An Image Processing Portal and Web-service for the Study of Ancient Documents

Ségolène M. Tarte, David Wallom, Pin Hu, Kang Tang, Tiejun Ma
Oxford e-Research Centre
University of Oxford
7, Keble road
Oxford OX1 3QG, UK
{segolene.tarte, david.wallom, pin.hu, kang.tang, tiejun.ma}@oerc.ox.ac.uk

Abstract

Linking up two projects that are dedicated to facilitate the work of documentary scholars, this paper presents image processing algorithms tailored to the study of ancient documents and how they have been made available to the users through a portal that calls upon a web-service exploiting grid computational power. To that end, image processing algorithms were wrapped to fit into the National Grid Service (NGS) Uniform Execution Environment; the data model of an existing Virtual Research Environment (VRE-SDM) was extended; JSR-168 compliant portlets were developed to facilitate secure and seamless distributed image analysis; and a GridSAM interface between the portal and the NGS-installed algorithms was developed. The outcomes of the project include: a web-based application, a proof of concept for the usability of the VRE-SDM platform, an opportunity for wider dissemination for the eSAD algorithms, and a proof of feasibility for the use of the NGS for Humanities applications.

1. Introduction

Ancient and scarcely legible documents represent a challenge for both the Classics community and the Image processing community. The documents of interest here are wooden stylus tablets as used in Classical Antiquity. The original text was incised on tablets consisting of a wax-coated wooden blocks, of which only the wood and the incisions that traversed the wax and marked the wood remain as traces of the script.

For Classical historians, adding to the legibility challenge, access to the document is often limited. Collaborative work on the documents is one factor that facilitates their deciphering, transcription and interpretation. The Virtual Research Environment for the Study of Documents and Manuscripts pilot software (VRE-SDM[1]) [1] was developed to promote non-colocated work between documentary scholars, by providing them with a web-based interface allowing them to visualize and annotate documents in a digitized form, share annotations, exchange opinions and access external knowledge bases.

For the Image processing community, enhancing the visibility of the script based on the digital pictures of the text-bearing artefacts presents challenges of a different nature than enhancement and feature detection in images such as medical, synthetic or natural images. Appropriate image capture protocol allows to reproduce the real-world processes that the Classicists draw upon when in possession of such a tablet, taking advantage of the shadow-stereo principle to amplify the script-signal [2]. Yet, the unavoidable noise in the images due to the imaging technology is combined to the noise with respect to the text under scrutiny: texture of the wood, stains, damage, illumination conditions. Specifically tailored enhancement algorithms have thus been developed as part of the e-Science and Ancient Documents (eSAD[2]) project [8]. They enable the minimization of the interferences due to these noises of various nature and help make the text more legible.

In this paper, we present the latest eSAD image processing algorithms and how they are made accessible to the Classics community, via a web-service that uses the UK National Grid Service (NGS[3]) resources and can be called through a portal from within the VRE-SDM application.

2. Material and Methods

In this section we detail the technology involved in the design and development of our image processing portal dedicated to documentary research.

[1] http://bvreh.humanities.ox.ac.uk/VRE-SDM.html —last checked: 17/07/2009

[2] http://esad.classics.ox.ac.uk —last checked: 17/07/2009

[3] http://www.grid-support.ac.uk —last checked: 17/07/2009

2.1. Image processing algorithms

Various algorithms have been developed to assist Classical historians in their detection of the text. The first algorithm aims at rebalancing illumination throughout the image. This is achieved by homomorphic filtering via a low-pass filter, based on the reasonable assumption (given the imaging protocol [8]) that the brightness in an image varies slowly [6]. The texture of the wood, in the form of the grooves of the wood grain, also represents noise that complicates text detection. In order to remove the wood grain, we have developed two algorithms. Mostly, the wood grain runs parallel to the direction of the writing (the X axis), meaning that an underlying stripe pattern occurs, following the changes in direction of the normal to the wooden surface which condition the reflection of the light on the tablet. The first method removes the wood grain additively by performing horizontal smoothing and equal leveling by a moving average method. The second method compensates for the dark wood grain stripes multiplicatively by using Lambertian reflectance considerations [4, 7]. The wood grain running horizontally, the gray values in the image I locally follow: $\forall y, \sum_x I(x, y) = c \cdot \cos\theta_y$, where c is the albedo (locally constant colour), and θ_y is the angle between the normal to the surface and the incident light, which varies only in its Y component. Locally, in this context, because the wood grain runs horizontally, means in a reasonably small portion of the image (e.g., a span of 50 pixels). For each local line in the image, the algorithm identifies the value of θ_y and resets it to the value given by the elevation angle of the light source, thus flattening the grooves of the wood grain (c.f. Fig. 1). A third algorithm was implemented, enabling to better visually isolate the incised text. As the imaging protocol amplifies brightness and shadows around the incisions of the text, a two-level Otsu segmentation [5] enables to statistically identify the pixels that are likely to be of interest and flatten all the others to a single gray value. All these algorithms were implemented in Matlab, and tested on various images of wooden tablets. They are the algorithms that are made available through the tool presented in this paper. More advanced feature detection approaches for further enhancement of the text's legibility is one current focus of research of the eSAD project.

2.2. Portal and data model extension

To integrate the functionality offered by the algorithms presented in section 2.1 within the VRE-SDM application, we expanded both its interface and its underlying data model, designing and developing our own compatible portal. Within this portal, users can: search original or processed images; display thumbnails and image metadata; display single or tiled images; view and edit annotations on images; select the desired algorithms to perform image analyses; select a region of interest for local image analyses; run analyses on individual images or on a group of images belonging to the same object view; view real-time job status; and compare images in a dual display window (c.f. Fig. 1). Complying with the VRE-SDM's architecture, and enabling easy integration, we developed an uPortal to hold JSR-168 portlets and, manage user accounts and access rights. At the server end, a generic Java Servlet and a set of related classes were developed to handle client-side requests such as information retrieval from TripleStore and job submission to the NGS. A Direct Web Remoting (DWR) based COMET implementation is used to push server data to the browser in order to deliver real-time message exchanges and job status to the users. On the client side, a JavaScript library delivers AJAX dynamic functions to allow actions such as image display, mapping of image co-ordinates, etc... A Linux file system holds the images, and a Tomcat Web container allows remote access to the uPortal and to the stored images. Like the metadata from the original VRE-SDM data model, the metadata from the extended model uses TripleStore for storage and the Resource Description Framework (RDF) format for retrieval. The extension of the data model allows to differentiate between non-processed and processed images, as well as to trace the nature of the analysis that a processed image has undergone (c.f. Fig. 2). This extended data model enables the user to keep track of the imaging parameters of a given image, of the extent of any defined region of interest, and of the sequence of algorithms that have been applied to a given image to produce a specific processed image.

2.3. Web-service and communication protocols

Our image processing portal was developed to facilitate secure and seamless distributed image analysis to users. In a first instance, the original image processing algorithms were wrapped into the NGS's Uniform Execution Environment. To that end, the Matlab code was compiled and installed on the NGS, along with the Matlab Run Time Environment, thus avoiding licensing problems. The submission and monitoring web service we chose is GridSAM[4], an open source OMII-UK software. At the client/server interface, all the relevant information is passed along in an OGF JSDL file containing: the executable name and arguments; the standard output and error files; the data staging in/out information; and the access information needed for proxy identification. The standard JSDL web service interface and extensive client side API that GridSAM provides also helps to reduce the development effort typically required by using JavaCoG Kit. Another benefit we found during this project when using GridSAM is that, it helps

[4] http://gridsam.sourceforge.net —last checked 17/07/2009

to keep a clean boundary between normal web application development and grid development. A job is submitted by calling the submitJob() method on the Job Manager wrapper object, which returns a unique job ID once the submission is completed. The job ID acts as a unique handler for each submitted job. The portal can query the current status of each submitted job by providing this job ID handler. Once a job has been received by the GridSAM server, it identifies the data staging information from the JSDL file and starts to download the source image file(s) onto the NGS (c.f. Fig. 3). Currently, the JSDL file needs to contain the data staging information, due to the fact that the latest version of GridSAM (v2.1.14) supports only user name and password authentication when using SFTP. This was identified as a security risks and was mitigated by using SSL to encrypt all the plain SOAP messages between the portal and the GridSAM server. The actual executable is then invoked by GridSAM on behalf of the portal when computing resources are available on the NGS. During the period of job execution, the portal periodically queries GridSAM for job status updates. If a job has been successfully completed, the output images are staged out via SFTP, based on the access information from the JSDL file. If, on the other hand, errors occur before a job is finished, a brief error message is returned back to the portal.

3. Results and Discussion

Taking advantage of the VRE-SDM platform, we have developed a portal through which a web-service can be invoked to perform grid-enabled image processing operations on images of ancient documents.

Regarding the technology used during this project, two interesting findings were made pertaining to: (1) the submission of multiple jobs to the NGS, and (2) a security work-around that was necessitated by the use of GridSAM in combination with SFTP.

At present, on the NGS, it is not possible to submit multi-jobs via JSDL, a multi-job (such as batch-processing of a group of images) needs to be split into a set of single JSDL jobs. A multi-job, containing a set of sub-jobs (as supported by GT4), could relatively speed-up the whole execution process and make job management easier, however it is not supported by the NGS to date.

Our second finding relates to the data staging mechanism through GridSAM, which supports FTP, SFTP, HTTP and GridFTP. In this project, the portal front end is responsible for submitting jobs and we wanted the processed output images to be securely transferred back to portal. In order to keep a clear boundary between the portal and the grid, the portal needed to not be aware of the existence of the grid at the backend, nor to be required to provide support to the GridFTP protocol. Hence, we decided to use SFTP, rather than GridFTP for data staging. Indeed SFTP is supported by most UNIX based operating systems out-of-box. Its relative simplicity and built-in security makes it an ideal choice in this project. Typically SFTP transfers files securely through the SSH-2 protocol, which supports not only simple username and password authentication but also public-key based authentication. One of the many advantages the latter has, is that it does not require the client to provide their remote login details and then transfer the credentials on the wire. However, the current (and latest, to date) version of GRIDSAM 2.1.x supports only simple username and password authentication when using SFTP for data staging, so the username and password information have to be embedded within JSDL as plain text. Although we have been using SSL to encrypt all JSDL messages between the client and GridSAM to mitigate the security risk, we found that, in practice, it did increase the complexity and time required with regard to installation and configuration. If public-key authentication was supported by GridSAM then we would be able to not only take full advantage of public key authentication but also to avoid transferring login credentials through JSDL on wire.

Altogether, the software enables documentary scholars to seamlessly make use of the computational power of the NGS to run elaborate image processing algorithms through a web-browser. The Classics community is one already sensitized to ICT [3], and this new tool, benefiting from the experience and implementation of the VRE-SDM, which was based on extensive requirements gathering, allows the use of web-technologies to be widened from knowledge bases and collaborative environments to the execution of advanced image processing algorithms to enhance hard-to-read documents. In addition to its web-based software output, the research outcome of this work is three-fold.

Firstly, this projects constitutes a proof of concept for the VRE-SDM pilot. By deploying our image processing portal within the VRE-SDM, we were able to show that the VRE-SDM design is not only sound, but that it was also easy to extend its data model in order to integrate new functionality.

Next, this project allows the eSAD project to benefit from a wider dissemination of its research results. Contrary to the VRE-SDM project, which is quite mature, the eSAD project is still young; being able to offer these algorithms to the Classics community at an early stage also allows for wider use, thus feeding back user appreciations on their appropriateness in the eSAD algorithms design loop.

Finally, throughout this project we have been able to interact with Classicists and Image processing researchers, neither of whom were acquainted with the use of grid-technology and web-services, without overwhelming them with the details of the communication protocols between the web-based application and the NGS. Further, the image processing researcher is now provided with the necessary

knowledge to wrap future algorithms into NGS executables, and adapt the portal, along with its associated data model, to enable it to offer any newly developed algorithm as a functionality.

4. Conclusion

This short project was able to develop within four months a portal through which image processing functionality is offered as a web-service, deploying it as an extension of the VRE-SDM platform and integrating research output from the eSAD project. It also substantially reduced the "fear-factor" generated by the idea of the use of the NGS amongst researchers not familiar with grid-technology, thus opening the possibility of taking advantage of distributed computing for future more complex feature detection algorithms dedicated to Humanities applications. The VRE-SDM project was proved a valuable, easy to extend tool, and GridSAM proved to be a robust interface for communication with the NGS. Future work will look into integrating more of the results of the eSAD project within the developed portal. For sustainability purposes, we will also have to adapt to the changes that the VRE-SDM pilot is going through, not least of which is the porting of the VRE-SDM tools to the OpenSocial API.

5. Acknowledgments

This work was conducted as part of the Engage Initiative[5], supported by the National Grid Service (NGS) and the UK Open Middleware Infrastructure Institute (OMII-UK); it was funded by the Joint Information Systems Committee (JISC).

The authors further wish to thank S. Brewer of the OMII-UK, Dr M. Turilli of the NGS, Dr M. Terras, Prof. A.K. Bowman, Prof. M. Brady and H. Roued from the eSAD project and R. Kirkham and J. Pybus from the VRE-SDM project for facilitating this collaboration.

References

[1] A. K. Bowman, C. V. Crowther, R. Kirkham, and J. Pybus. A virtual research environment for the study of documents and manuscript. *Digital Classicist*, —forthcoming, 2009.

[2] M. Brady, X. Pan, V. Schenck, M. Terras, P. Robertson, and N. Molton. *Images and Artefacts of the Ancient World*, chapter Shadow stereo, image filtering and constraint propagation, pages 15–30. British Academy, London, 2005.

[3] G. Crane. *A Companion to Digital Humanities.*, chapter Classics and the Computer: An End of the History. Blackwell Companions to Literature and Culture, Oxford, 2008.

[4] N. Molton, X. Pan, M. Brady, A. K. Bowman, C. Crowther, and R. Tomlin. Visual enhancement of incised text. *Pattern recognition*, 36(4):1031–43, 2003.

[5] N. Otsu. A threshold selection method from gray level histograms. *IEEE Trans. Systems, Man and Cybernetics*, 9:62–66, Mar. 1979. minimize inter class variance.

[6] X. Pan, M. Brady, A. Bowman, C. Crowther, and R. Tomlin. Enhancement and feature extraction for images of incised and ink texts. *Image and Vision Computing*, 22:443–51, 2004.

[7] S. Tarte, M. Brady, H. R. Olsen, M. Terras, and A. Bowman. Image acquisition & analysis to enhance the legibility of ancient texts. *E-Science All Hands Meeting 2008*, Dec 2008.

[8] S. M. Tarte. Papyrological investigations: Tranferring perception and interpretation into the digital world. *Literary and Linguistics Computing*, —forthcoming, 2009.

[5] http://engage.ac.uk/ —last checked 17/07/2009

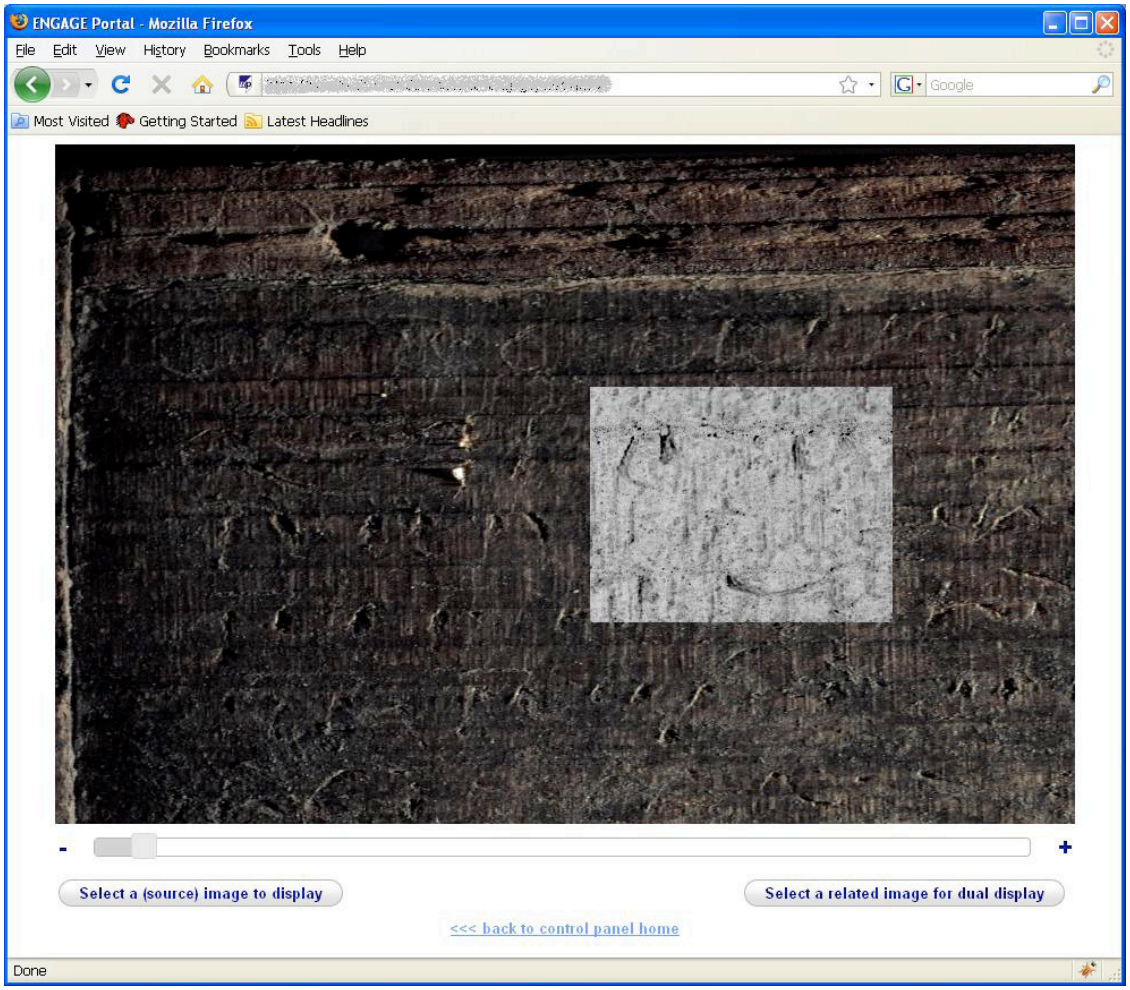

Figure 1. The image processing portal in dual display mode. The background image is a portion of the original image of a wooden stylus tablet, the windowed image is a smaller overlaid portion of a processed version of the same image. The image processing, in this specific case, involved illumination correction, wood grain removal and color inversion (for better visibility).

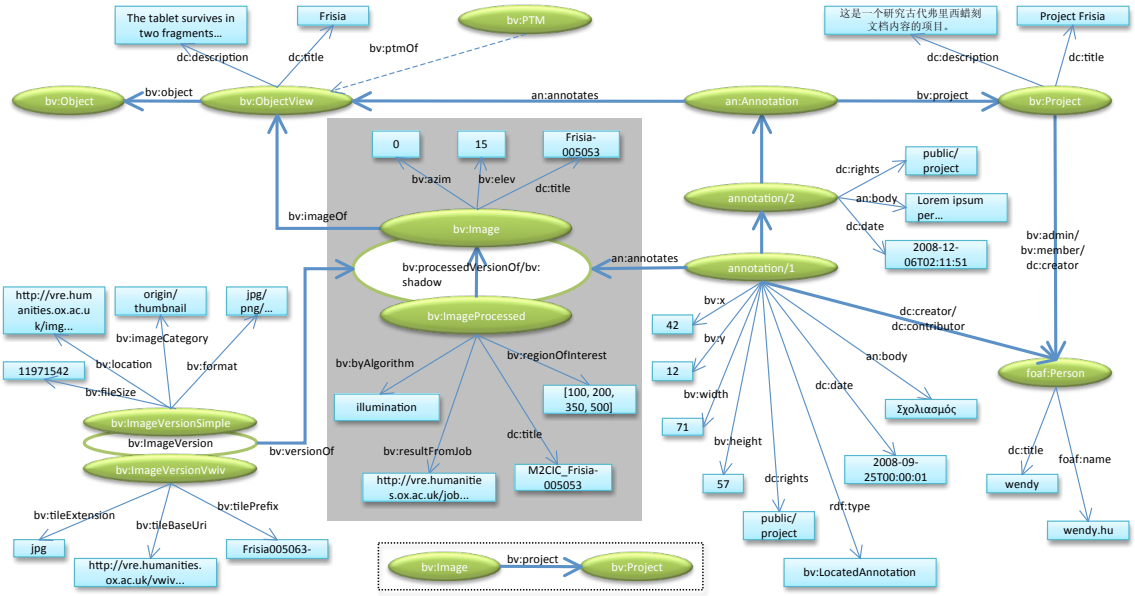

Figure 2. Overview of the VRE-SDM data model and our additions to it. The elements in the grayed out box represent our extension of the original model; they allow the user to trace the provenance of any given processed image.

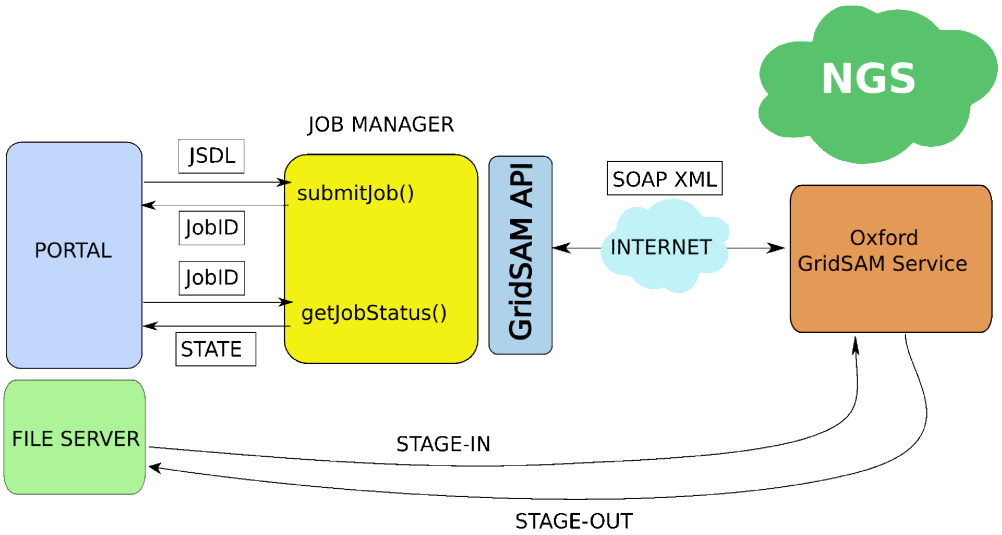

Figure 3. Communication protocol between the portal, the file server hosting the images, and the NGS via GridSAM.