Automatic image network design leading to optimal image-based 3D models

Enabling laymen to capture high quality 3D models of Cultural Heritage

Bashar Alsadik & Markus Gerke, ITC, University of Twente, NL-Enschede

b.s.a.alsadik@utwente.nl; m.gerke@utwente.nl

Content

1. Motivation
2. Automatic camera network design
3. Automatic guidance of camera operator
4. Robust image correspondences
5. Evaluation
6. Conclusions & Outlook
7. References

Motivation

Aim: Create highly detailed and accurate 3D models of cultural heritage by using imaging techniques ➔ Image-based modeling (IBM).

IBM comprises:

- -Network design (needs experience), mostly manual still
- -Image orientation (exterior – interior) (Automatic)
- -Image matching/surface reconstruction etc.

Motivation

If the *network design* is automated one can expect that subsequent 3D modeling…

- uses not more images (and not less) than needed for a pre-specified accuracy
- can easily be done even by nonprofessionals
- becomes potentially real time
- will be suitable for autonomous vehicles

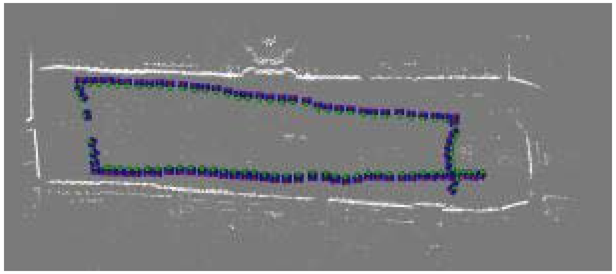
Further: bundle adjustment can make use of knowledge on initial camera position

Motivation

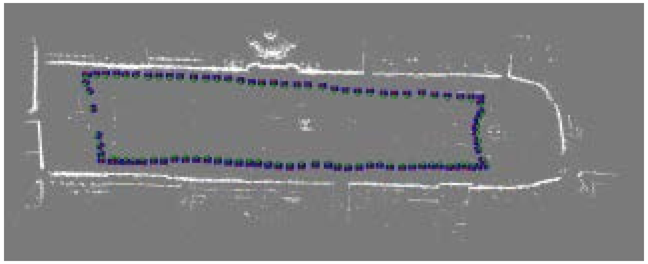
- Why we need planning? Why don't just take a bunch of images and process automatically by C.V. approaches?

- Accuracy for C.H application (not just a visualization).
- To have a complete model.
- Satisfy the project requirements, eg concerning the needed geometric and radiometric details.
- Avoid the failure of the image orientation by SfM and\or wrong or deteriorated network.

http://3dom.fbk.eu/sites/3dom.fbk.eu/files/pdf/Remondino_etal_Euromed2012_preprint.pdf



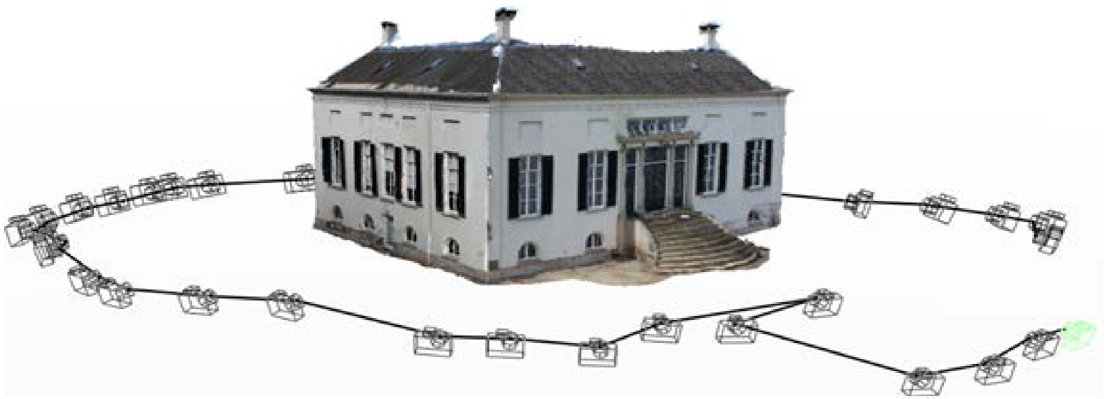
SfM



Photogrammetric based approach

Motivation

In cultural heritage we have two major "large" object types:



Buildings (can be modeled from inside or outside)

Monuments and statues

Motivation

In this presentation: three steps

1) Automatic design of camera network

2) Guidance of the camera operator to the optimal location

3) Robust image correspondences

➔ For both: building and statues

Automatic camera network design

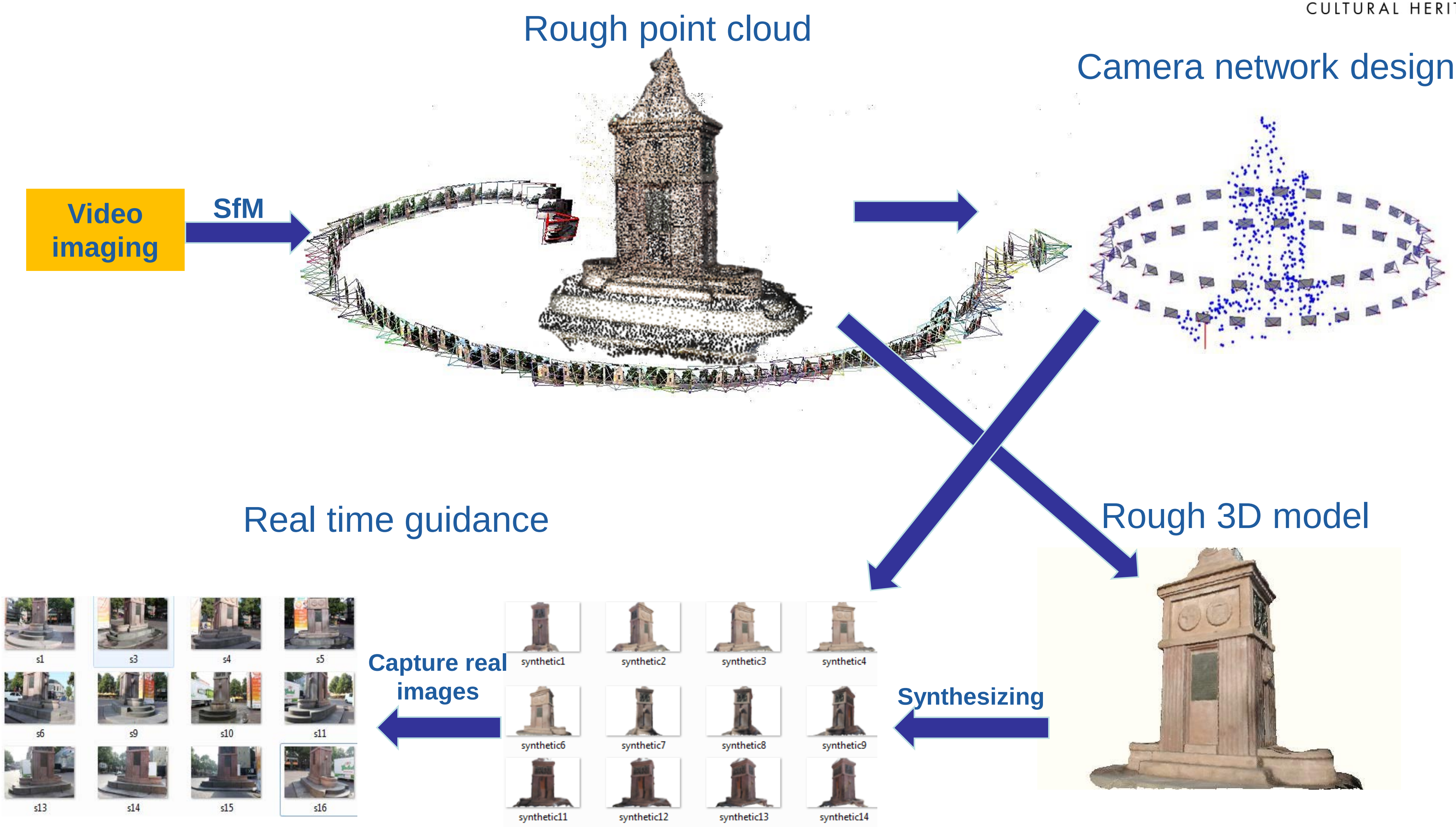
Problem :
Where to place the cameras?
How many cameras?

Accuracy and coverage

Methodology

a)Create initial point cloud
b)Design a dense initial camera network
c)Filter the dense camera network
d)Optimize (if needed) the filtered camera network
e)Guiding procedure
f) image correspondences matching + orientation
g)3D modeling

Methodology

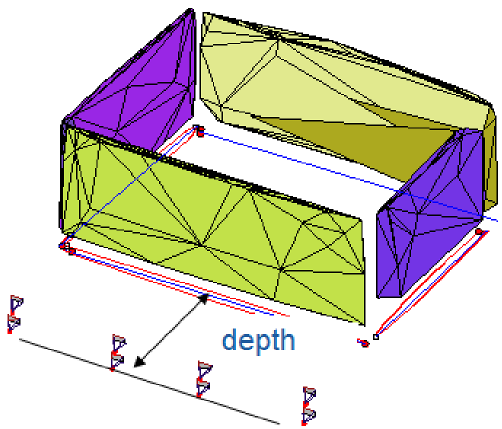


Automatic camera network design

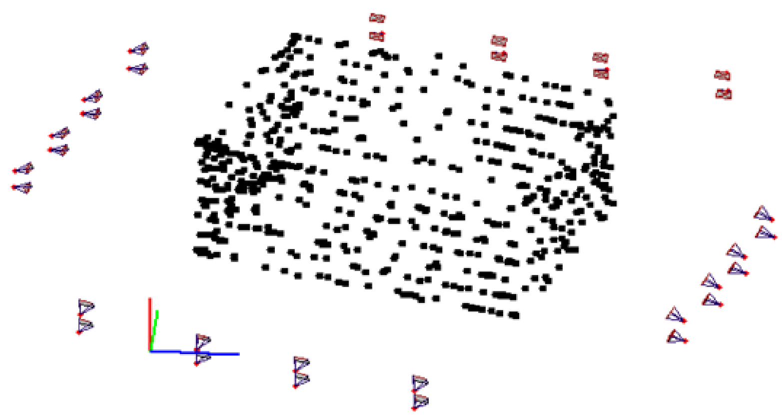
a) Design a *dense* initial camera network

Put a regular planar or circular grid around the object and design cameras, according to pre-defined accuracy demands

$$\sigma_{X,Y} = \frac{q.S.\sigma_{x,y}}{\sqrt{K}}$$



60% endlap
20% sidelap
Camera specification

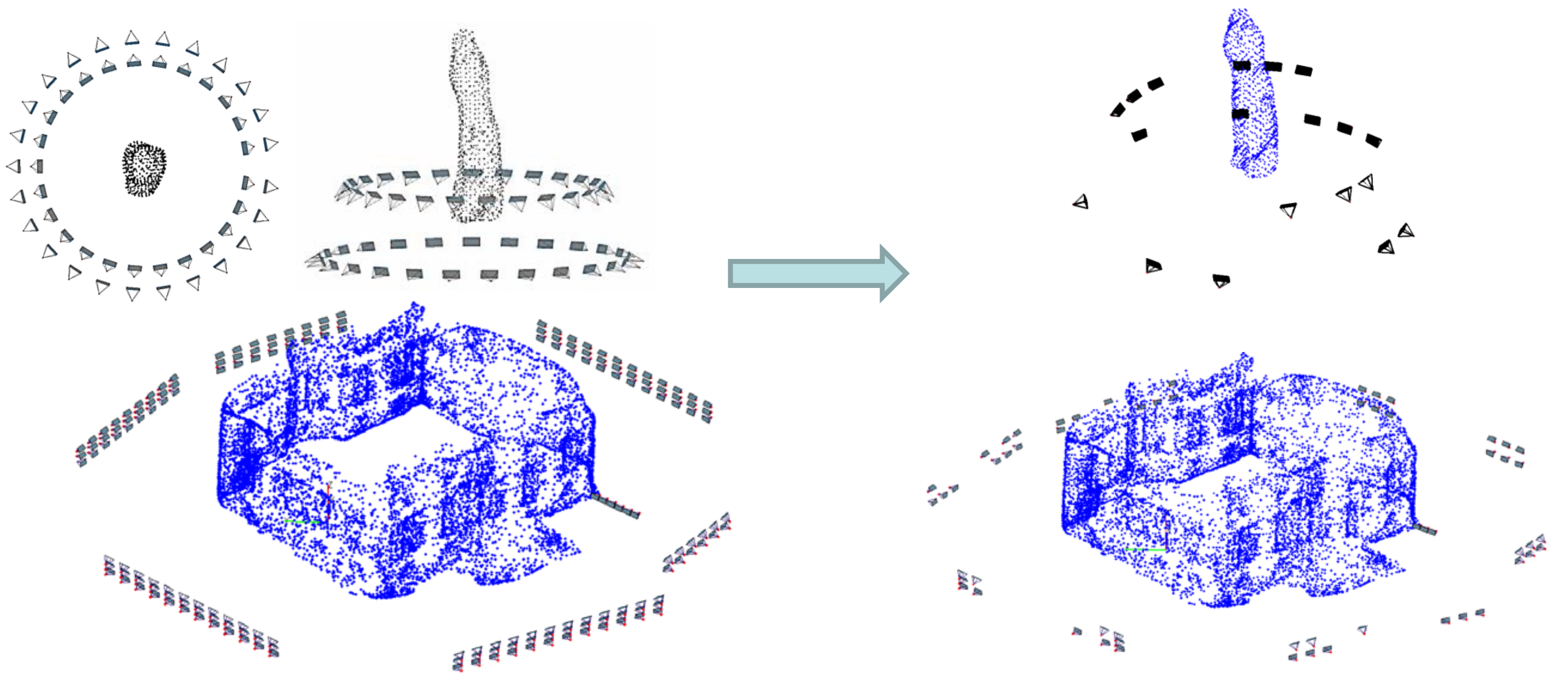


The initial exterior orientation parameters $\omega_i{}^o, \varphi_i{}^o, \kappa_i{}^o, Tx_i{}^o, Ty_i{}^o, Tz_i{}^o$ for each camera i.

Automatic camera network design

b) Filter the dense camera network

➔ Iteratively filter out cameras which only provide "useless" redundancy in sense of coverage and accuracy



Guided image capture

So far we only talked about *planned images*, based on the initial point cloud

➔aim now: help the camera operator capture the designed high resolution images in reality:

a) Create *synthetic images*, based on initial video images, point cloud, and computed optimal camera locations

b) Use those synthetic images *to guide* the camera operator to the correct location, let him take the image

c) *Check* if taken image corresponds to the planned image

Guided image capture

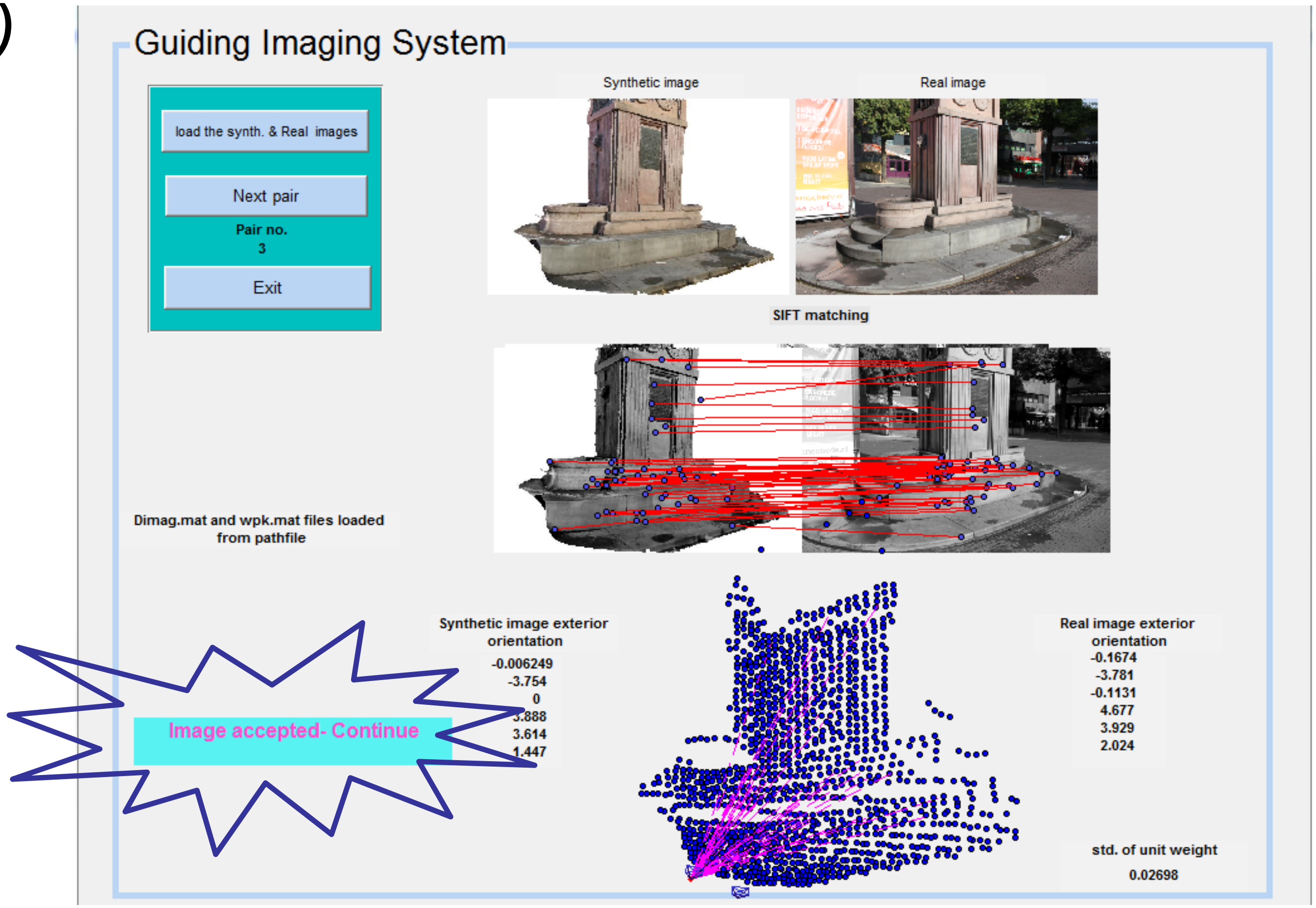
a) Create synthetic images, based on initial video images, point cloud, and computed optimal camera locations

COSCH
COLOUR & SPACE IN CULTURAL HERITAGE

Guided image capture

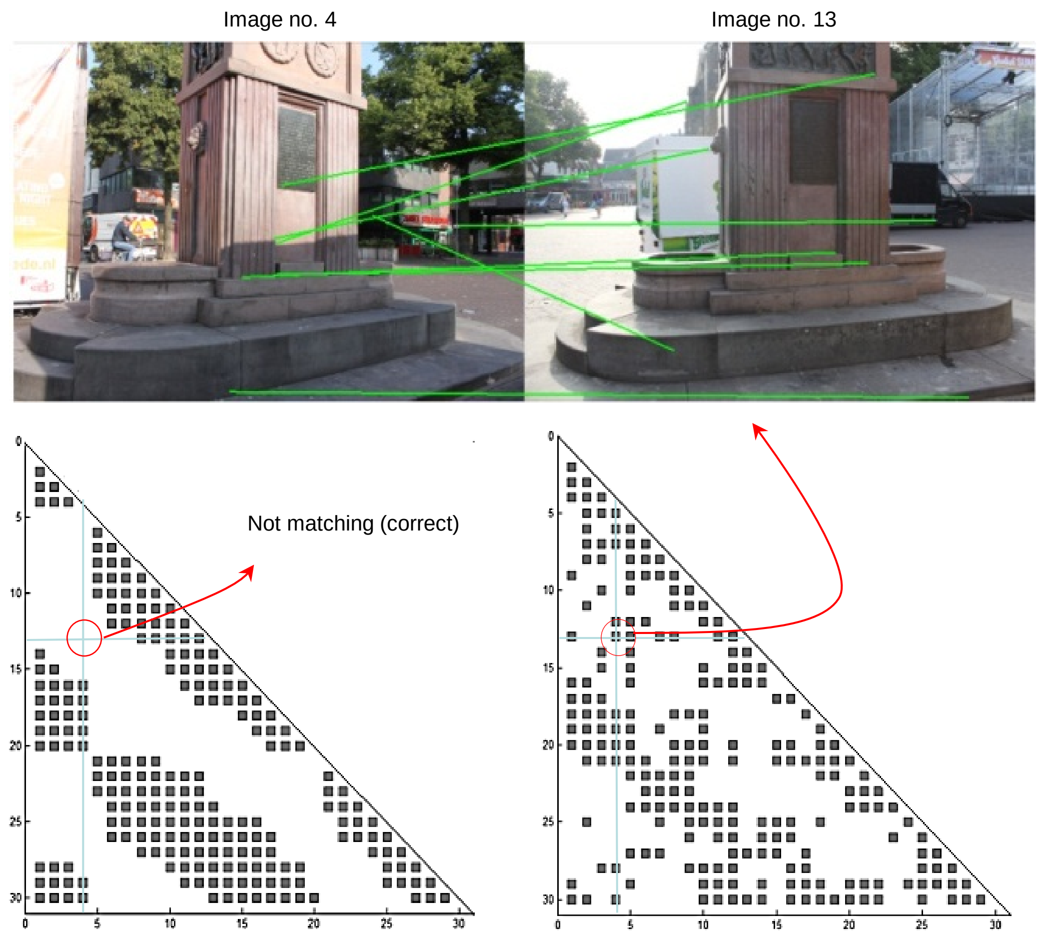
b) Use those synthetic images to guide the camera operator to the correct location, let him take the image

(prototype)



Robust extraction of image correspondences

Motivation:
- Reduce wrong matches by avoiding to compare images which do not show the same part of the object
- Especially at buildings/statues (repetitive patterns)
- Exploit knowledge about approx. 3D Model and camera locations for robust correspondence search

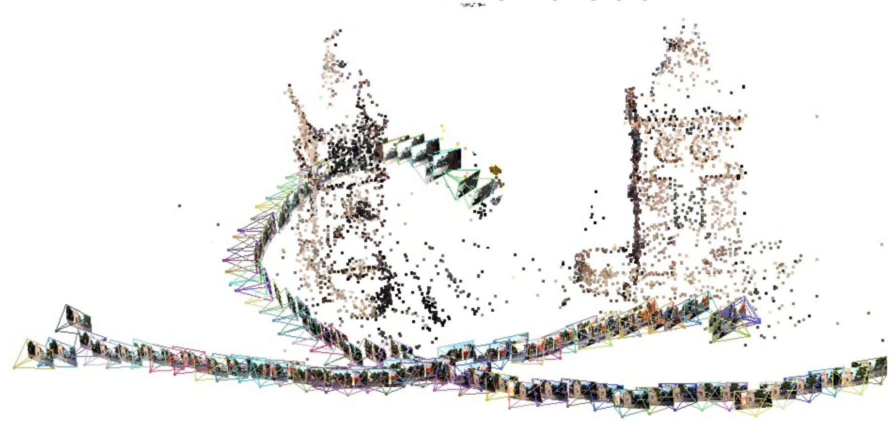


Full pairwise versus guided matching

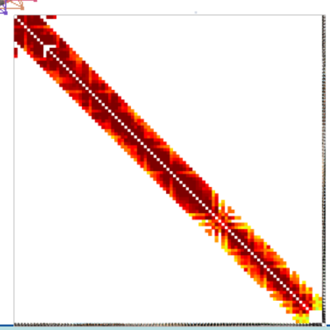
COLOUR & SPACE IN CULTURAL HERITAGE

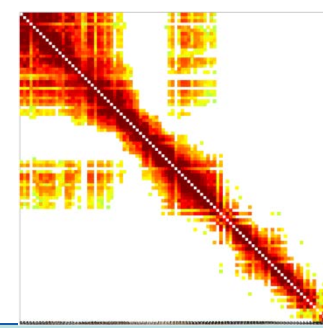
- Video stream (1920x1088) → 96 frame
- SFM with SIFT GPU.

Full pair-wise:
orientation unsuccessful
due to mismatching
826 sec.



Guided matching consumes
319 sec + orientation successful



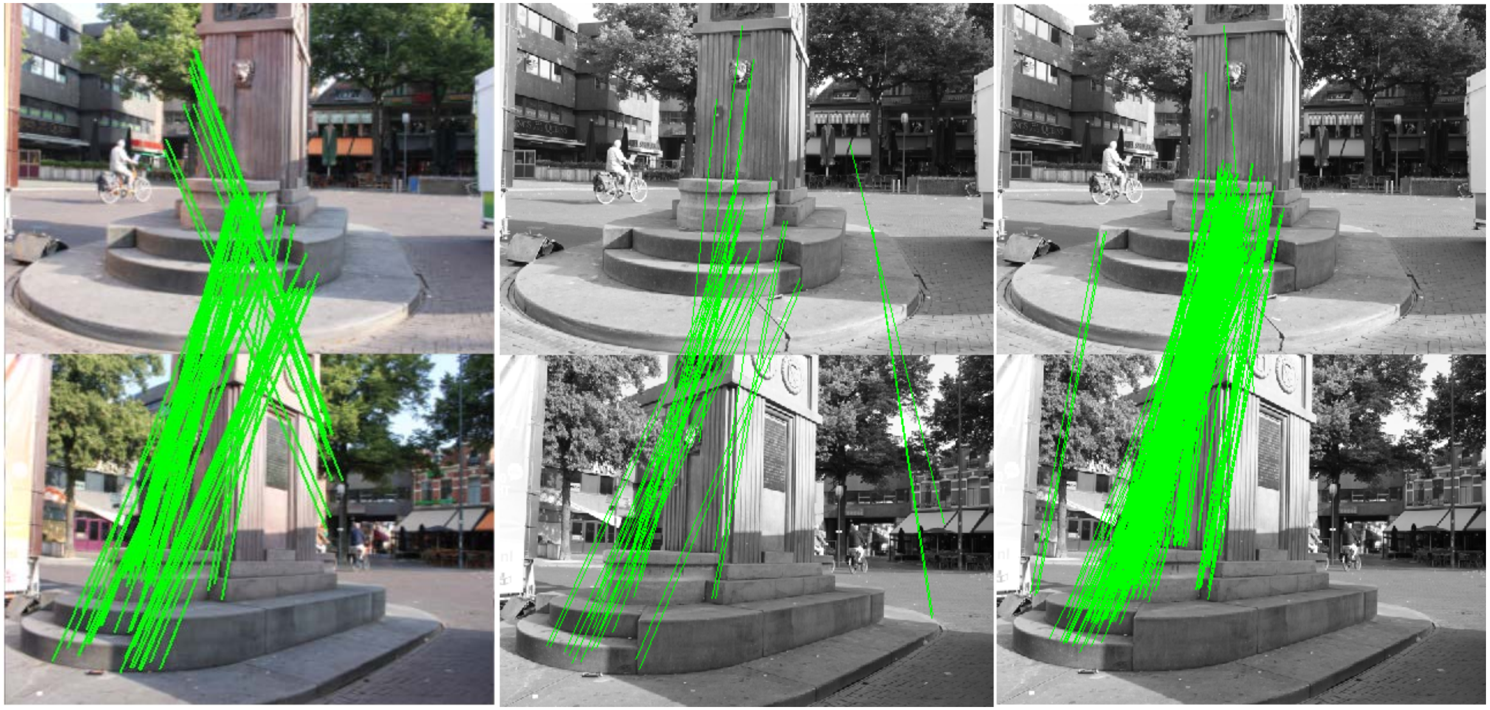


COSCH
COLOUR & SPACE IN CULTURAL HERITAGE

comparison
State – of – the – art

Wide based imaging + moving objects in background: challenge

Vsfm : 66/145 | Photomodelr 38/41 | Guided method 285/1138



Evaluation

Most interesting question for now: Does the thinned network comply to pre-defined accuracy demands?

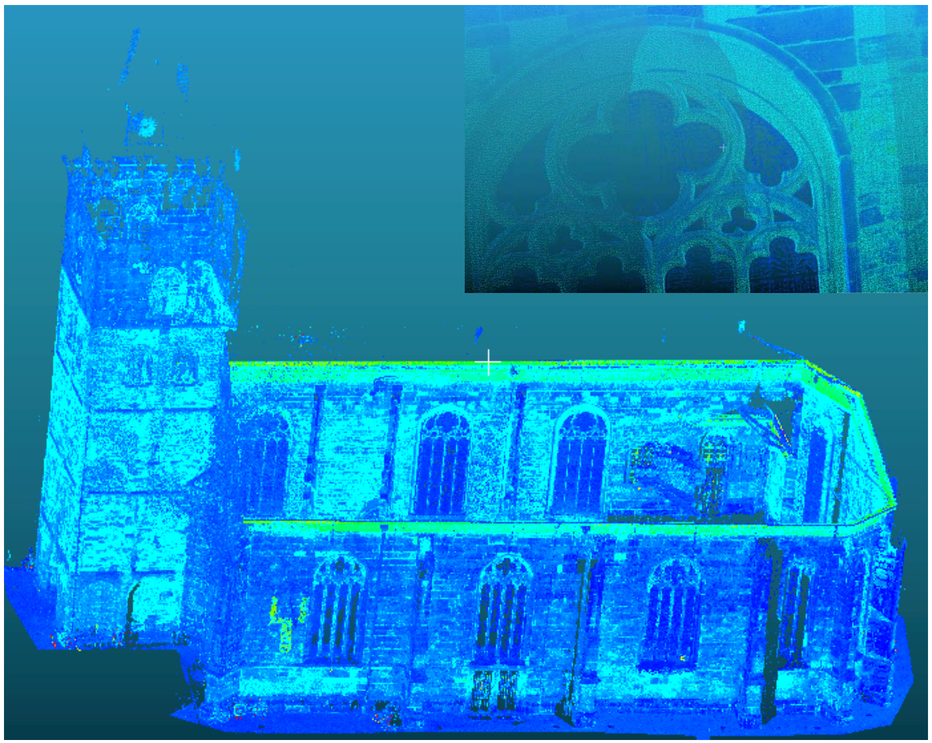
Experiment: the Grote Kerk in Enschede – Netherlands.

Reference: high accurate survey with a terrestrial laser scanning TLS - with “TRIMBLE CX SCANNER” by www.landmeetdienst.nl company

The manufacturer single point accuracy standards are position = 4.5 mm @ 30 m; 7.3 mm @ 50m distance = 1.2 mm @ 30 m; 2 mm @ 50m

Evaluation

- 23309511 points – church object



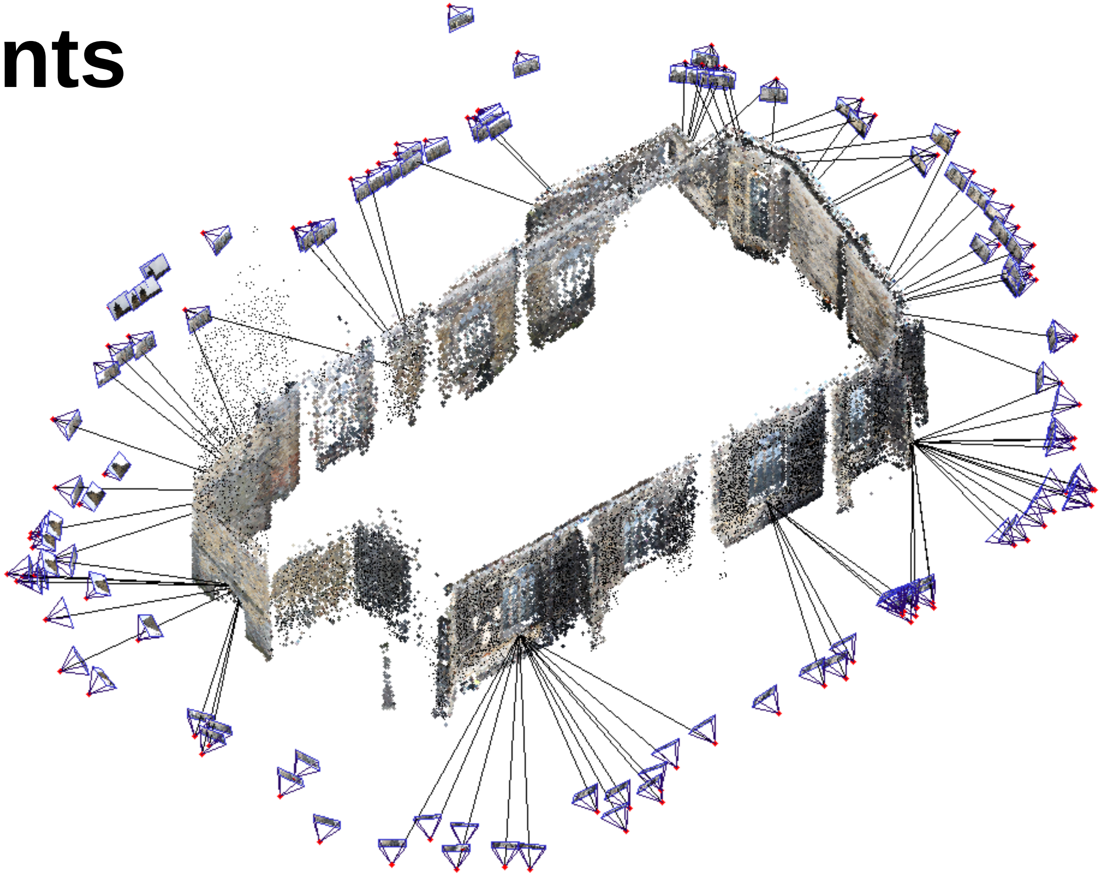
Evaluation

- **Target control points**

Point	X	Y	Z	Error
Pt-24	1000.007	20015.5	302.628	0.014
Pt-51	976.074	20026.1	302.56	0.011
Pt-66	987.689	20056.6	302.665	0.015
Pt-16	995.682	20038.8	302.675	0.011

RMSE(gcp) ≈1.2 cm

RMSE(check)≈6.5 cm



GCP and check points distribution

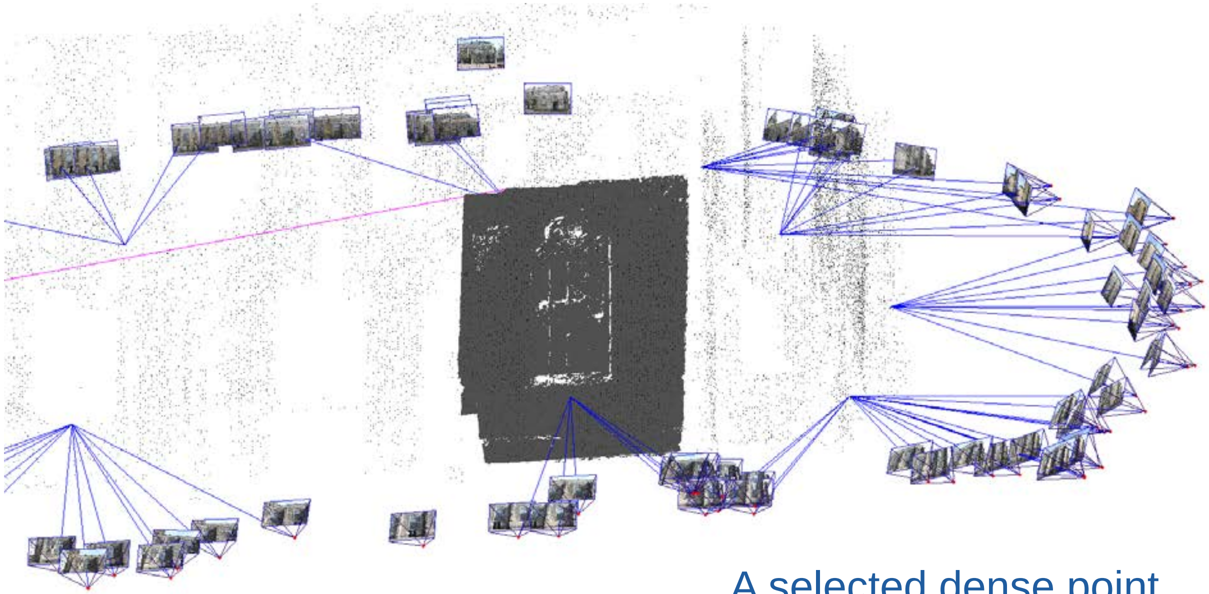
3D model – Point cloud

3D dense point cloud of the church.

-Is it sufficient and accurate for cultural heritage documentation and preservation?

Need evaluation!

Dense point cloud- evaluation



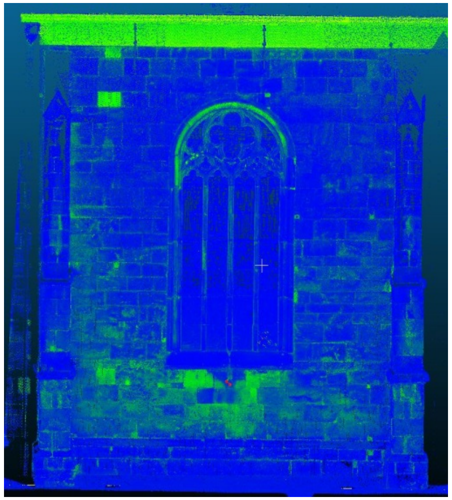
A selected dense point cloud patch for the comparison and evaluation with reference to the laser scan survey.

Dense Image Matching- SURE (Open source)

http://www.ifp.uni-stuttgart.de/publications/software/sure/index.en.html

Cloud comparison

TLS – 4144647 points

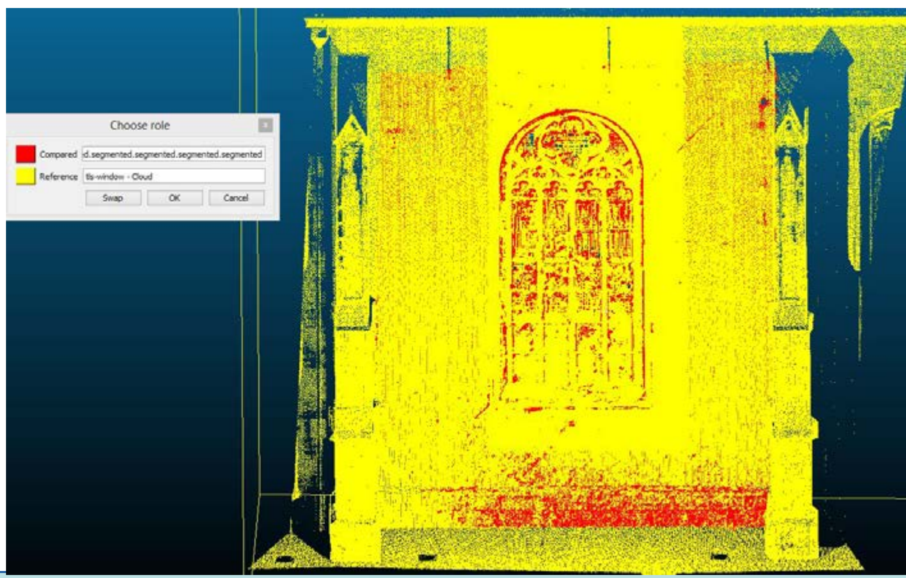


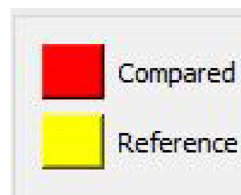
2980117 points

The comparison between two dense clouds created by the laser scan (Reference) and the image dense matching. Using

Cloud compare- open source

http://www.danielgm.net/cc/





Cloud compare – C2C

The graph shows the total number of points differences in a histogram plot.

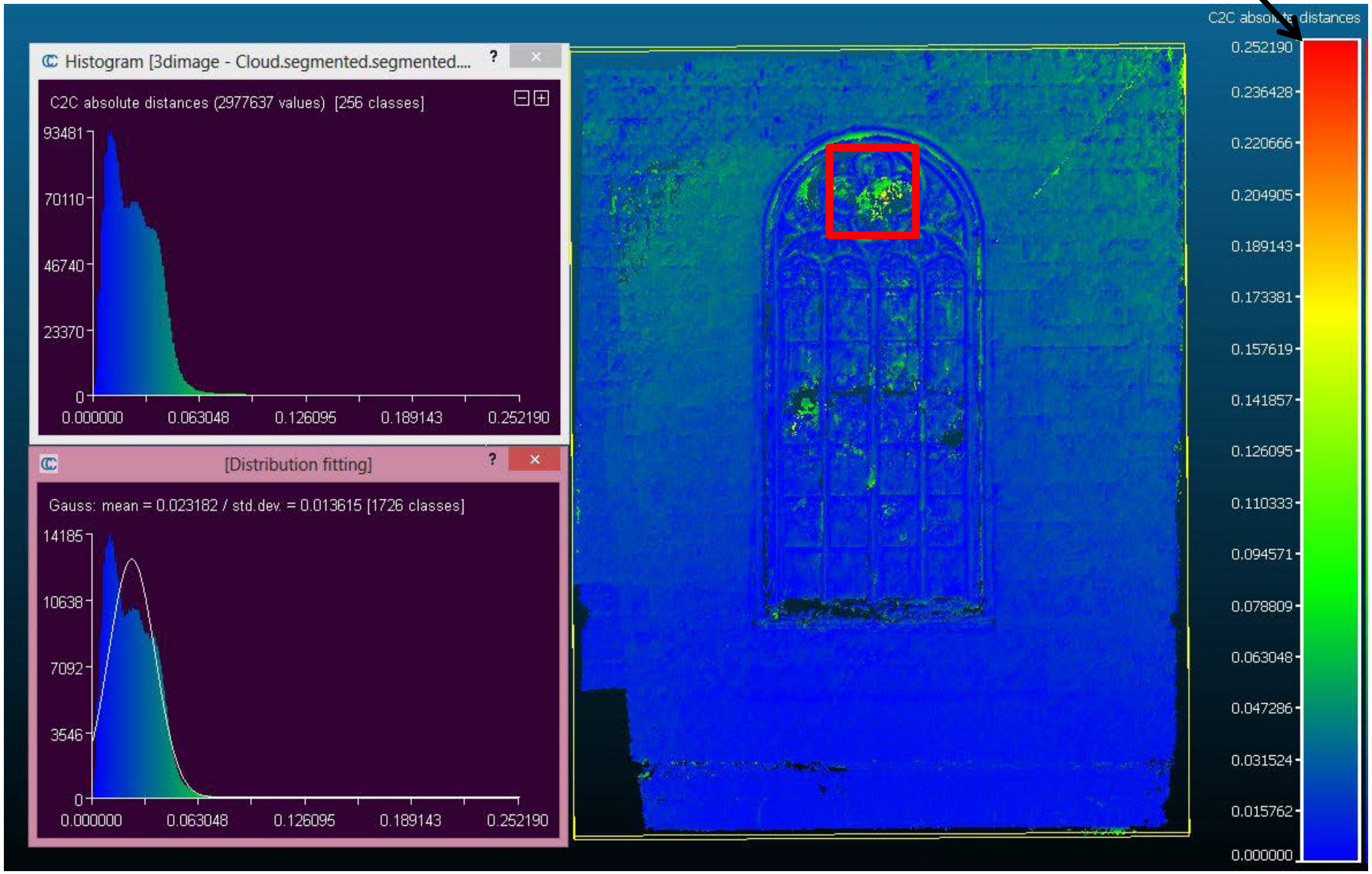
And the Gaussian curve fit which results in:

Mean distance = 2.1 [cm]

Std. deviation = ±1.4 [cm]

- The largest difference is in a small part of the window. This is expected because of the non-texture area effect.

Scale of shift [m]



Conclusions, current and future work

- An efficient and practical automated workflow for image network planning is introduced
- Experiments show that filtered network compares to fully planned photogrammetric network
- An efficient matching approach which exploits the pre-knowledge is shown
- Currently we work on gap identification in dense matching and accuracy evaluation/improvement

Outlook:

- Further: Test approach using an Unmanned Aerial Vehicle (UAV)

Thanks for your attention!

Questions?

References

Alsadik, B.S., Gerke, M. and Vosselman, G. (2013) Automated camera network design for 3D modeling of cultural heritage objects. In: Journal of Cultural Heritage, (2013)IN PRESS, 12 p.

Alsadik, B.S., Remondino, F., Menna, F., Gerke, M. and Vosselman, G. (2013) Robust extraction of image correspondences exploiting the image scene geometry and approximate camera orientation. In: ISPRS Archives Volume XL-5/W1 : Proceedings of 3D-ARCH 2013 : 3D Virtual Reconstruction and Visualization of Complex Architectures, 25-26 February 2013, Trento, Italy

Alsadik, B.S., Gerke, M., Vosselman, G., Daham, A. and Jasim, L. (2013) Minimal camera network for the 3D image-based modeling of cultural heritage. Submitted to Survey Review Journal