Image Capture and Analysis for Inscribed Documents

Image, Text, Interpretation:
e-Science, Technology, and Documents

Ségolène Tarte

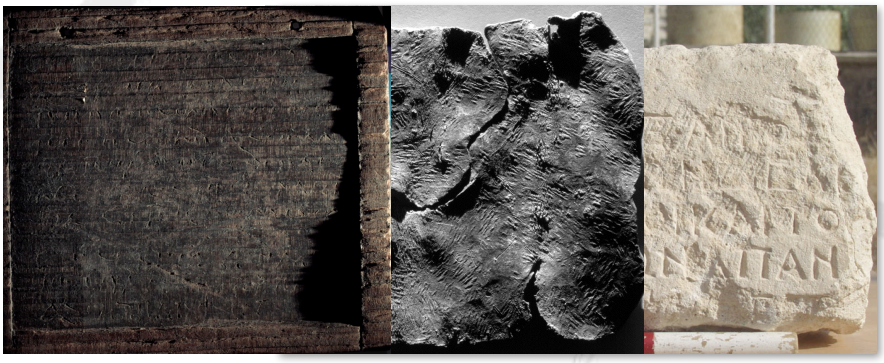
eSAD Workshop, OeRC, Oxford
13th May 2009

Oxford University
Centre for the Study of Ancient Documents
UCL Department of Information Studies

Oxford e-Research Centre

Image, Text, Interpretation: e-Science, Technology and Documents

- Targeted texts (*in this talk*) inscribed in:
  - Stone
  - Lead
  - Wood



Where can image processing help?

Reading Level	Thematic Subject
8	Meaning or sense of document as a whole
7	Meaning or sense of a group or phrase or words
6	Meaning or sense of a word
5	Discussion of grammar
4	Identification of possible word or morphemic unit
3	Identification of sequence of characters
2	Identification of possible character
1	Discussion of features of character
0	Discussion of physical attributes of the document
-1	Archaeological or historical context

'Increasing' meaning

Image analysis Low level percpets

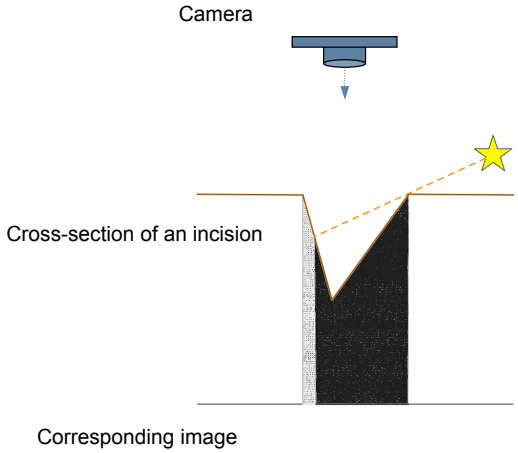
- Melissa's talk presents how these levels were identified
- Mike's talk covers the image analysis for reading level 2
- Henriette's talk covers reading levels 2, 3 and 4

Main characteristics of inscribed texts

- Script is made of incisions
  - Out-of-plane information, i.e., volumetric information
- Legibility affected by heterogeneity of the writing support
  - Damage and stains
  - Bumpiness (e.g., lead tablets); woodgrain (wood)
  - Palimpsest

Capturing volumetric information with digital photographs

- Shadow-stereo principle:
  - The highlight and shadow areas move according to the position of the light
  - Stains stay in place
- ➢ Yields volumetric information



→ Mimics Alan and Roger's tablet rocking in raking light!

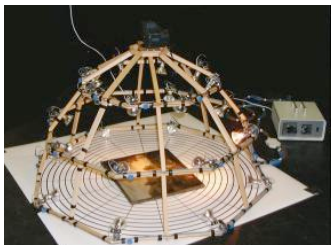
Frisian tablet. Original transcription by C.W. Vollgraff, *1917 -- "De tabella emptionis aetatis Traiani nuper in Frisia reperta"*

New transcription by A.K. Bowman, R.S.O. Tomlin, and K.A. Worp, *2009 forthcoming -- "Emptio bovis Frisica: the 'Frisian ox sale' reconsidered"*

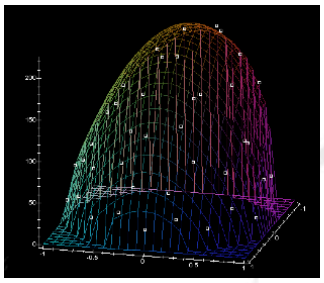
→ Digital pictures by Dr. C. Crowther

Specific application of the shadow-stereo principle

- 'Polynomial Texture Mapping'
  - Image collection (fixed camera, various light positions)
  - Summarise the information of all images in one *.ptm image:
    - Image model: *Image = L(l)xRGB*, where only *L* depend on the light position *l*
    - Polynomial interpolation of *L* to evaluate the image at an unknown light position l_u
    - Size of the *.ptm image is 3 times the size of one given image



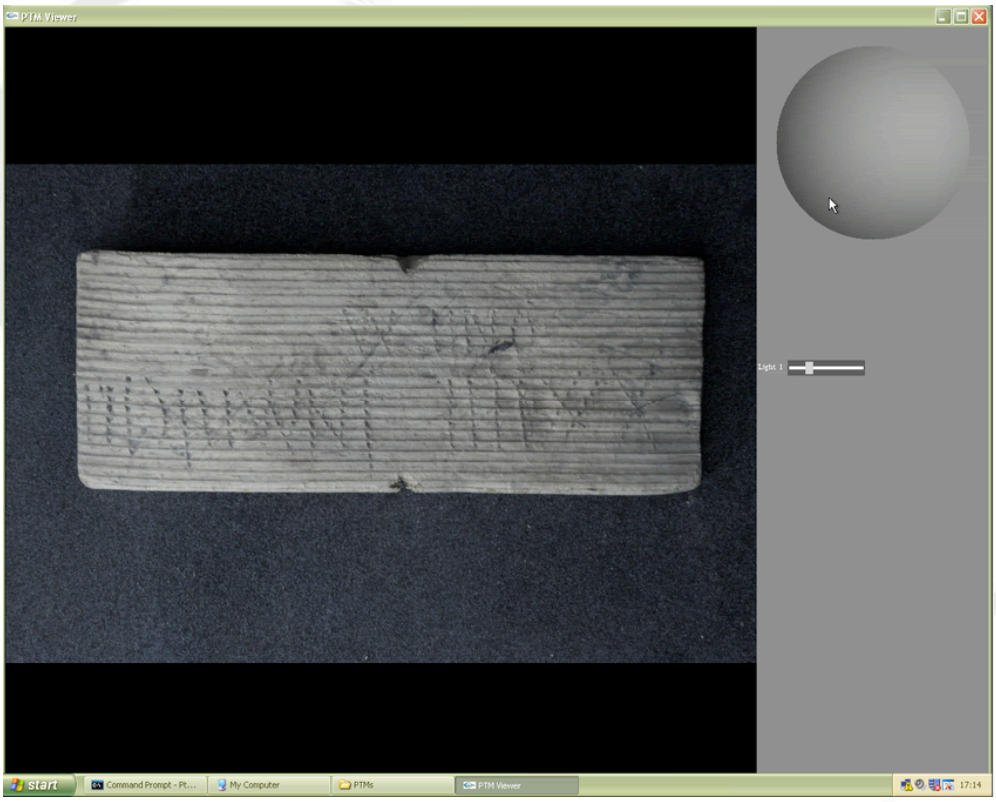
PTM setup used at the National Gallery in London for examining the surfaces of paintings



L as a polynomial surface for a given pixel Each point on the surface represents a value of *L* for a given light position.

Siggraph 2001; Malzbender et al.

→ Collaborative project with Dr. G. Earl (Southampton), for further capture of Vindolanda tablets at the British Museum



PTM viewer in action on a PTM image of a Vindolanda tablet imaged and created by Dr. Earl's team.

From pure signal image to image features

- Image enhancement
  - Correct background for better homogeneity and improved incisions visibility
    - Illumination correction
    - Woodgrain removal (for wooden tablets!)
- Feature detection
  - Identification of image features as potential ‘strokelets’ that are part of a character
    - Characterisation of a ‘strokelet’ as an incision

Illumination correction

- Homomorphic filtering

Original image

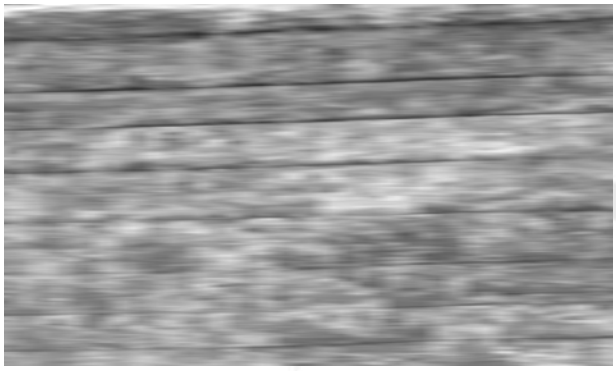
After illumination correction

Woodgrain removal

- Multiplicative removal

Original image

Woodgrain

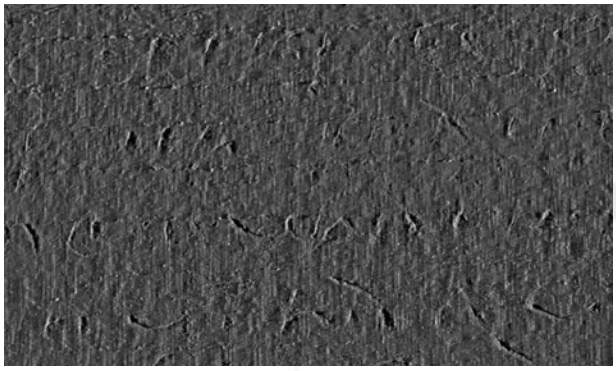


Woodgrain removal

- Multiplicative removal

Original image

After woodgrain removal



Feature detection

Two approaches:

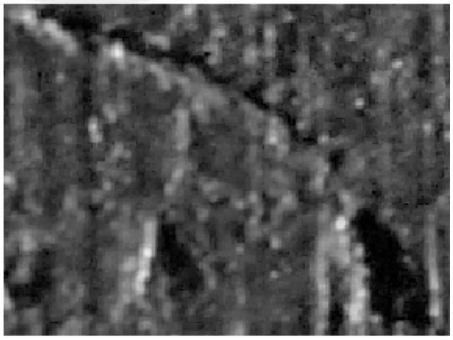
- Study images individually and then combine the processed images:
  - Phase congruency
- Use the combination of images:
  - Simple combination (after image enhancement)
  - PTM
  - Statistical information-theory based approach

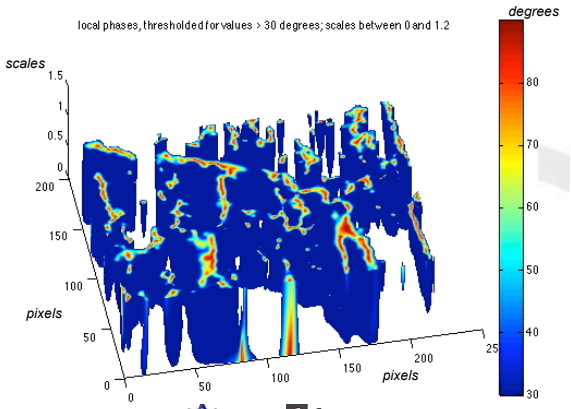
Single-image processing

- Feature detection via “phase congruency” (monogenic signal/Fourier analysis and scale-space)
  - Enables to classify pixels in an image regardless of the lighting conditions, and based on their “profile”
    - If they correspond to a feature
      - A type of feature (a value of phase)
      - An orientation of the feature

Single-image processing-2

- Some of the background texture has the same characteristics as the features
- Nice theory, but images **very** noisy





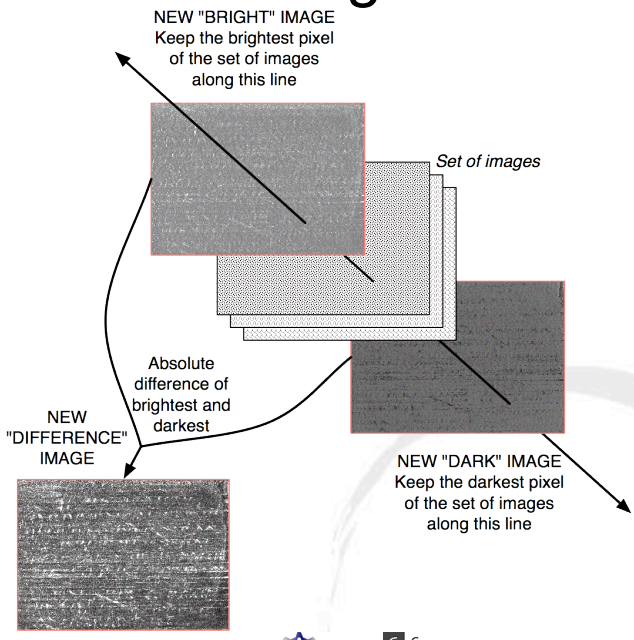
The challenge

- Images are extremely noisy, even after homogenisation of the background:
  - They are actually noisier than medical images on which this approach works!
  - Problem: how to remove noise without interfering with the contained information?
    - Normally, the answer would be: “Know what you’re looking for!”
    - However, here, it’s actually to better differentiate what we’re looking for from noise that we wanted to identify the noise in the first place!!!
    - ➢ Elements of an answer: the noise should not change in nature, from one image to the next, whereas the visual aspect of the incisions does

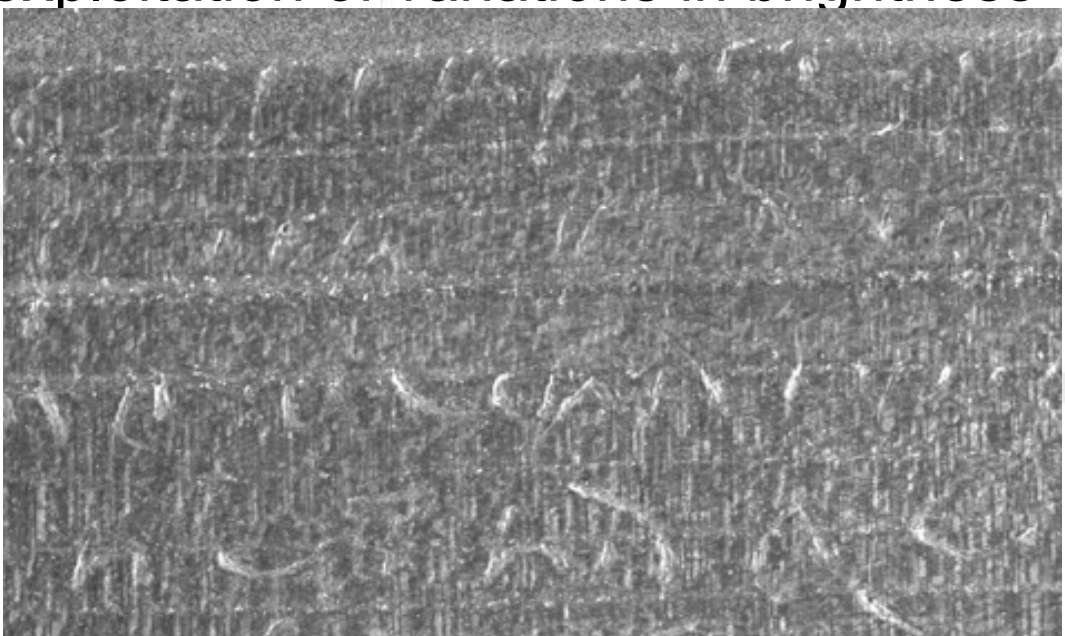
  - ➢ Not only for localising features is shadow-stereo useful, it will also allow to find similitude of noise in the images and identify those areas as background

Multiple images approach 1: naïve exploitation of variations in brightness

- Work with the original stack of images after they each have undergone image enhancement (illumination correction and woodgrain removal):
  - Noise and stains have similar brightness from one image to the next
  - Shadows and highlights of the incisions move according to the light position



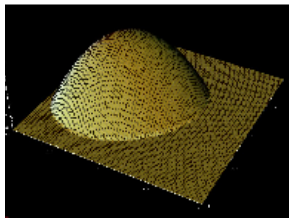
Multiple images approach 1: naïve exploitation of variations in brightness

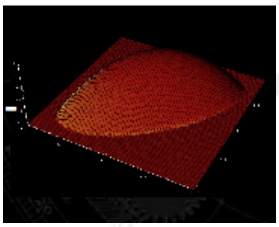


Multiple images approach 2: PTM

- Work with PTMs based on the image model
  Image = L(l)xRGB
  - At pixels corresponding to flat background: *L* is isotropic
  - At pixels corresponding to a groove or incision: *L* is anisotropic, skewed in the direction of the incision, and given by the principal directions of curvature at the apex of the surface *L*

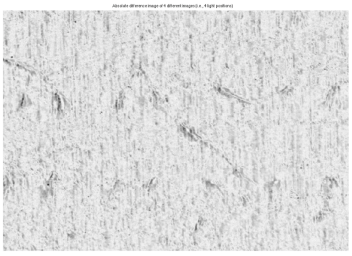
➢ Nice theory, but here again, the images are too noisy

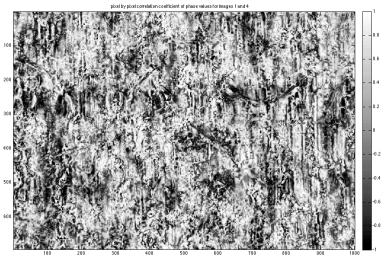




Other attempts at multiple images approach (*work in progress*)

- Combine results of single-image phase congruency method
- Use statistical methods such as local correlation or information theory (local joint entropy) to identify the most dissimilar areas in two images as likely locations for incisions



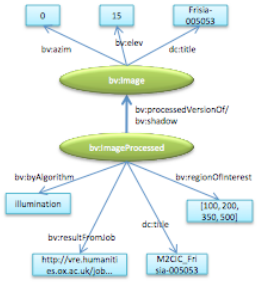


Dissemination

- Offer the background correction algorithms as functionality from within a portlet in the VRE-SDM (BVREH) application
  - Illumination correction
  - Woodgrain removal
  - Dual display
  - Expanded the BVREH data model to handle processed images and keep track of their provenance

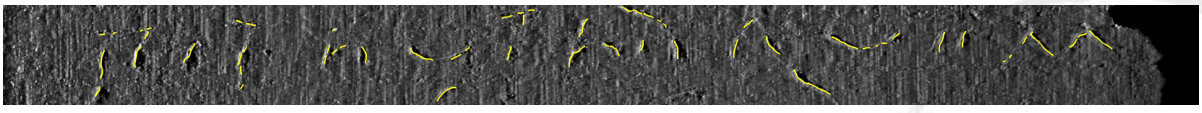
→ BVREH poster
→ Ruth, John and Pin's demos
→ ENGAGE eSAD+VRE-SDM poster

- Enable use of Grid resources (NGS) for heavy computations
  - Project supported by the OMII-UK under the ENGAGE umbrella

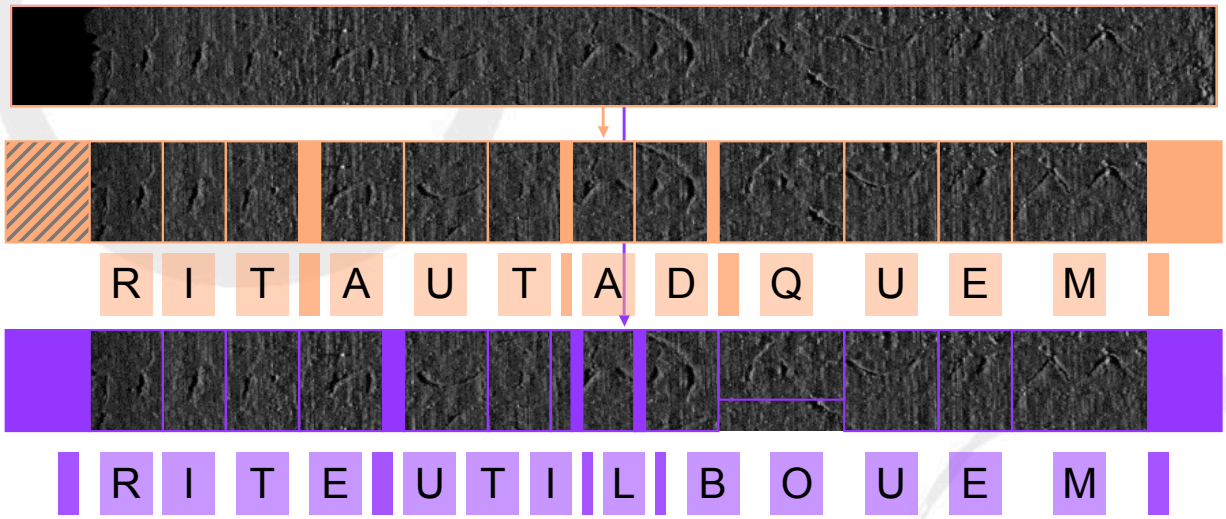


Conclusion

- Background correction algorithms available as web-based interface through VRE-SDM
- Still some work to do to be able to extract features that are character strokelets!! …
  - Exploit better the multiple light positions and combine the images!
  - Strokelets are the input for the next level of reading… when debates can start as to which strokelet is part of which character



Example of two groupings of strokelets yielding two different interpretations



→ Mike and Henriette's talks

GRATIAS VOBIS AGO PRO VESTRI INTENTIO

http://esad.classics.ox.ac.uk