-adopted, and presents an opportunity to lower cost and gain a competitive advantage
- Realize that ICR can be applied in most instances where context is available
- Identify high-level characteristics to determine if automation is right for your organization

Summary

This paper provides a brief overview of automated handwriting recognition, its current state of adoption, and key points to understand and evaluate this software.

What's in this Paper?

6273 Monarch Park Place, Longmont, CO 80503 USA
T: 303.381.3100 | toll free: 888.225.0169
F: 303.381.3101 | info@parascript.com

Introduction

Businesses and organizations have grown and prospered by converting from manual to automated processes to save time and money and increase productivity—typically by reducing reliance on paper. Yet despite tremendous gains in process efficiency, paper is still a crucial medium for bridging external and internal processes; which could be as simple as taking a new customer application and entering it into a computer.

With rising labor costs and competitive pressures for faster customer service, automation of data extraction from paper is becoming crucial to businesses, organizations and government agencies that process huge volumes of forms, applications, checks and other documents.

These factors have led to the continued growth of document capture and electronic content management (ECM) solutions, as well as recognition solutions, which are critical to automating the extraction of content from paper forms and documents. This software is used across several industries to read and digitally convert machine- and hand-written content, and even validate signature authenticity. The requirement to efficiently process this data is increasing exponentially.

Key Terms – OCR & ICR

OCR – Optical Character Recognition is the technology that does electronic conversion of scanned images of text from a paper source into computer-encoded text. Digitized data can then be stored, displayed, analyzed, searched and otherwise manipulated. However, OCR is constrained to machine print text only, which significantly limits its application.

ICR - Intelligent Character Recognition is a more advanced technology that reads handwriting, which may include constrained and unconstrained handprint, in addition to cursive script. This is a significantly more complicated technical challenge since handwriting is unique to each individual. The latest generation of ICR software greatly broadens the technology's reach into more diverse applications, offering benefits of improved cost savings, productivity and accuracy.

With rising labor costs and competitive pressures for faster customer service, automation of data extraction from paper is becoming crucial to businesses.

Automated Handwriting Recognition

It is important to understand the differences between OCR and ICR to appreciate the operations and value of these related technologies.

OCR – Optical Character Recognition
The United States Postal Service has used OCR scanners to sort mail since the 1960s. OCR software translates printed characters into electronic text, word processing files and other digital file formats. The technology is becoming ubiquitous in our society, especially with the growth in distributed scanning and mobile platforms.

Value:
Electronic documents can be stored with significantly reduced storage requirements than hard copy materials. Even more important in today's business environment, digitized data can be extracted, sorted and efficiently distributed for multiple uses. Since printed characters and fonts are standardized, OCR can easily and accurately identify and convert data.

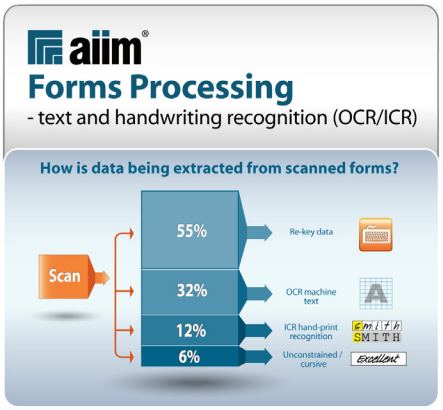
ICR – Intelligent Character Recognition
Intelligent Character Recognition (ICR) is a significant step forward compared to OCR technology. The ability to recognize hand-printed characters significantly broadens the range of applications that may benefit from an ICR solution. Hand-printed characters are created by humans, so understanding and interpreting the patterns of human writing is far more complicated than converting simple machine print, because no two people ever write identical characters. With the latest technology advancements, ICR can recognize both unconstrained handprint and cursive handwriting.

Value:
Despite the value offered by these technologies, adoption rates are still relatively low. As a result, there is an opportunity to gain an advantage over competition. By automating and improving the time it takes to process inbound forms, organizations can provide better customer service and lower their costs.

ICR – Past, Present & Future

ICR technology has been in use for nearly 20 years, but it has not been widely adopted outside of very specific applications such as check and mail processing. Despite its capabilities, relatively few organizations or service providers leverage this technology for processing forms.

AIIM (Association for Information and Image Management), a global community of information professionals wanted to find out why. In March 2012 the organization conducted a study with Parascript to better understand ICR adoption. The study of 255 corporate and government respondents involved in document management provided these results:



AIIM Forms Processing White Paper
User experiences of text and handwriting recognition (OCR/ICR).

Click Here to Download or Visit
http://info.parascript.com/aiim-forms-processing-white-paper

- 88% of organizations scan their forms, but only 32% perform automated OCR on machine print
- 12% automate the recognition of constrained hand-printed fields
- 6% automate recognition of unconstrained or handwritten fields

Despite the low use of ICR, respondents replied that handwriting was prominent on many of their forms. They also reported an expected 35% improvement in productivity with automation tools.

Why Low Adoption Rates?

Based on this study, adoption rates for ICR are surprisingly low considering its benefits in accuracy, cost and improved productivity and reliability. The study surfaced some reasons:

1. Decision making for this type of process software for forms and documents is often made on the local level, not by a centralized IT department. Middle managers may be reticent to recommend such a substantial process change.

2. OCR technology is not widely adopted (32% of study respondents) and OCR adoption is generally a pre-requisite to implementing ICR.

3. Many executives and department heads simply have not evaluated ICR recently and are basing their decision, or non-decision on outdated information and results.

Another observation is that there is lack of awareness in the market as to how it ICR works and its current capabilities.

How ICR Works

There are infinite numbers of variations encountered in human handwriting. The most advanced ICR software is designed to replicate the human recognition process - programmed to think like the human brain. Similar to the way a person reads a document; the recognition engine creates a hypothesis about a word or number series, evaluates it against a database of possible answers (name, address, date of birth, social security number, etc.), and proves the answer that exists in the database. It views handwriting in a form field as an entire word or phase, not individual characters. This “holistic approach” for automated handwriting recognition combines and integrates many components:

- Highly advanced mathematical algorithms
- Rules that regulate the interpretation process
- Data modeling tools
- Image interpretation using a special description language
- Content validation within its context

ICR Evaluation and Adoption Considerations

There are several reasons to consider implementing ICR technology. Those organizations that already have massive form and document volume will need to evaluate the impact of reduced costs and increased productivity if they expect continued growth. With low market adoption, there is a significant opportunity to gain a competitive advantage over others in the market segment. And most importantly, by automating and reducing the time it takes to process inbound forms, organizations and companies can provide a much better customer service experience.

In order to develop an ROI model, there are three essential points of evaluation:

1. Understand what documents may or may not qualify for recognition
2. How these documents will enter your process
3. How exceptions will be handled

Consider the answers to these questions to properly set expectations:

- What is the current volume and thruput of forms or document processing?
- Is the document classifiable and/or does it appear with regularity?
- Do the documents have standardized content or formats that exist in your databases, or commonly available databases (names, addresses, phone numbers, product names, etc.)?
- How will exceptions be handled (it is likely that large teams of people are employed to perform manual entry)?
- Will you encourage distributed capture?

Conclusion

Now, most certainly is the time to evaluate and consider automated handwriting recognition. Businesses and governments whose daily activities include processing a large volume of forms and documents can and should consider evaluating handwriting recognition software. This assessment is especially critical if your company receives payments in-house, applications or any other form or document that requires intensive processing. Not only do companies stand to gain internal process improvement, and all that that implies, but will also generate a higher level of customer experience and satisfaction. An investment in automated recognition technology can deliver to both top and bottom line considerations.

About Parascript

Parascript is a leading developer of cursive, handprint, and machine print recognition solutions. Leveraging digital image analysis and advanced pattern recognition, its software enables business automation in forms processing, postal and financial automation, and fraud prevention; and supports cancer screening in medical imaging. Parascript's award-winning technology draws on a proven 15+ year track record and processes billions of document images annually. Fortune 500 companies, postal operators (including the U.S. Postal Service), major government and financial institutions rely on Parascript products, which are distributed through its OEM and Value Added Reseller networks, including partners such as: IBM, EMC, Bell and Howell, Fiserv, Selex Elsag, Lockheed Martin, NCR, Siemens, and Burroughs. Visit Parascript online at http://www.parascript.com.

6273 Monarch Park Place, Longmont, CO 80503 USA
T: 303.381.3100 I toll free: 888.225.0169
F: 303.381.3101 I info@parascript.com