Style characterization of machine printed texts
Bagdanov, A.D.

Citation for published version (APA):
Bagdanov, A. D. (2004). Style characterization of machine printed texts

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
Contents

0 Prelude .................................................. 1
1 Introduction ............................................. 3
  1.1 Elements of style ..................................... 3
  1.2 Context and scope ................................... 4
    1.2.1 Context ........................................ 4
    1.2.2 Scope .......................................... 7
  1.3 Organization of this thesis ......................... 11
2 Characterizing layout style using first order Gaussian graphs 13
  2.1 Definitions and basic concepts ....................... 14
    2.1.1 First order Gaussian graphs .................. 14
    2.1.2 Technicalities ................................ 17
    2.1.3 Reflections .................................... 20
  2.2 Clustering and classification ....................... 21
    2.2.1 Hierarchical clustering of FOGGs .............. 21
    2.2.2 Classification using first order Gaussian graphs 23
  2.3 Experiments .......................................... 25
    2.3.1 Test data ...................................... 25
    2.3.2 Classifiers .................................... 26
    2.3.3 Experimental results .......................... 27
    2.3.4 Computational efficiency ..................... 28
    2.3.5 Analysis ....................................... 29
  2.4 Discussion .......................................... 31
3 Multi-scale visual style characterization with rectangular granulometries 35
  3.1 Document genre ...................................... 37
  3.2 Granulometries ...................................... 39
  3.3 Document representation ............................ 41
    3.3.1 Rectangular size distributions ............... 42
    3.3.2 Efficiency ..................................... 43
    3.3.3 Feature space reduction and interpretation .... 44
  3.4 Experimental results ................................ 46
    3.4.1 Genre classification .......................... 47
    3.4.2 Document image retrieval ...................... 47
## Contents

### 3.5 Discussion

### 4 Probing textual style with local vertical granulometries

#### 4.1 Another look at granulometries

- 4.1.1 The key observation
- 4.1.2 Introducing localization

#### 4.2 An efficient word spotter

#### 4.3 A Generative model

#### 4.4 Experiments and illustration

#### 4.5 Discussion

### 5 Autocorrelation-driven restoration of scanned color halftones

#### 5.1 Halftone process color

- 5.1.1 Halftone color reproduction
- 5.1.2 Scanned color halftones

#### 5.2 Diffusion of scanned color halftones

- 5.2.1 Linear diffusion filtering
- 5.2.2 Nonlinear diffusion filtering

#### 5.3 Measuring local autocorrelation

#### 5.4 Experiments

#### 5.5 Discussion

### 6 A functional approach to software design in image processing research environments

#### 6.1 Introduction

#### 6.2 A critique of pure reason

- 6.2.1 Analysis

#### 6.3 Design considerations

- 6.3.1 Goals
- 6.3.2 Choice of language
- 6.3.3 Previous work

#### 6.4 Architecture

#### 6.5 Primitive types and operations

- 6.5.1 Types and typing
- 6.5.2 Primitive image operations

#### 6.6 Backend substitution

#### 6.7 Case studies

- 6.7.1 Linear scalespace
- 6.7.2 Complete lattice morphology
- 6.7.3 An algebraic expression compiler

#### 6.8 Discussion
7 Summary and concluding remarks 131
  7.1 Summary .................................................. 131
  7.2 Concluding remarks ...................................... 133

Bibliography 135

Samenvatting 143

Acknowledgements 145