

An On-Line Japanese Handwriting Recognition System integrated
into an E-Learning Environment for Kanji

Diplomarbeit

zur Erlangung des Grades
eines Diplom-Linguisten
der
Fachrichtung 4.7 Allgemeine Linguistik
der Universität des Saarlandes.

Anfertigt von Steven B. Poggel
sbp@coli.uni-saarland.de

unter Leitung von
Prof. Dr. Dr. h.c. mult. Wolfgang Wahlster
und
Dr. Tilman Becker

Saarbrücken, den 31.03.2010

Contents

Summary	3
Zusammenfassung	5
Danksagungen	7

Summary

In this work, a Japanese handwriting recognition system is being developed. The system is integrated into an e-learning environment in order to provide a Kanji teaching application with automated error correction. Conceptually, the application is an e-learning environment for Japanese characters, intended for the western learner of the Japanese language. Most e-learning systems of Japanese Kanji provide only a multiple choice method for the learner to reproduce characters. The present prototype offers the ability to enter characters with a stylus on a touch screen system.

The study seeks to determine to what extent it is possible to use modern NLP methods for language learning. While other studies mainly focus on grammatical correction, this application is targeted on the Kanji characters. It will be examined if a handwriting recognition engine can generate informed feedback, suitable for a learner. Additionally, the study examines if that feedback helps obtaining the ability to actively reproduce the Kanji characters.

The prototype developed in this work combines e-learning methods with natural language processing applied to the Japanese script. In order to prepare the task of creating an interdisciplinary software that spans across the aforementioned fields of study, the work reviews the structure of the Japanese script, the current state of the art in handwriting recognition methods and e-learning techniques. The recognition engine implements a structural approach to Kanji character identification. The recognition performs partial analysis of substructures and binds the recognised elements together to form a character. Because of the structural approach it becomes possible to create an informed error recognition that considers linguistic units of the Kanji characters.

The result of the study is

Zusammenfassung

In dieser Arbeit wird eine Handschriftenerkennung für japanische Kanji entwickelt. Der Handschriftenerkennung ist in eine E-Learning-Umgebung integriert und liefert eine automatisch generierte Fehlerkorrektur für Lernende.

Das System ist in konzeptioneller Hinsicht eine E-Learning-Anwendung für das Erlernen der japanischen Schrift. Letztere weist aufgrund ihrer morphemischen Struktur einen hohen Komplexitätsgrad auf und benötigt daher besonderen Lernaufwand. Die meisten E-Learning-Systeme für asiatische Schriftzeichen bieten Zeichenabfrage als Multiple-Choice an, da die Eingabe der Zeichen für einen Lernenden sonst ein technisches Problem darstellen würde. Der in dieser Arbeit erstellte Prototyp bietet die Möglichkeit zur handschriftlichen Eingabe von Zeichen auf einer dafür geeigneten Bildschirmoberfläche. Das ist ein Alleinstellungsmerkmal unter den E-Learning-Anwendungen für die japanische Sprache.

Die Studie untersucht, inwieweit es im Bereich des Schrifterwerbs möglich ist, NLP und Lernmethoden zusammenzubringen. Dabei wird nicht mit Parsing-Methoden die grammatische Struktur der Sprache untersucht, sondern vielmehr die interne Struktur der Kanji zugrunde gelegt und durch einen Handschriftenerkennung erfasst. Dabei sollen Schreibfehler strukturell erkannt werden. Intelligentes Feedback soll dem Lernenden dabei helfen, die Fähigkeit der aktiven Reproduktion der Kanji zu erwerben.

Da der Prototyp eine disziplinübergreifende Software ist, die in den Bereichen Handschriftenerkennung und E-Learning angesiedelt ist, wird in der vorliegenden Arbeit der Forschungsstand der beiden untersucht. Weiterhin wird die Struktur der japanischen Schrift linguistisch analysiert und dargestellt. Die Kombination der drei Disziplinen in einer Studie führt dazu, dass die Substrukturen der Kanji überhaupt programmatisch analysiert werden können, wodurch die Fehlererkennung ermöglicht wird.

Danksagungen

Ich möchte mich bei allen bedanken, die mir bei der Erstellung dieser Arbeit geholfen haben. Jeder hat einen wertvollen Beitrag geleistet, den ich sehr zu schätzen weiß. Besonders bedanken möchte ich mich bei Dr. Tilman Becker, der mir über die gesamte Zeit wertvolle Hinweise gegeben hat, mit dem ich hochinteressante fachliche Diskussionen hatte, die diese Arbeit vorangetrieben und verbessert haben. Weiterhin bedanke ich mich bei Prof. Wahlster, in dessen Abteilung *Intelligente Benutzerschnittstellen* am DFKI ich längere Zeit als HiWi und als Diplomand tätig war, für das Ermöglichen dieser Diplomstudie. Mein besonderer Dank gilt allen Korrekturlesern, die mich vor dem ein oder anderen Rechtschreibfehler und inhaltlichen Unstimmigkeiten bewahrt haben, sowie allen, die mir mit Rat und Tat zur Seite gestanden haben. Diejenigen, die ich vergessen habe, mögen mir verzeihen. Matthew Brown, Özgür Demir, Jacky Freiheit, Marco Geiger, Sebastian Germesin, Matthias an der Heiden, Gerd Herzog, Thomas Kleinbauer, Matthias Kwaterski, Hartmut Nebe, Tino Ortega-Gomez, Peter Poller, Sigrid Poggel, Klaus Poggel, Gerhard Sonnenberg, Rainer Stahlmann, Anke Steffen, Hideki Yamaguchi. Alle in der Arbeit verbleibenden Fehler sind natürlich meine eigenen.

References

**Document created on Wednesday 31st
March, 2010 at 01:58**