Chapter 1

On-Line Handwriting Recognition

1.1 Introduction

Handwriting is a very personal skill to individuals. It consists of graphical marks on a surface, it can be used to identify a person, it has the main purpose of communication. This is achivied by drawing letters or other *graphemes*, which in turn represent parts of a language. The characters have a certain basic shape, which must be recognisable for a human in order for the communication process to function. There are rules for the combination of letters, which have the ability - if known to the reader - to help recognise a character or word.

Handwriting was developed as a means of communication and to expand one's own memory. With the advent of each new technologies the question arose, if handwriting was going to survive. However, the opposite seems to be the truth: For example, the printing press increased the number of documents available and therefore increased the number of people who learned to read and write. Through the increased rate of alphabetisation, naturally there was an increased use of handwriting as a means of communication.

In various situations handwriting seems much more practical than typing on a keyboard. For instance children at school are using notepads and pencils or ink pens, which are regarded as a better tool to teach writing by German teachers. Therefore it can be concluded that there is little danger of the extinction of handwriting as a communication tool. In fact, as the length of handwritten messages decreases, the number of people using handwriting increases (Plamondon and Srihari 2000).

1.2 Handwriting Features

Any script of any language has a number of features. The fundamental characteristic of a script is that the differences between the features of different characters are more decisive than the different features of drawing variants of the same letter in individual handwriting styles. There might be exeptions, because *0* and *O* or *1* and *I* respectively, can be written alike. However, in those cases, context makes clear which one was intendet by the writer. Despite the exception, written communication can only work with that fundamental quality (Tappert et al. 1990).

1.2.1 Handwriting Properties of Latin Script

In the latin script we have 26 letters, each of which has two variants, a capital and a lowercase variant. When writing a character in the latin script, there are four main areas, in which the character can reside. All characters have their main part between a top line and a ground line. There is also a middle line. Capital characters stretch out to use the full space between the ground line and the top line, whereas lowercase characters usually use the space between the ground line and the middle line. Some lowercase characters (like lowercase *b, d, f, h, k, l, t*) have an ascender and use the area above the middle line as well, some lowercase characters have a descender and use the area below the ground line (like lowercase *g, j, p, q, y*). In handwritten cursive script, there are writing variants where also some lowercase letters (*f, z*) and certain uppercase characters (*G, J*) expand below the ground line. For all latin-based alphabets, usually one character is finished before the next one starts, however, there are exceptions: In cursive handwriting, the dots on *i* and *j* and the crosses of *t* might be delayed until the underlying portions of all the characters of the word are completed.

XXX Graphic with exmaple of expanding cursive letters here.

1.2.2 Handwriting Properties of East Asian Scripts

Generally, a handwriting is a formed of a number of strokes, that are drawn in a time sequence. Opposed to the latin-based alphbets, consider Chinese and Japanses script. Chinese has a larger alphabet, up to 50000 characters, 3000-5000 of which are in active use. There are also two writing styles, block style - which corresponds to printed characters in latin alphabets, even if handwritten. The other style is cursive style. In block style the individual parts of the character are usually written in proper stroke order, and abide by the proper stroke number. In cursive style the characters are written faster, with less care and don't necessarily abide to stroke number or order. In fact, they are usually written with fewer strokes, connecting some block-style strokes by using simpler radical shapes (Tappert et al. 1990).

In Japanese, three different scripts are in active use at the same time, mixed and next to each other. They are called *Hiragana* (ひらがな), *Katakana*

(カタカナ) and *Kanji* (漢字). Hiragana and Katakana are syllabic alphbets, each containing 46 characters (see **??**), whereas Kanji are essentially the Chinese *Hànzì* characters (▪字) as they were imported into the Japanese language (see **??**).

The different scripts can even be blended with each other within one word. Take for instance the verb *taberu* (食べる - *to eat*). The first character is a kanji character, pronounced */ta/*, also bears the meaning of the word, the second and third characters are the hiragana characters *be* and *ru* which are there for conjugation only as well as for phonetic reasons. However, without them, the character 食 still bears the meaning of the concept *eat*, but the character alone does not result in the verb *taberu*.

1.3 Automated Recognition of Handwriting

1.3.1 Short History of Handwriting recognition

Handwriting recognition (HWR) as a technological discipline performed by machines has been around for many years. The quality of the systems recognising handwriting has improved over the decades. It is the key technology to pen-based computer systems. The first research papers concerned with *pattern recognition* on computers were published in the late 1950ies, *Handwriting recognition* as an individual subject in the early 1960ies. (Goldberg 1915) describes in a US Patent a machine that can recognise alphanumeric characters as early as 1915. However, despite the surprise of how early such a device was invented, it should be taken into consideration that that was before the times of modern computers, therefore the methods he employs are quite different from the algorithms used after the advent of computers, more concretely, computers with screens.

(Tappert et al. 1990) describe in their review the development of handwriting recognition, which was a popular research topic in the early 1970ies and then again in the 1980ies, due to the increased availability of pen-input devices. Generally speaking, handwriting recognition (HWR) involves automatic conversion of handwritten text into a machine readable character encoding like ASCII or UTF-8. Typical HWR-environments include a pen or stylus that is used for the handwriting, a touch-sensitive surface, which the user writes on and a an application that interprets the strokes of the stylus on the surface and converts them into digital text. Usually, the writing surface captures the x-y coordinates of the stylus movement.

1.3.2 Pattern Recognition Problems

The general problem of *pattern recognition* is to take a non-symbolic representation of some pattern, like mouse or pen coordinates and transform it into a symbolic representation like a *rectangle* with its coordinates, or in the case of handwriting recognition, a character. Pattern recognition is a symbol manipulation procedure, that tries to generate discrete structures or sub-sets of discrete structures. Some see it as a game theory problem, where machine 'players' try to match the interpretation of an input produced by machine 'experts' (Zanibbi et al. 2005).

Related Problems

There are several related problems, the recognition of equations, line drawnings and gestures symbols. The recognition of language symbols includes the different large alphabets of Chinese, the different scripts of Japanese, alphabetic scripts like Greek, or Arabic and other non-alphabetic scripts like the Korean Hangul, but also various writing styles of the latin-based alphabets, and diacritics that are used to denote pronunciation variants in different languages using latin script, like Turkish or Vietnamese.

Other Problems that are related to handwriting recognition include for example mahetmatical formula recognition, where mathematical formulas are analysed and put into a computable format (Chan and Yeung 2001). In diagram recognition both the characters and the diagram layout are recognised (Blostein and Haken 1999).

Probem of Similar characters

XXX remake fig 2. in tappert1990 which he had stolen from his reference 240

There are several subproblems to the task of pattern recognition of a character. Different styles of handwriting can be seen in xxx LABELOFGRAPHICABOVE. Towards the bottom of the graphic characters are harder to recognise. In the case of boxed discrete characters, segmentation is done for the machine by the user. Run-on discrete characters are easier to recognise than pure cursive handwriting, because there is a pen-up and pen-down between each character. In cursive handwriting, segmentation between the characters becomes a more difficult task. Some parts of the writing may be delayed, like the crosses of *t* or the dots on *i* or *j*. Besides the segmentation, the discrimination of shapes is often not trivial. Humans may or may not be able to decipher somebody else's handwriting and clearly distinguish between, say *U-V*, but most of the times, context helps with that task. Other characters have similar shapes, too, like *C-L*, *a-d* and *n-h*. Confusion can arise between characters and numbers like *O-0*, *I-1*, *l-1*, *Z-2*, *S-5*, *G-6*, *I-1*, *Z-2*. Lowercase and uppercase are hard to distinguish in the cases of *C-c*, *K-k*, *O-o*, others are mainly distinguished by their position relative to the base line of the rest of the text: *P-p*, *Y-y*. Therefore, context helps the human reader to identify the correct character. This could be used as an advantage in automated pattern recognition, as well. However, the other characters nearby would have to be recognised first, which creates a (solvable) hen and egg problem. In the Japanese scrpit the problem is taken to another dimension. For example 本(root)-木(tree) is only a very simple example that might lead to confusion of the different symbols. However, due to the hierarchical organisation of the characters with their radicals (see **??**) there are many more shapes that look very much alike. From the shape recognition perspective, minor changes to the shape of a character can change its meaning drastically. Compare 嘸 (how, indeed), 撫 stroke, pat), 蕪 (turnip) and 憮 (disappointment). They all contain the radical 無 (nothingness, none), which doesn't seem to have any semantic connection with the characters, however, it can be seen as the main radical in those characters.

1.3.3 Hardware requirements

xxx: see santosh2009 basic tools / techniques / digitizer technology,

Several different hardware commercial products are available in order to capture the x-y coordinates of a stylus or pen. Graphics tablet like the products of the Wacom Co., Ltd.[1] are popular input devices for hand motions and hand gestures. The use of pen-like input devices has also been recommended, since 42% of mouse users report feelings of weakness, stiffness and general discomfort in the wrist and hand when using the mouse for long periods (Woods et al. 2002). Moreover there are PDAs and Tablet PCs, where the writing suface serves as an output device, i.e. an display at the same time. New generation mobile phones also contain touch-displays, but for those it is more common to be operated without a stylus. Those devices interpret user gestures, however the input is given directly with the users fingers. Another rather new development are real-ink digital pens. With those, a user can write on paper with real ink, and the pen stores the movements of the pen-tip on the paper. The movements are transferred to a computer later. It can be expected that with technologies like Bluetooth it may be possible to transfer those data in real-time, not delayed.

1.3.4 Recognition vs Identification

Handwriting recognition is the task of transforming a spatial language representation into a symbolic representation. In the English language (and many others) the symbolic representation is typically 8-bit ASCII. However, with *Unicode* being around for more than a decade now, storage space on harddisks not being as much of an issue any more and *RAM* being readily available to the Gigabytes, it has become more common to use a *UTF-8* encoding, which is a variable-length character encoding for Unicode (The Unicode Consortium 2000). Akin disciplines to handwriting recognition are *handwriting identification*, which is the task of identifying the author of a handwritten text sample from a set of writers, assuming that each handwriting style can be seen as individual to the person who wrote it. The task of *signature verification* is to determine if a given signature stems from the person who's name is given in the signature. Thus, handwriting identification and verification can be used for analysis in the field of jurisdiction. They determine the individual features of a handwritten sample of a specific writer and compare those to samples given by a different or the same writer. By analysing those features one can find out if a piece of handwritten text is authentic or not.

[1] www.wacom.com

1.3.5 Interpretation of Handwriting

Handwriting recognition and interpretation are trying to filter out the writer-specific variations and extract the text message only. This conversion process can be a hard task, even for a human. Humans use context knowledge in order to determine the likeliness of a certain message in a certain context. For instance, a handwritten message on a shopping list that could be read as *bread* or *broad* due to the similarities of the characters for 'e' and 'o' in some cursive handwriting styles, will be interpreted as *bread*, since it is a much more likely interpretation in the shopping list domain. However, if the next word on the shopping list is *beans*, the likelyhood for the interpretation of the first word as *broad* rises, because the collocation *broad beans* is a sequence that is likely on a shopping list, at least more likely than having the interpretation *bread* and then *beans* without a clear separation between the two. Even with non-handwritten, but printed characters, the human mind can be tricked because of the brain's ability to perform these interpretations within miliseconds without conscious thinking. An example of that are modern T-Shirt inscriptions that state things like *Pozilei* in a white font on a green ground (the German police colours in most federal states are green and white), which German native speakers usually read as *Polizei* (police), because that is the most likely interpretation.

1.3.6 On-Line vs. Off-Line recogniton

xxx: see plamondon2000 1.5. blend systems! xxx: see santosh2009: off-line / on-line chapter

On-line HWR means that the input is converted in *real-time*, *dynamically*, while the user is writing. This recognition can lag behind the user's writing speed. (Tappert et al. 1990) report average writing rates of 1.5-2.5 characters/s for English alphanumerics or 0.2-2.5 characters/s for Chinese characters. In online systems, the data usually comes in as a sequence of coordinate points.

Off-line HWR is the application of a HWR algorithm after the writing. It can be performed at any time after the writing has been completed. That includes recognition of data transferred from the real-ink pens (see 1.3.3) to a computing device after the writing has been completed. The stadard case of off-line HWR, however, is a subset of optical character recognition (OCR). An scanner tranfers the physical image on paper into a bitmap, the character recognition is performed on the bitmap. An OCR system can recognise several hundred characters per second.

On-line devices have the dynamic information of the writing, since each point coordinate is captured at a specific point of time. Also, the system know the input stroke sequence, their direction and speed of writing. All these information can be an advantage for an on-line system, however, off-line systems have used algorithms of line-thinning, such that the data consits of point coordinates, similar to the input of online systems (Tappert et al. 1990).

1.4 A Typical On-Line HWR application

A typical HWR application has several parts that follow up on each other in a procedural fashion.

- **Data capturing**: The data is captured through an input device like a writing surface and a stylus.
- **Preprocessing**: The data is segmented, noise reduction like smoothing and filtering are applied.
- **Character Recognition**: Feature analysis, stroke matching, time, direction and curve matching.

1.4.1 Data capturing

how is the data captured? what format? hardware? xxx: see plamondon2000 1.4. xxx: see santosh2009 sampling

1.4.2 Preprocessing

xxx: see santosh2009 pre-processing xxx: see tappert1990 preprocessing: segmentation, noise reduction. xxx: see santosh2009 noise elimination xxx: see santosh2009 normalization xxx: see santosh2009 repetition removal

1.4.3 Character Recognition

xxx: see tappert1990 VI shape recognition. xxx: see plamondon2000: 3.1.1 different models xxx: see all the substroke stuff, santosh, shimodaira2003, nakai2003: very short, properly in OLCCR

1.4.4 Postprocessing

xxx: what happens after the recognition process? xxx: see tappert1990 again in postprocessing chapter.

1.5 HWR of Hanzi and Kanji

- Warum: Um einen Ueberblick ueber HWR-Techniken fuer Japanische Schriftzeichen und verschiedene Herangehensweisen zu verschaffen.
- Nutzen: Leser kann sich ein Bild darueber verschaffen, in welchem Kontext sich die Applikation bewegt.
- Was: research different approaches, see what the focus on, what their specialty is and report about them. Take different specialist papers and compare them.
- Wie: Wiss. Report. / Zusammenfassung. Vergleich.

1.5.1 The current State-of-the-Art in Japanese and Chinese Character Recognition

From the 1990s onwards, On-Line Japanese and Chinese Character Recognition (OJCCR) systems have been aiming at loosening the restrictions imposed on the writer when using an OJCCR system. Their focus shifted from recognition of block style script ('regular' script) to fluent style script, which is also called 'cursive' style. Accuracies of up to about 95% are achieved in the different systems.

(Nakagawa, Tokuno, Zhu, Onuma, Oda, and Kitadai 2008) report their recent results of online Japanese handwriting recognition and its applications. Their article gives important insights into character modeling, which are employed in this application.

xxx: bla. says the opposite. (Chen and Lee 1996) oder auch xxx: (Nakagawa, Tokuno, Zhu, Onuma, Oda, and Kitadai 2008) und (Nakai, Shimodaira, and Sagayama 2003) xxx: zu guter letzt: (Santosh and Nattee 2009)

1.5.2 Overview of a typical OJCCR system

xxx: (Liu, Jaeger, and Nakagawa 2004) have said:

xxx: graphic: handwritten -> character segmentation -> ... -> character codes see fig. 3 of liujaegernakagawa2004

Broadly speaking, from an abstract viewpoint, typical handwriting recognition systems for Chinese and Japanese characters have the same structure like the systems for latin-based alphbets. The process begins with *Character segmentation*, goes on with *Preprocessing*, *Pattern descrpition*, *Pattern recognition* and ends with *Contextual processing*, if applicable. However, there are differences to the standard process, due to the nature of the Chinese characters (see 1.2.2). Especially the pattern representation is divers in the different OJCCR systems, whereas it is naturally more alike in the systems focussing on latin characters. This is due to the fact that the latin alphabet is rather small, but has more variation concerning writing style, whereas the Chinese alphabet has a larger inventory of characters, but less variation in how to write a character - at least - it is widely agreed upon a 'proper' stroke sequence for a character, even across country borders.

xxx: liujaeger2004: 4.1. structural representation. statistical representation.

xxx: liujaeger2004: character classification xxx: liujaeger2004: very short: contextual processing.

References

Blostein, D. and L. Haken (1999). Using diagram generation software to improve diagram recognition: A case study of music notation. *IEEE Trans. Pattern Anal. Mach. Intell. 21*(11), 1121--1136.

Chan, K.-F. and D.-Y. Yeung (2001). Error Detection, Error Correction and Performance Evaluation in On-Line Mathematical Expression Recognition. In *Pattern Recognition*, Volume 34, pp. 1671--1684.

Chen, J.-W. and S.-Y. Lee (1996). A Hierarchical Representation for the Reference Database of On-Line Chinese Character Recognition. In *Advances in Structural and Syntactical Pattern Recognition*, Volume 1121 of *Lecture Notes in Computer Science*, pp. 351--360. Berlin/Heidelberg, Germany: Springer.

Goldberg, H. E. (1915, December). Controller. *United States Patent 1,116,663*.

Liu, C.-L., S. Jaeger, and M. Nakagawa (2004). Online Recognition of Chinese Characters: The State-of-the-Art. *IEEE Transactions on Pattern Analysis and Machine Intelligence 26*, 198--213.

Nakagawa, M., J. Tokuno, B. Zhu, M. Onuma, H. Oda, and A. Kitadai (2008). Recent Results of Online Japanese Handwriting Recognition and Its Applications. In D. Doermann and S. Jaeger (Eds.), *Arabic and Chinese Handwriting Recognition*, Volume 4768 of *Lecture Notes in Computer Science*, pp. 170--195. Berlin/Heidelberg, Germany: Springer.

Nakai, M., H. Shimodaira, and S. Sagayama (2003). Generation of Hierarchical Dictionary for Stroke-Order Free Kanji Handwriting Recognition Based on Substroke HMM. In *Proc. Seventh Int'l Conf. Document Analysis and Recognition*, pp. 514--518.

Plamondon, R. and S. N. Srihari (2000). On-line and off-line handwriting recognition: A comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence 22*(1), 63--84.

Santosh, K. and C. Nattee (2009). A Comprehensive Survey on On-Line Handwriting Recognition Technology and its Real Application to the Nepalese Natural Handwriting. *Kathmandu University Journal of Sience, Engineering and Technology 6*(I), 30--54.

Tappert, C. C., C. Y. Suen, and T. Wakahara (1990). The State of the Art in Online Handwriting Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence 12*(8), 787--808.

The Unicode Consortium (Ed.) (2000). *The Unicode Standard. Version 3.0.* Addison-Wesley.

Woods, V., S. Hastings, P. Buckle, and R. Haslam (2002). *Ergonomics of using a mouse or other non-keyboard input device*, Chapter 3, pp. 23. Number 045 in HSE research report. London: Health and Safety Executive.

Zanibbi, R., D. Blostein, and J. R. Cordy (2005). Recognition Tasks are Imitation Games. In S. Singh, M. Singh, C. Apte, and P. Perner (Eds.), *Pattern Recognition and Data Mining*, pp. 209--218.