Scientific Service Project Z03

Image Processing Methods for Determining Visual Manuscript and Character Features

Prof. Dr.-Ing. H. Siegfried Stiehl
Dr.-Ing. Volker Märgner

Thomas Konidaris, PhD
Hussein A. Mohammed, M.Sc.

Aim of the Project

Piloting of a web-based platform (Advanced Manuscript Analysis Portal, AMAP) to support the research programme of the SFB with well-established and innovative methods of image processing (IP)

Key Aspects

- development of the workbench (ARMA, first phase) towards a portal-based solution (AMAP)
- further development of an intuitive user interface
- integration with the repository and long-term archiving (see also sub-project INF)
- validation and evaluation in collaboration with sub-projects (see also aim 2)
- advice on, and training for, use of the portal AMAP
- networking with research groups and service centres

Continuation of aims of first phase:

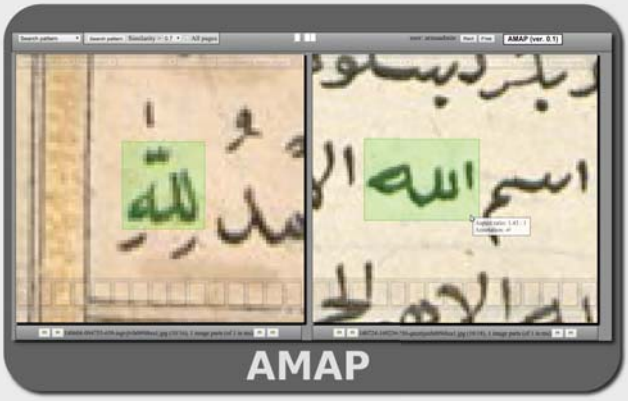
- 1 Development of a workbench ARMA
- 2 Application of image processing methods
- 3 Development of innovative IP methods

Aim 1: Further Development of the Workbench (AMAP)

Implementation of a browser-based, interactive, extendable and intuitively usable system for manuscript research:

- computer-aided identification of palaeographic features
- methods for the statistical analysis of features
- analysis of large pools of data (see also aims 2 and 3)
- integration with the SFB's repository (see also INF)

Example: Comparison of text samples in two manuscripts (AMAP screen shot of prototype user interface)



Aim 2: Application of IP Methods

A requirement analysis for the second phase of the SFB revealed at least eight potential collaborations between Z03 and other sub-projects:

- A03: attribution of palm-leaf manuscripts originating from northern Laos and northern Thailand to monasteries and scribes
- A04: stylistic grouping of colophons in Tamil palm-leaf manuscripts
- B07: inspection of manuscripts from the Codex Florentinus to identify identical scribes (in collaboration with Z02)
- B08: detection of magical patterns in voluminous Hebrew manuscripts and check for identical scribes
- C04: identification of deviations from norms in respect of layout and design characteristics within large data pools
- C05: support of dating in respect of medieval Greek manuscripts by means of 'word spotting' and the evaluation of palaeographic characteristics (in collaboration with Z02)
- C07: identification of scribes in Swahili manuscripts (additional materials analysis by Z02)
- C08: identification of passages by identical scribes in manuscripts from the Frankish Empire (in collaboration with Z02)

Aim 3: Development of Innovative IP Methods

Besides an interactive measurement of e.g. palaeographic features, the second phase will primarily involve developing methods for the identification of scribes, the analysis of writing styles, and for 'word spotting':

- analysing entire pages for similarity of writing styles,
- specifying a template and searching for similar patterns .

The findings will be useful for the purposes of identifying scribes with known writing styles and/or with known, characteristic ways of writing particular characters.

Example: Employing the spatial configuration of 'interest points' to find similar patterns

Template

Manuscript section with direct matches