

## Guide to OCR for Arabic scripts /

Märgner, Volker El Abed, Haikal

Springer, 2012



Monografía

Optical Character Recognition (OCR) is a key technology enabling access to digital text data. This technique is especially valuable for Arabic scripts, for which there has been very little digital access. Arabic script is widely used today. It is estimated that approximately 200 million people use Arabic as a first language, and the Arabic script is shared by an additional 13 languages, making it the second most widespread script in the world. However, Arabic scripts pose unique challenges for OCR systems that cannot be simply adapted from existing Latin character-based processing techniques. This comprehensive Guide to OCR for Arabic Scripts is the first book of its kind, specifically devoted to this emerging field. Presenting state-of-the-art research from an international selection of pre-eminent authorities, the book reviews techniques and algorithms for the recognition of both handwritten and printed Arabic scripts. Many of these techniques can also be applied to other scripts, serving as an inspiration to all groups working in the area of OCR. Topics and features: Contains contributions from the leading researchers in the fieldWith a Foreword by Professor Bente Maegaard of the University of CopenhagenPresents a detailed overview of Arabic character recognition technology, covering a range of different aspects of pre-processing and feature extractionReviews a broad selection of varying approaches, including HMM-based methods and a recognition system based on multidimensional recurrent neural networksExamines the evaluation of Arabic script recognition systems, discussing data collection and annotation, benchmarking strategies, and handwriting recognition competitionsDescribes numerous applications of Arabic script recognition technology, from historical Arabic manuscripts to online Arabic recognitionThis authoritative work is an essential reference for all researchers and graduate students interested in OCR technology and methodology in general, and in Arabic scripts in particular

https://rebiunoda.pro.baratznet.cloud:38443/OpacDiscovery/public/catalog/detail/b2FpOmNlbGVicmF0aW9uOmVzLmJhcmF0ei5yZW4vMzE1MzU0OTg

Título: Guide to OCR for Arabic scripts Volker Märgner, Haikal El Abed, editors

Editorial: London New York Springer 2012

Descripción física: 1 online resource

**Tipo Audiovisual:** Computer science Text processing (Computer science) Computer vision Optical pattern recognition Image Processing and Computer Vision Document Preparation and Text Processing

Bibliografía: Includes bibliographical references and index

Contenido: Part 1.) Pre-Processing -- An Assessment of Arabic Handwriting Recognition Technology Sargur N. Srihari and Gregory Ball. -- Layout Analysis of Arabic Script Documents Syed Saqib Bukhari, Faisal Shafait and Thomas M. Breuel. -- A Multi-stage Approach to Arabic Document Analysis Eugene Borovikov and Ilya Zavorin. -- Pre-processing Issues in Arabic OCR Zhixin Shi, Srirangaraj Setlur and Venu Govindaraju. -- Segmentation of Ancient Arabic Documents Abdel Belaïd and Nazih Ouwayed. -- Features for HMM-Based Arabic Handwritten Word Recognition Systems Laurence Likforman-Sulem, Ramy Al Hajj Mohammad, Chafic Mokbel, Fares Menasri and Anne-Laure Bianne-Bernard, et al Part 2.) Recognition -- Printed Arabic Text Recognition Irfan Ahmed, Sabri A. Mahmoud and Mohammed Tanvir Parvez. -- Handwritten Arabic Word Recognition Using the IFN/ENITdatabase Mario Pechwitz, Haikal El Abed and Volker Märgner. -- RWTH OCR: A Large Vocabulary Optical Character Recognition System for Arabic Scripts Philippe Dreuw, David Rybach, Georg Heigold and Hermann Ney. -- Arabic Handwriting Recognition Using Bernoulli HMMs Ihab Alkhoury, Adrià Giménez and Alfons Juan. -- Handwritten Farsi Word Recognition Using Hidden Markov Models Puntis Jifroodian Haghighi and Ching Y. Suen. -- Offline Arabic Handwriting Recognition with Multidimensional Recurrent Neural Networks Alex Graves. -- Application of Fractal Theory in Farsi/Arabic Document Analysis Saeed Mozaffari. -- Multi-stream Markov Models for Arabic Handwriting Recognition Yousri Kessentini, Thierry Paquet and AbdelMajid Ben Hamadou. --Toward Distributed Cursive Writing OCR Systems Based on a Combination of Complementary Approaches Maher Khemakhem and Abdelfettah Belghith Part 3.) Evaluation -- Data Collection and Annotation for Arabic Document Analysis Ilya Zavorin and Eugene Borovikov. -- Arabic Handwriting Recognition Competitions Volker Märgner and Haikal El Abed. -- Benchmarking Strategy for Arabic Screen-Rendered Word Recognition Fouad Slimane, Slim Kanoun, Jean Hennebert, Rolf Ingold and Adel M. Alimi Part 4.) Applications -- A Robust Word Spotting System for Historical Arabic Manuscripts Mohamed Cheriet and Reza Farrahi Moghaddam. -- Arabic Text Recognition Using a Script-Independent Methodology: A Unified HMM-Based Approach for Machine-Printed and Handwritten Text Premkumar Natarajan, Rohit Prasad, Huaigu Cao, Krishna Subramanian and Shirin Saleem, et al. -- Arabic Handwriting Recognition Using VDHMM and Over-segmentation Amlan Kundu and Tom Hines. --Online Arabic Databases and Applications Houcine Boubaker, Abdelkarim Elbaati, Najiba Tagougui, Haikal El Abed and Monji Kherallah, et al. -- On-line Arabic Handwritten Word Recognition Based on HMM and Combination of On-line and Off-line Features Sherif Abdelazeem, Hesham M. Eraqi and Hany Ahmed

## Lengua: English

Copyright/Depósito Legal: 801364279 985032750 990483332 1005748691 1111026863 1203988174 1259066130

**ISBN:** 9781447140726 electronic bk.) 1447140729 electronic bk.) 1447159764 9781447159766 9781447140719 1447140710 1280996110 9781280996115 9786613767721 6613767727

**Materia:** Optical character recognition Arabic character sets (Data processing) Reconnaissance optique des caractères Jeux de caractères arabes (Informatique) optical character recognition. COMPUTERS- General. Informatique. Arabic character sets (Data processing) Optical character recognition.

Autores: Märgner, Volker El Abed, Haikal

**Enlace a formato físico adicional:** Print version Märgner, Volker. Guide to OCR for Arabic Scripts. Dordrecht : Springer, 2012 9781447140719

## **Baratz Innovación Documental**

- Gran Vía, 59 28013 Madrid
- (+34) 91 456 03 60
- informa@baratz.es