**On-line handwriting recognition**

**Introduction**

Although the problem of handwriting recognition has been considered for more than 30 years [http://www.ocr.eu/], there are still many unsolved issues, especially in the task of unconstrained recognition. This domain is traditionally divided into on-line and off-line recognition. In on-line recognition a time ordered sequence of co-ordinates is captured, while only the image is available in the off-line mode.

The wide spread use of pen-based hand held devices such as PDAs, smartphones, and tablet-PC, increases the demand for high performance on-line handwritten recognition systems. This man machine interface method is an alternative for the traditional keyboard with the advantages of being more easy, friendly, and natural. This technology has great potential markets in friendly learning environments, business applications and more.

**1-Latin Script Status**

***State of the art:***

On-line handwriting recognition for Latin script is a rich and huge field in both research and commercial products domains.

Researchers, research groups, and research centers working in this field are spread overall the world.

Datasets for training and comparison of results are found easily.

A lot of magazines ,journals , and conferences can be found in this area (They are not restricted for Latin script but Latin script is the major script) .

Many companies such as Vision Objects, A2IA, ABBYY, [Readiris](http://en.wikipedia.org/wiki/Readiris), etc. are working in this field.

Very high performance commercially products can be found in many applications. Users can enjoy entering data on handheld devices using pen instead of the keyboard.

***Future trends:***

Improving the performance of the following:

a) Unconstrained character recognition

b) Writer independent recognition

c) Limited computational and memory resources (specially for handheld devices).



**2-Applications and market priorities**

* ***Numeral recognition:***
* ***Isolated character recognition:***
* ***Lexicon-based cursive recognition:***
* ***constrained cursive recognition***:
* ***unconstrained cursive recognition***
* ***Writer Identification and verification***
* ***Signatures Verification***

**Devices**

* **Tablet PCs**
* **SmartPhones**
* **PDAs**

**3- Applications Performance (Latin and Arabic)**

Products for Latin script have very high performance. The following is an example for a commercial product:

 **Parascript Inc. (**<http://www.parascript.com/>**)**

Parascript began life as the company ParaGraph International, which developed the handwriting recognition features of the Apple Newton--the first commercially available natural handwriting recognition engine. Following the Newton, their next handwriting recognition application was CalliGrapher. CalliGrapher technologies were sold to to Microsoft. Today, the CalliGrapher technology forms a part of Transcriber, the handwriting recognition software included in all Pocket PCs and Tablet PCs.

The Parascript Pen&Internet division is developing the next generation of handwriting recognition software, called riteScript. Currently, though, the division's main product is riteMail, an application for creating, storing and exchanging handwritten notes and drawings on a computing device (known as electronic ink) and sending them to any email address. riteMail is currently in beta, and includes riteShape, a technology that recognizes and automatically perfects common shapes like circles, rectangles, triangles, and arrows.
Users can try out riteMail, as well as the riteShape and riteScript recognition engines, online at: <http://www.ritemail.net/>

Parascript defines natural handwriting recognition (NHR) as the ability for a computer to recognize and convert to text any freehand writing, cursive or print, unconstrained by any boxes or combs. Most commercially available natural handwriting software is based on ParaGraph or Parascript technology.

Parascript Pen&Internet NHR does not look at characters; it looks at words and phrases. CalliGrapher (and now Transcriber), as well as other Parascript applications such as TRS, uses the XR elements (64 designed elements can be combined to form any character or set of characters in cursive handwriting) to break cursive words down into a linear series of elements. Each series of elements then represents a word or phrase that can be matched against a database of expected or common words or phrases.

riteScript, however, does not use the XR technology. riteScript's new NHR is much more dependent on contextual information than the XR-based system. As each word is processed, the NHR engine begins comparing it to multiple databases of known or expected words and phrases--a process known as lexical support. Since the interpretation of the strokes is much broader under the new riteScript NHR engine, It can achieve much better results with good lexical support than the XR system. The new technology has greater much lexical support than older recognition engines, including multiple lexical sources that operate in parallel.

Key features of the riteScript technology include:

Style independence: riteScript recognizes handwriting in connected "cursive", separate letter "print" and the mix of both styles. Users don't have to learn artificial shorthand symbols or change their writing styles

Writer independence: riteScript recognizes handwriting with high recognition rates without requiring users to train it on a lengthy sample text

Multi-lexical support: riteScript recognizes vocabulary and non-vocabulary words as well as arbitrary combinations of letters, digits and special symbols

Page location independence: Users can write anywhere on the page or writing surface and do not need to confine their writing to a restrictive baseline, boxes or combs.

Phrase and document awareness: riteScript provides automatic word and line segmentation and baseline detection in the handwritten text

riteScript is designed as a plug-in component for Web and wireless solutions that require handwriting recognition. Examples of applications are Web-based forms recognition and advanced processing of personal notes taken on pen-enabled handhelds, Webpads, intelligent pens and other devices.

ritePen® is an advanced handwriting recognition software for Microsoft Windows-based pen-enabled computers. Users of ritePen can write anywhere on their screen or other input surface and have their handwriting instantly converted to text for use in any Windows application, including Word, Excel, Outlook, and numerous others. ritePen is a seamless extension of normal writing because it accurately recognizes virtually any handwriting style, does not require learning or training, and allows you to write in whole sentences, while automatically segmenting your handwriting into words and lines.

**Arabic Products**:

Practically functional Arabic handwritten OCR systems are rare, and the product of Arabic Writer© form ImagiNet® can be selected as a representative one. The underlying methodology of this system is to train and deploy artificial Neural Networks to decide on the most likely character sequences corresponding to the dynamically sensed features sequences of curvature, with a preprocessing of short strokes corresponding to dots and diacritics. For more details on this system; the reader can visit:

<http://www.imaginet-software.com/index.aspx> (for xType and iScript products)

Other products exist, namely by VisionObjects, QuickScript and Sakhr.

The following table summarizes the main products for handwritten OCR which include Arabic or has the potential to include Arabic soon (e.g., ritepen).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| product | Applications | License | Languages | Performance | Platforms | Price | Notes |
| ritepen |  |  | Dutch, English, French, German, Italian, Portuguese, Russian and Spanish handwriting.(No Arabic Yet) | Recognizes with high accuracy unrestricted and continuous writing.Requires no training.Lists alternative word options and allows easy in-line corrections. | Tablet PC, Ultra-Mobile PC, "Netvertible", electronic whiteboard, pen or pen&touch tablet, and any other interactive pen input device |  |  |
| VisionObjects:Myscript Stylus (Lingo) |  (MyScript Lingo’s language packs are designed particularly for: [Form processing](http://www.visionobjects.com/products/software-development-kits/myscript-builder/myscript-form/) and [note taking](http://www.visionobjects.com/products/software-development-kits/myscript-builder/) applications that require highly accurate results [Embedded platforms](http://www.visionobjects.com/products/software-development-kits/myscript-builder-embedded/) using **natural cursive handwriting** as a text input method.The use of lexicons, data formats and language models enables the MyScript engine to recognize text in combs and boxes as well as in free written text |  | 26 languages including Arabic. |  | Devices running on Windows, MAC and Unix.Pocket PC. |  | **MyScript Lingo is a set of 26 language packs available with MyScript Builder or MyScript Builder Embedded software development kits. MyScript Lingo is designed particularly for form processing, note taking and other applications requiring cursive handwriting, including for embedded platforms.** |
| VisionObjects:Myscript Stylus (Letra) | MyScript Letra is specially designed for integration into embedded platforms using touch screen and stylus-based interfaces such as:* Smartphones,
* PDAs,
* GPS,
* Electronic Tablets,
* Gaming devices and so on.
 |  | 83 languages including Arabic |  |  |  | **MyScript Letra is a pack of resources available for MyScript Builder Embedded software development kit.MyScript Letra provides language specific sets of characters in order to recognize hand-printed and isolated characters in more than 80 languages.** |
| Quickscript |  |  | 26 languages including **Arabic.** |  | MAC only. |  |  |
| Imaginet |  |  | **Arabic** |  | HTC phone, smart phones, PDA's |  |  |
| Sakhr | Sakhr is a leader in Arabic handwriting recognition, online and offline. Sakhr’s online intelligent character recognition (ICR) recognizes Arabic cursive handwritten input through a normal pen with 85% word accuracy.Sakhr ICR runs on any Tablet PC using Windows XP. It can also be integrated with other handheld devices such as Palm, Pocket PC and other smartphones. Sakhr’s offline recognition technology is available for recognizing specific data fields on defined forms. |  | **Arabic** |  | Tablet PC, smart phones. |  |  |

**4-Applications required modules**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Isolated Numeral recognition | IsolatedCharacter recognition | LexiconCursive recognition | Restricted cursive recognition | Free cursive recognition | Writer Identification | Signatures Verfication |
| **Preprocessing** | **++** | **+** | **+++** | **+++** | **+++** | **+** | **+** |
| **Segmentation** |  |  | **+++** | **+++** | **+++** | **+** | **+** |
| **Feature extraction** | **++** | **++** | **+++** | **+++** | **+++** | **++** | **++** |
| **classification** | **+++** | **+++** | **+++** | **+++** | **+++** | **+++** | **+++** |
| **PostProcessing** | **+** | **+** | **+++** | **+++** | **+++** |  |  |
| **Lexicon** | **++** | **++** | **+++** | **++** | **++** |  |  |
| **Character LM** |  |  | **++** |  | **++** |  |  |
| **Word LM** |  |  |  |  | **++** |  |  |

The above table shows the relation between required modules and the applications

**5-Modules and the language resources**

The main modules that requires language resources are the language models and the classifiers. The sufficient amount of resources required for training such modules and for benchmarking are in section 6.

**6- Available language resources:**

The ADAB-database:

Please see:

<http://www.icdar2007.org/>

<http://www.ifn.ing.tu-bs.de/competition2007/>

<http://www.ifn.ing.tu-bs.de/cfp-icdar2009/>

[www.**cvc.uab.es**/icdar2009/papers/3725b383.pdf](http://www.cvc.uab.es/icdar2009/papers/3725b383.pdf)

The database ADAB (Arabic DAtaBase) was developed to advance the research and development of Arabic **online** handwritten text recognition systems. This database

is developed in a cooperation between the Institute for Communications Technology (IfN) and the Ecole Nationale d’Ing`enieurs de Sfax (ENIS), Research Group on Intelligent Machines (REGIM), Sfax, Tunisia.

The database in version **1.0** consists of 15158 Arabic words handwritten by more than 130 different writers, most of them selected from the narrower range of the l’Ecole Nationale d’Ing`enieurs de Sfax (ENIS). The text written is from 937 Tunisian town/village names. Special tools for the collection of the data and verification of the ground truth are developed. These tools give the possibilities to record the online written data, to save some writer information, to select the lexicon for the collection, and re-write and correct wrong written text. Ground truth was added to the text information automatically from the selected lexicon and verified manually.

The database in version **2.0** patch level 1e (v2.0p1e) consists of 32492 Arabic words handwritten by more than 1000 writers. The words written are 937 Tunisian town/village

names. Each writer filled one to five forms with preselected town/village names and the corresponding post code. Ground truth was added to the image data automatically

and verified manually.

The test datasets which are unknown to all participants were collected for the tests of the ICDAR 2007 competition. The words are from the same lexicon as those of IfN/ENIT-database and written by writers, who did not contribute to the data sets before. The test data is composed of about 10,000 Arabic names (City and Town names).

**Best Performances:**

The best achieved performance at the 2009 competition was obtained by the MDLSTM system, with 93.4% on set ***f*** (about 8500 names, collected in Tunisia, similar to the training data), and 82% on set ***s*** (about 1500 names collected in UAE).

The MDLSTM system is developed by Alex Graves from Techische Universitat Munchen, Munchen, Germany. This multilingual handwriting recognition system is based on a hierarchy of multidimensional recurrent neural networks [http://www.idsia.ch/~juergen/nips2009.pdf]. It can accept either on-line or off-line handwriting data, and in both cases works directly on the raw input without any preprocessing or feature extraction. It uses the multidimensional Long Short-Term Memory network architecture, an extension of Long Short-Term Memory to data with more than one spatio-temporal dimension. The basic structure of the system, including the hidden layer architecture and the hierarchical subsampling method is described in [http://www.idsia.ch/~juergen/nips2009.pdf]: available online.

The second best system obtained about 89.9% and 77.7% for the two sets mentioned above. The system is called Ai2A.

The A2iA Arab-Reader system was submitted by Fares Menasri and Christopher Kermorvant (A2iA SA, France), Anne-Laure Bianne (A2iA SA and Telecom ParisTech,

France), and Laurence Likforman-Sulem (Telecom Paris-Tech, France). This system is a combination of two different word recognizers, both based on HMM. The first one

is a Hybrid HMM/NN with grapheme segmentation: Please see: <http://portal.acm.org/citation.cfm?id=1006603>.

It is mainly based on the standard A2iA word recognizer for Latin script, with several adaptations for Arabic script. The second one is a Gaussian mixture HMM based on HTK, with sliding windows (no explicit pre-segmentation). The computation of features was greatly inspired by Al-Hajj works on geometric features for Arabic recognition. The results of the two previous word recognition systems are combined so as to compute the final answer:

http://alqlmlibrary.org/LocalisationDocument/O/Off-LineArabicCharacterRecognitionAReview.pdf

**7-Sufficient required resources:**

 For a specific application, such as recognizing city names (ADAB database), with a lexicon of about 1000 words, it was sufficient to collect data from 1000 writers, with a total of about 35,000 words (average of 35 words by each writer). If we look at the Part of Arabic Words (PAW) frequency, we find that it was also about 35,000 in the whole test set. This shows that it was sufficient to train the system with an average of one PAW occurrence. However, there is no analysis of the training data coverage of the different PAWs. We think that synthesizing balanced coverage of the PAWs would give better results.

As for the benchmarking data, the lexicon of 1000 words corresponded to a total set of 10,000 instances, with an average of 10 occurrences for each word in the lexicon.

This competition benchmark information can be taken as a good starting point for developing more benchmarks with different lexicons for other domains.

**8-A gap analysis between available and required:**

Data Sets:

The only standard database available is the ADAB online Arabic handwritten one used in the ICDAR 2007 and 2009 competitions. Also, there are some individually collected data such that the one available from Dr. Nagy Fatey in his Ph.D. thesis, and from Dr. Hazem AbdelAzeem and Dr. Sherif Abdou students at Cairo university.

Personal contacts with these esteemed researchers and with the ICDAR colleagues will be done to see the level of availability of these data sets.

It would be beneficial to do a new data collection at ALTEC with some specific application in mind and the target will be around 3000 writers each writing around 50 words, selected carefully to cover most existing PAWs.

**9- Research Approaches:**

Many approaches have been tried, namely, neural networks, dynamic time warping, hidden Markov models, string similarity measures, and more.

The best systems that have competed in the most recent ICDAR 2009 competition were described earlier.

**10-Strengths, Weaknesses, Opportunities and Threats:**

1. **Strengths** :

There are a few researchers in Egypt who have worked in online handwritten recognition and who can contribute in the future research and products.

The tools required to train systems are mostly available (Matlab, HTK, other neural network and Graphical model tools).

1. **Weaknesses :**

No standard benchmark for the online technology except the one by ICDAR 2009 using the ADAB city names data.

There is no available reliable databases for training systems for various applications.

It is not practical to develop Omni handwritten online OCR systems, rather, systems should be application dependent to limit the complexity of the system in order to obtain good performance.

1. **Opportunities:**

The wide spread use of pen-based hand held devices such as PDAs, smartphones, and tablet-PC, increases the demand for high performance on-line handwritten recognition systems. In particular, in educational domains and businesses where handwritten notes are taken frequently.

Also, there are no obvious systems that support free cursive handwriting with multilingual capabilities.

1. **Threats:**

There are few companies that have already developed handwriting OCR for Arabic like ImagiNet, QuickScript, MyScript Stylus and Sakhr. Other companies may produce such systems soon.

**11-Suggestions for Survey Questionnaire**

1. **Specify the application that online handwriting recognition will be used for**
2. **What is the data used/intended to train the system?**
3. **What is the benchmark to test your system on?**
4. **Would you be interested to contribute in the data collection. At what capacity?**
5. **Would you be interested to buy online Arabic handwritten data?**
6. **Would you be interested to contribute in a competition**
7. **How many persons working in this area in your team? What are their qualifications?**
8. **What are the platforms supported/targeted in your application?**
9. **What is the market share anticipated in your application?**
10. **Would your application support any other languages? Explain.**

**12-List of people/organizations to contact in the survey (for OCR in general and online in particular):**

**Sakhr**

**RDI**

**Orange Labs Cairo**

**IBM Egypt**

**Microsoft CMIC lab Cairo**

**ImagiNet**

**(MoBiDev)**

**AUC**

**GUC**

**BUC**

**ERI (Dr. Samia Mashaly and her group)**

**Cairo university (Many researchers)**

**Ain shams university (Many researchers)**

**Al-Azhar university (Many researchers)**

**Arab academy company for science and technology**

**Dr. Haikal El Abed (http://www.ifn.ing.tu-bs.de/en/sp/elabed/)**

**Dr. Adel Alimi (http://adel.alimi.regim.org/)**

**Dr. Alex Graves (**[**http://www6.in.tum.de/Main/Graves**](http://www6.in.tum.de/Main/Graves)**)**

**13-Key persons to invite in a workshop**

1. **Dr. Hazem Abdel Azeem (Cairo university-ITIDA)**
2. **Dr. Haikal El Abed (IFN, Germany)**
3. **Dr. Adel Alimi (Sfax, Tunisia)**
4. **Dr. Alex Graves (Munich, Germany)**

**14- Suggestions for LR**

For a specific application, such as recognizing city names (ADAB database), with a lexicon of about 1000 words, it was sufficient to collect data from 1000 writers, with a total of about 35,000 words (average of 35 words by each writer).

If we look at the Part of Arabic Words (PAW) frequency, we find that it was also about 35,000 in the whole test set. This shows that it was sufficient to train the system with an average of one PAW occurrence. We think that synthesizing balanced coverage of the PAWs would give better results.

For the training data, we suggest 10,000 writers, one page per person. In the first phase, we will start with **2000** writers, each writing one page (average of **100** words per page), which gives about **200,000** words. We could retain 150,000 words for training and 50,000 for benchmarking.

The vocabulary issue must be addressed. Also, how to ensure the fair coverage of the PAWs. Cairo university has annotation tools to assist manual segmentation of the online data, and Dr. Sherif Abdou will kindly make it available to ALTEC.

We would need to buy 10 data collection boards, which costs 10,000LE.

The persons employed in the collection would cost about 2000\*20=40,000 LE.

The total cost is about 50,000LE without the annotation.

The annotation may take 2 months to complete the parallel with the collection. The annotation would take 1Man-Month i.e. 5000LE. The total cost is thus **55,000LE**.