Institute for Spatial Information
and Surveying Technology
Mainz University of Applied Sciences

COLOR & SPACE IN CULTURAL HERITAGE

Region Based Fusion of 3D and 2D Visual Data for Cultural Heritage Objects

Robert Frohlich[1], *Levente Tamas*[2], *Alain Trémeau*[3], *Zoltan Kato*[1]

[1]*University of Szeged, Hungary*
[2]*Technical University of Cluj-Napoca, Romania*
[3]*Laboratoire Hubert Curien, Université Jean Monnet, Saint-Etienne, France*

Introduction

COLOR & SPACE IN CULTURAL HERITAGE

- Recently a lot of new devices are available both for 3D sensing and 2D imaging domains.
- These range from the more affordable, consumer oriented to the highly professional very expensive equipment.
- Since there is no ONE device that can “do it all”, the use of multiple acquisition methods is usually necessary.
- The fusion of the captured multimodal data poses a big problem.

We show a **processing pipeline** that:

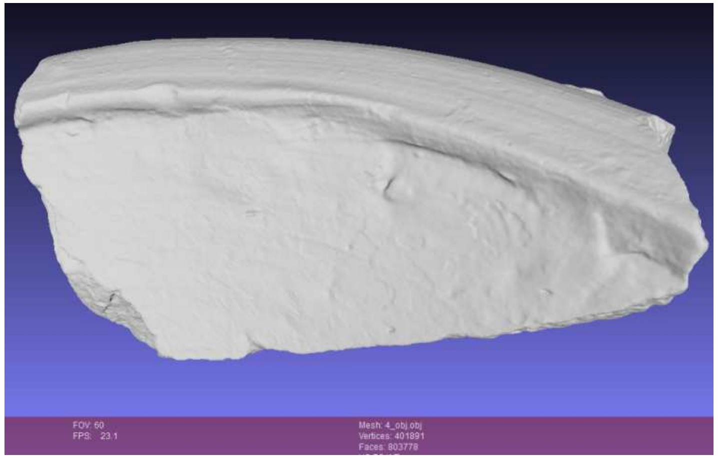
- uses a different approach than State-of-the-Art solutions,
- can be applied for various 3D pointcloud data, and 2D images,
- does not necessarily rely on intensity information of the 3D data,
- uses a novel region based registration algorithm, and an ICP alignment step for parameter refinement.

Problem statement

COSCH
COLOR & SPACE IN CULTURAL HERITAGE

Why is this needed?

- Some 3D scanners have an RGB camera integrated
- These usually have relatively small sensors, don't provide high resolution detail
- Relying only on this RGB information is not satisfactory for CH application
- Other 3D scanners may provide 3D data with no RGB information at all
- Any DSLR camera can capture RGB information in much higher detail


The presented ceramic fragment belongs to the MOM (i.e.. © Maison de l'Orient et de la Méditérannée).

Solutions

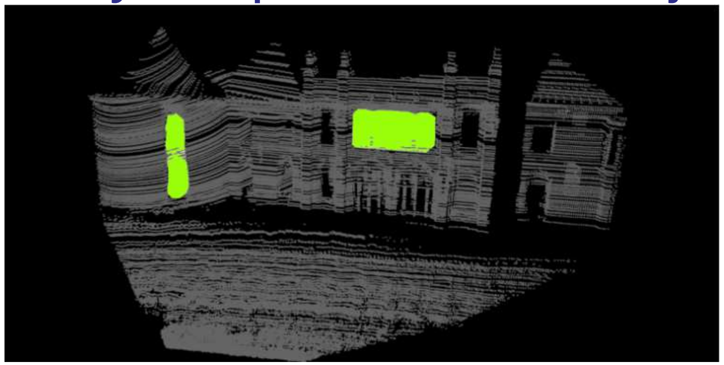
- A good approach is to **fuse the 3D pointcloud** data, captured by any means available, with high quality, possibly color calibrated **2D images**.
- Typically most State-of-the-Art methods rely on extracted point features, but it can be hard to find such descriptors if the surface properties of the object don’t satisfy certain conditions.
- The use of markers can be helpful, if it is possible at all.

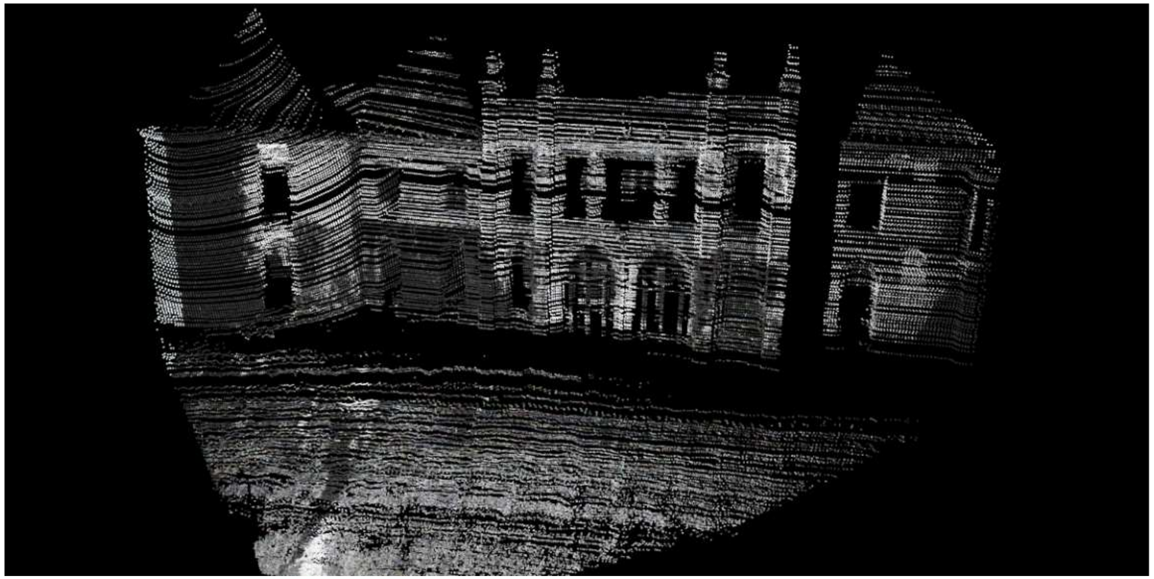
Our solution

- We use segmented smooth regions of the object, instead of key-points or markers.
- These may be easier to detect or to manually select in case of some unique or special objects.
- The regions may also come from holes in the 3D data (see windows on the next slide).

Previous work with our solution

Our algorithm was primarily designed and used for urban scenes, but it is easily adaptable for CH object and scenes of different scale.





- L. Tamas and Z. Kato, "Targetless Calibration of a Lidar - Perspective Camera Pair," in *Int. Conf. on Computer Vision,* Sydney, Australia, December 2013, pp. 668–675.
- Zoltan Kato and Levente Tamas. "Relative Pose Estimation and Fusion of 2D Spectral and 3D Lidar Images". In *Proc. of the Computational Color Imaging Workshop* (CCIW), Lecture Notes in Comp. Sci., Vol. 9016, Saint-Etienne, France, pages 33-42, March 2015. Springer.

Outline of what's coming

COLOR & SPACE IN CULTURAL HERITAGE

- Data acquisition
- Pipeline:
  - Segmentation
  - Registration
  - ICP refinement
  - Evaluating the registration results
  - Data fusion
- Results
- Conclusions

Datasets

We used 3D data produced by two different methods. First some ceramics captured with the **Artec Spider** 3D scanner, and then the Chinese warrior figurine captured with a **structured light scanner**. While the first may be considered a relatively cheap, easy to use device, the second was a much more professional solution, that requires special laboratory conditions.

Different quality data:

- While the Artec solution produces a generally good looking model, it lacks the details, the colors and textures are washed out and blurry.
- The structured light scanner on the other hand produced a much more detailed, high resolution model, with better color as well.

Data acquisition

Artec scanner - its advantage is also its weakness.

- It is a structured light scanner, with blue LEDs projecting a speckle pattern, but it is also a handheld device, that was designed to be easy to use.
- While it can capture easily a view, merging the different slices captured by the moving device is difficult. It needs a precise alignment of some common reference points. If detecting such points is not possible, or not precise enough, then the software will fail to produce results.

Main problems are with reflective surfaces, with homogeneous, texture-less surfaces.

Markers could help. In our acquisition test we didn't place any markers on the object itself, instead we placed the object on a well textured surface.

Data acquisition

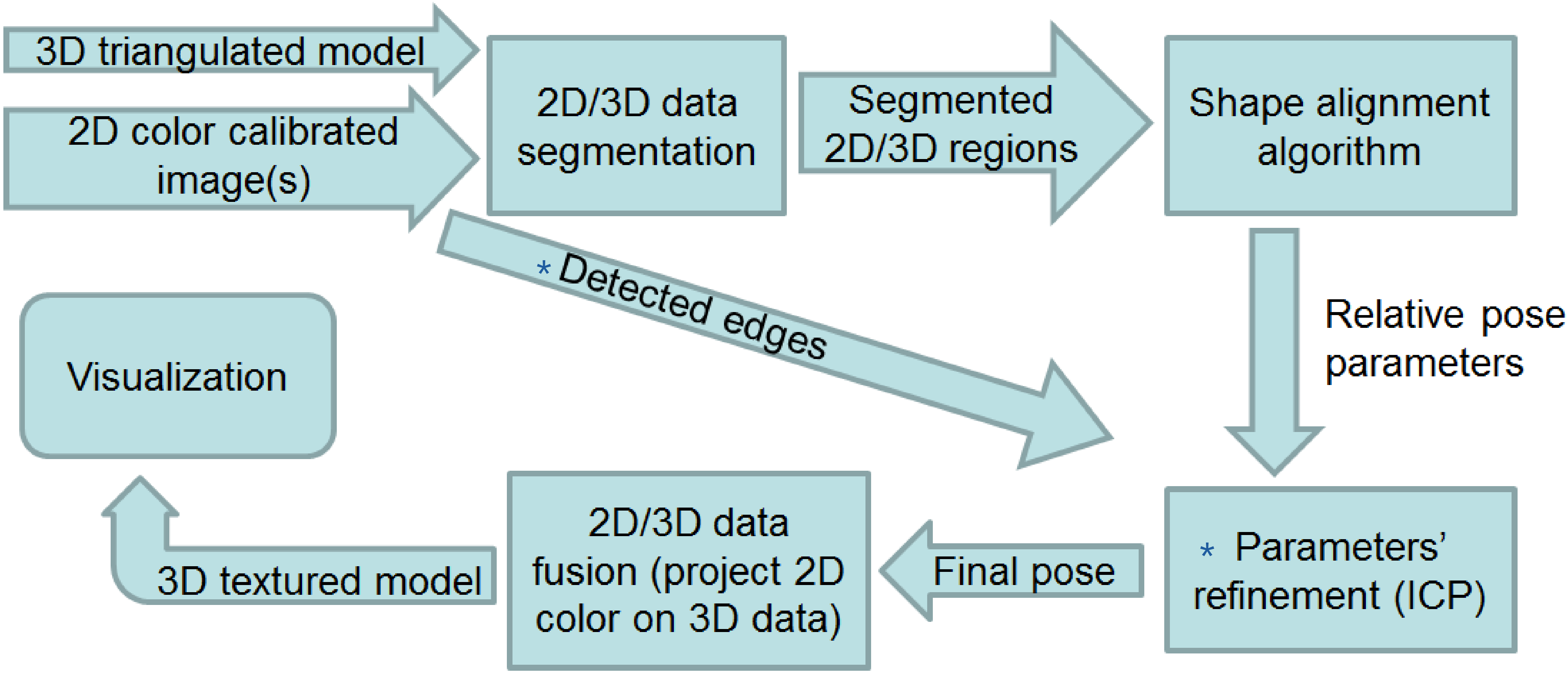
2D images were captured with commercial **DSLR cameras**.

Useful tips:

- long exposure, using a tripod and remote control or delayed exposure
- high aperture values avoid the depth of field effect, resulting in more sharpness over the whole image
- use multiple light sources if available, to avoid unwanted shades, and also to create useful shades
- shooting in RAW gives better control when correcting lens distortion and other unwanted effects

Pipeline

- The processing workflow that we propose has been applied for small ceramic fragments and objects with promising results.



Data processing

Segmentation

Automatic and **semiautomatic** segmentation algorithms are available for 2D and 3D, but we have to choose the best solution according to our objects surface properties. If the available specific segmentation methods are not suitable for our data, we can use assistive tools to **manually** select regions.

- In 3D we used the *Z-painting* tool in *Meshlab*, that makes it easy to select the desired surface by painting with a brush over it.
- In 2D we used the *Fuzzy selection* tool in *Gimp*, that applies an intelligent region growing algorithm based on the pixels color information.

Useful tips:

- Depending on the 3D data we have (pure geometric, or also including intensities), different approaches have to be taken while selecting the appropriate surface regions.

Data processing

Useful tips:

If we **don't have** color information, then usually well contoured, smooth surface patches, dents, holes, hollows can be reliably segmented, as these probably are also visible on the 2D images (1st image). Different lighting setups can also emphasize these regions on the images, a perfect lighting with no shades is not always the best solution.

If we do **have** color, even if it's low resolution, we can trust the devices internal calibration, that means the color information are well aligned with the 3D points, which can be useful. In this case we can select regions based on their color, which gives a slight advantage in segmentation (2nd image).

Data processing

COLOR & SPACE IN CULTURAL HERITAGE

Useful tips:

- Compared to the classical solutions that rely on key-points, region segmentation may be easier to manually perform, since the human observer can better discern bigger shapes than some specific points on pixel level. If using relatively large regions, then segmentation errors are also well tolerated.
- It has a clear advantage in case of textureless objects, which have some usable geometric properties (some holes, dents, edges, etc.).
- It may not be possible to use in case of objects with a rough surface, where no small region can be selected without self-occlusion of the data.

Data processing

Registration

- Given a corresponding set of segmented 2D-3D regions, our algorithm can automatically estimate the relative pose of the camera to the 3D coordinate frame.
- Having the internal projection parameters of the camera (calibration can be done separately with a free toolbox) we estimate the rotation and translation values that produce the best overlap of the projected 3D regions and the 2D segmented regions.

Technical description of the algorithm is presented in our paper, accepted for publication at the *ICPR 2016* conference.

Robert Frohlich, Zoltan Kato, Alain Tremeau, Levente Tamas, Shadi Shabo, and Yona Waksman, "Region Based Fusion of 3D and 2D Visual Data for Cultural Heritage Objects," in *International Conference on Pattern Recognition*, Cancun, Mexico, Nov. 2016

Data processing

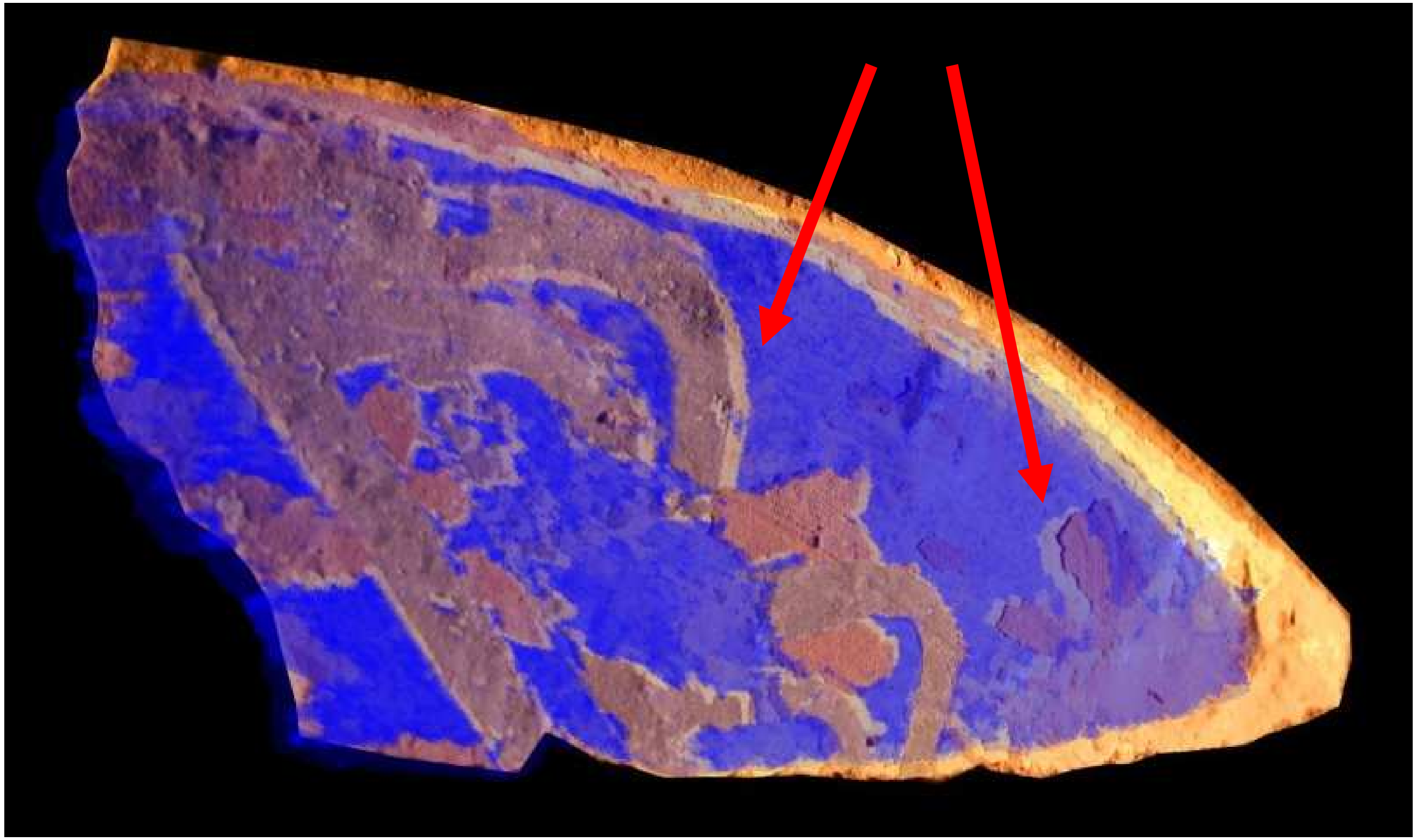
ICP refinement

- To avoid the errors introduced by the user's manual segmentation, we propose a parameter refinement step, that is based on the *Iterative Closest Point* algorithm. This step can only be applied **for color 3D data**.
- It uses the edge map of the projected 3D data and the 2D image.
- Edge detection is an automatic step, but needs some simple parameters, that were set experimentally.
- This step by itself wouldn't be able to solve the pose estimation problem, because ICP usually needs a good initialization. But since our algorithm can serve a good estimated pose, the ICP can further refine it.

If **no color** is available in the 3D data, our registration algorithm still provides a relatively good result, but that may not be suitable for professional documenting purpose, but only for visualization applications.

Evaluating the results

Data imperfections – We have found, that data captured with the handheld scanner may have some imperfections. Observe as the yellow regions of the 3D object are projected in blue on top of the original image, there is some misalignment of texture edges, that is not possible to correct by pose.



COSCH

COLOR & SPACE IN CULTURAL HERITAGE

Evaluating the results

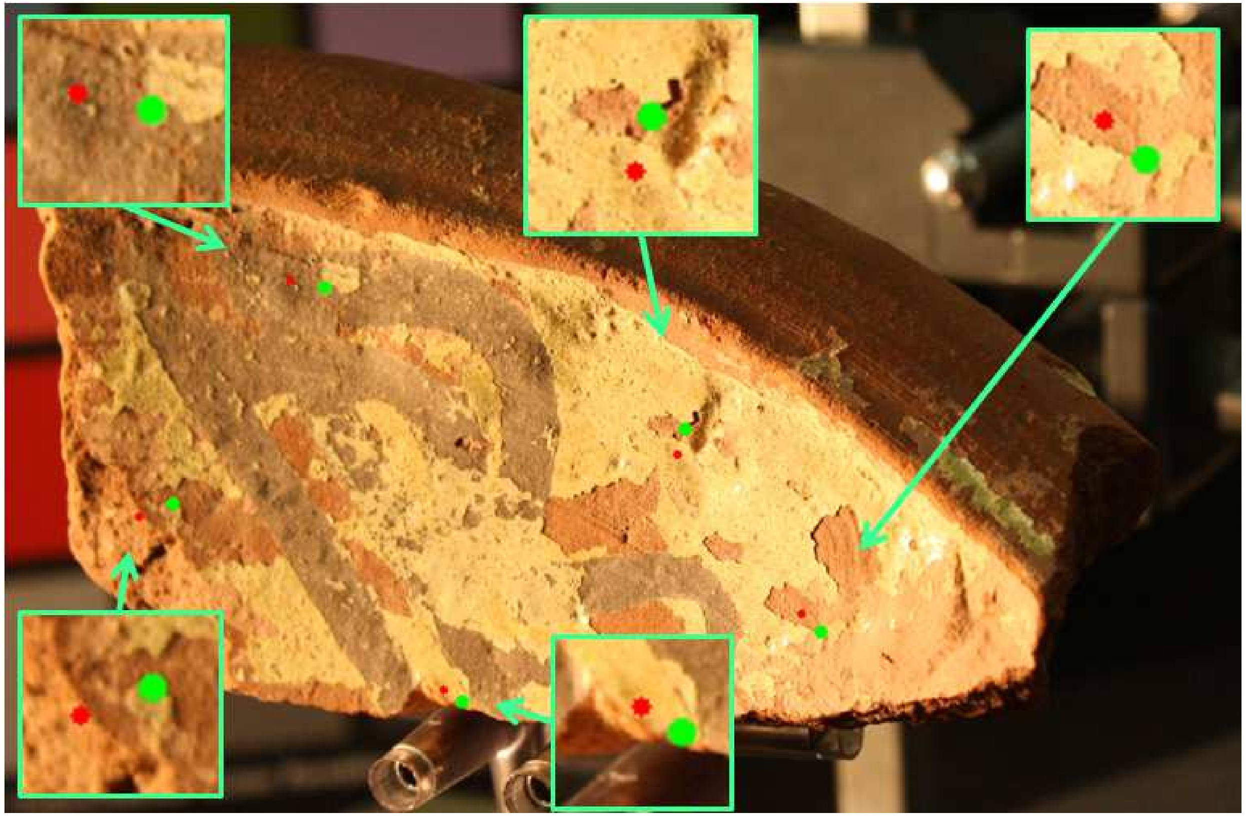
Useful tips:

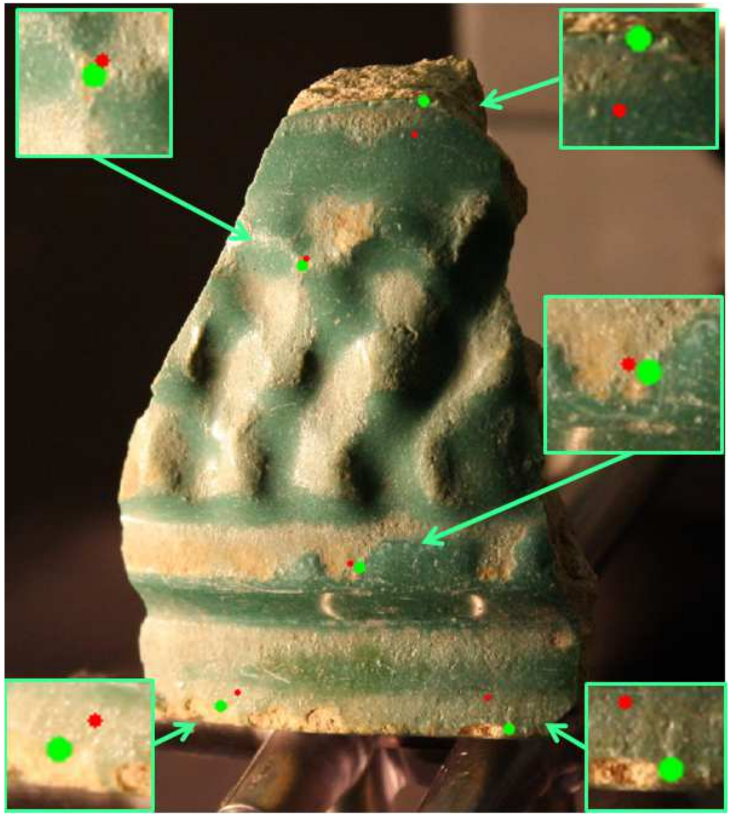
- Imperfections may be caused by the software's alignment, that implements a fast algorithm based on key-points to be able to stitch together scanned slices as you do the scanning. This may introduce some **approximations** at some steps, that translate to the shown imperfections.
- A highly skilled technician may be able to avoid these by a thorough planning and careful scanning operation.
- In most applications these imperfections will only cause issues, if the data is compared at a given point to a **reference**.

Evaluating the results

Registration result evaluated on custom landmark points (green dots show the position of the selected landmark point, red dots show where the 3D equivalent point is projected using the camera pose).

Back-projection **error** measured at **33px** and **28px** respectively.





The presented ceramic fragments belong to the MOM (i.e.. © Maison de l'Orient et de la Méditérannée).

Evaluating the results

Other data set

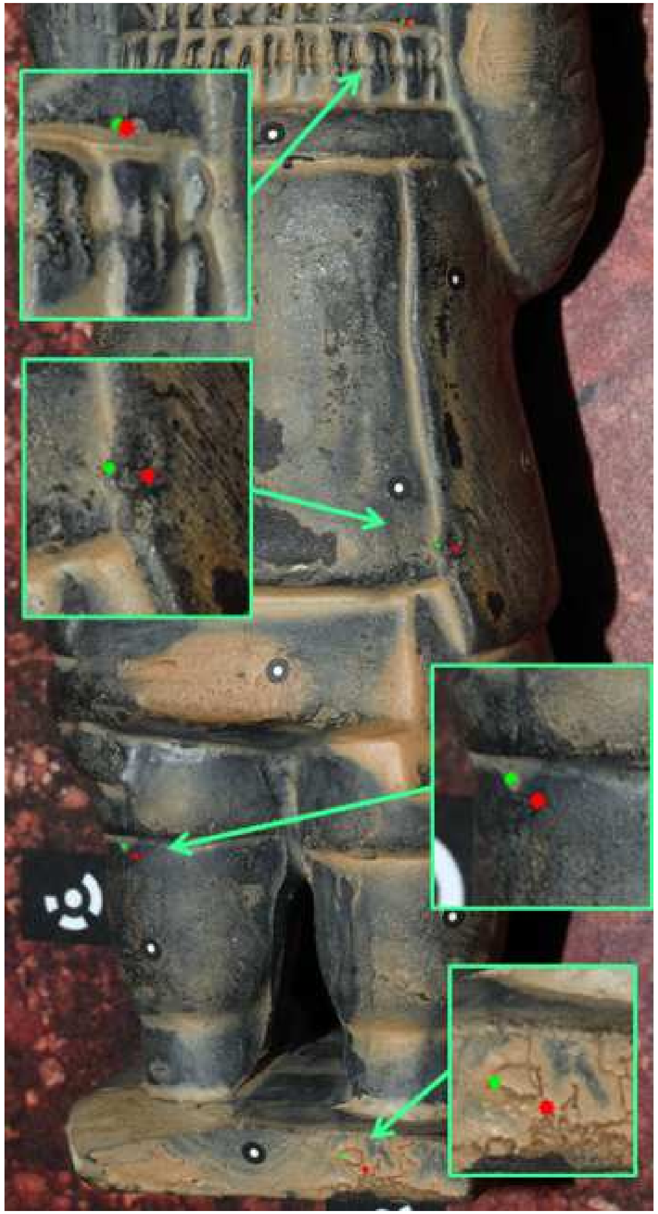
- To possibly avoid these kind of imperfections, we also tested our pipeline on the data provided by i3mainz. The small Chinese warrior figurine that we've used has been captured in strict laboratory conditions for the best precision.
- In this case we didn't observe such imperfections of the 3D data.

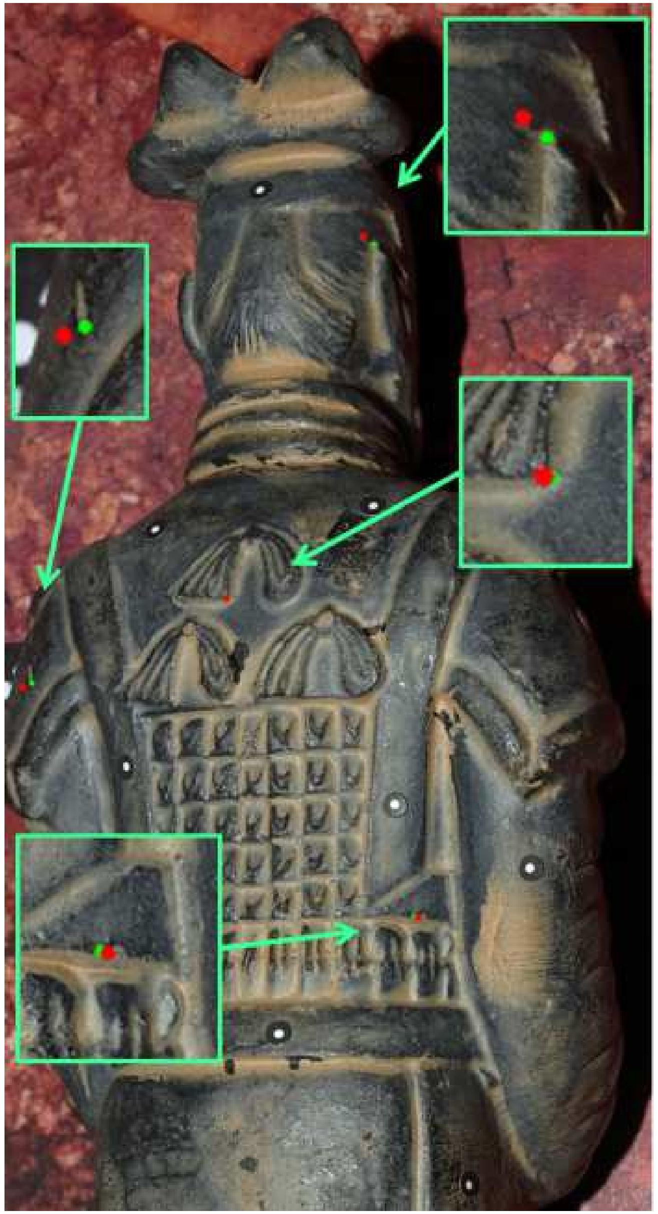
This object belongs to i3mainz (Institute for Spatial Information and Surveying Technology).

Evaluating the results

Only the **pose estimation** result evaluated on custom landmark points (green dots show the position of the selected landmark point, red dots show where the corresponding 3D point is projected using the camera pose).

Back-projection error measured at around **20-30px**, that translates to approx. **1mm** on a 18cm tall object.



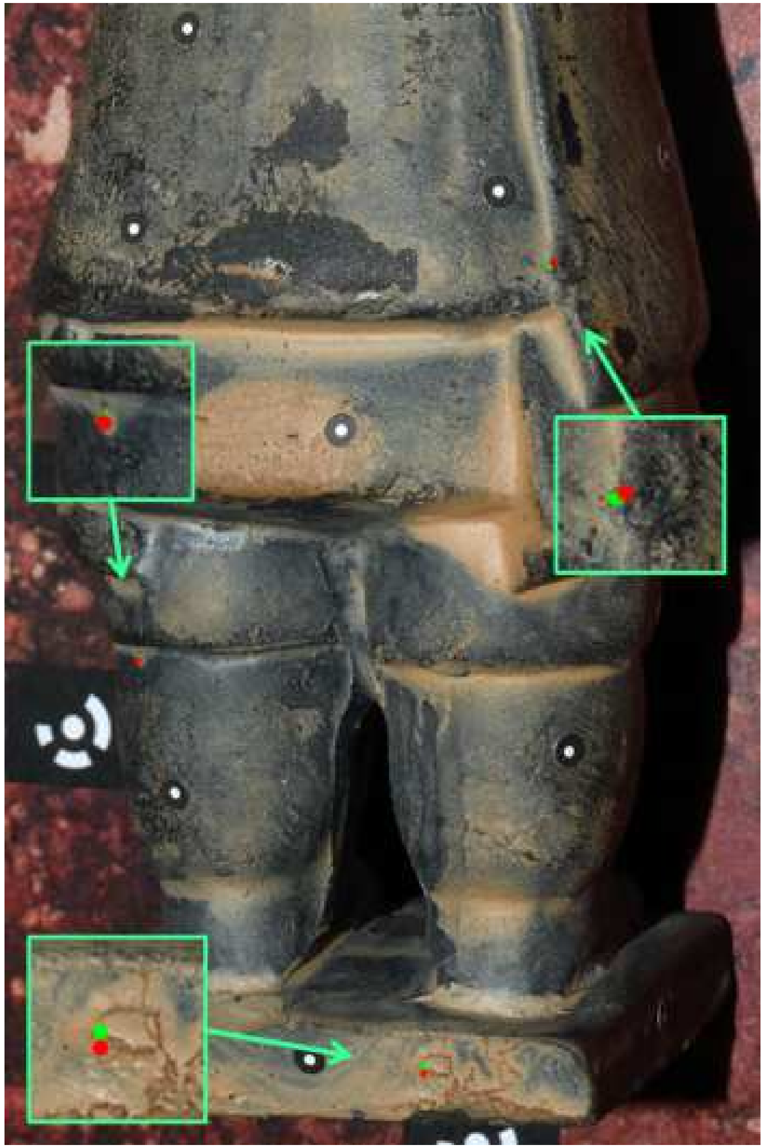


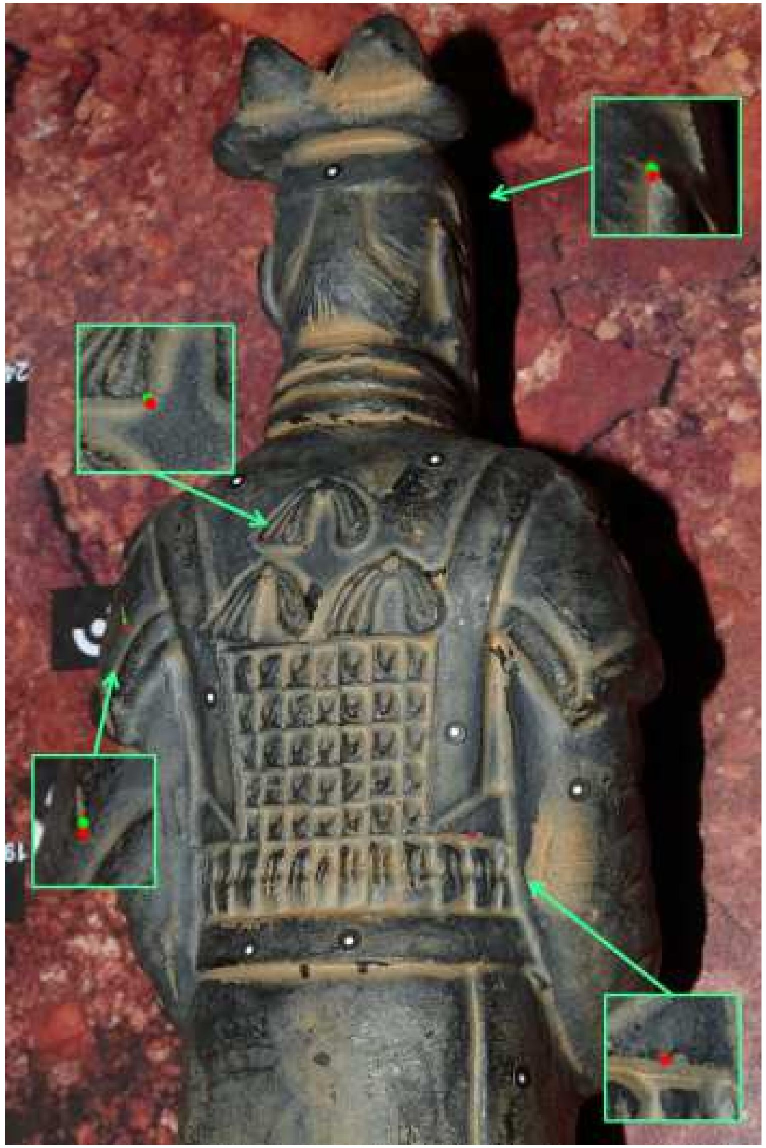
Evaluating the results

COSCH
COLOR & SPACE IN CULTURAL HERITAGE

ICP refinement result evaluated on custom landmark points.

Back-projection error measured at average of **8px**, that translates to approx. **0.2mm** on a 18cm tall object.





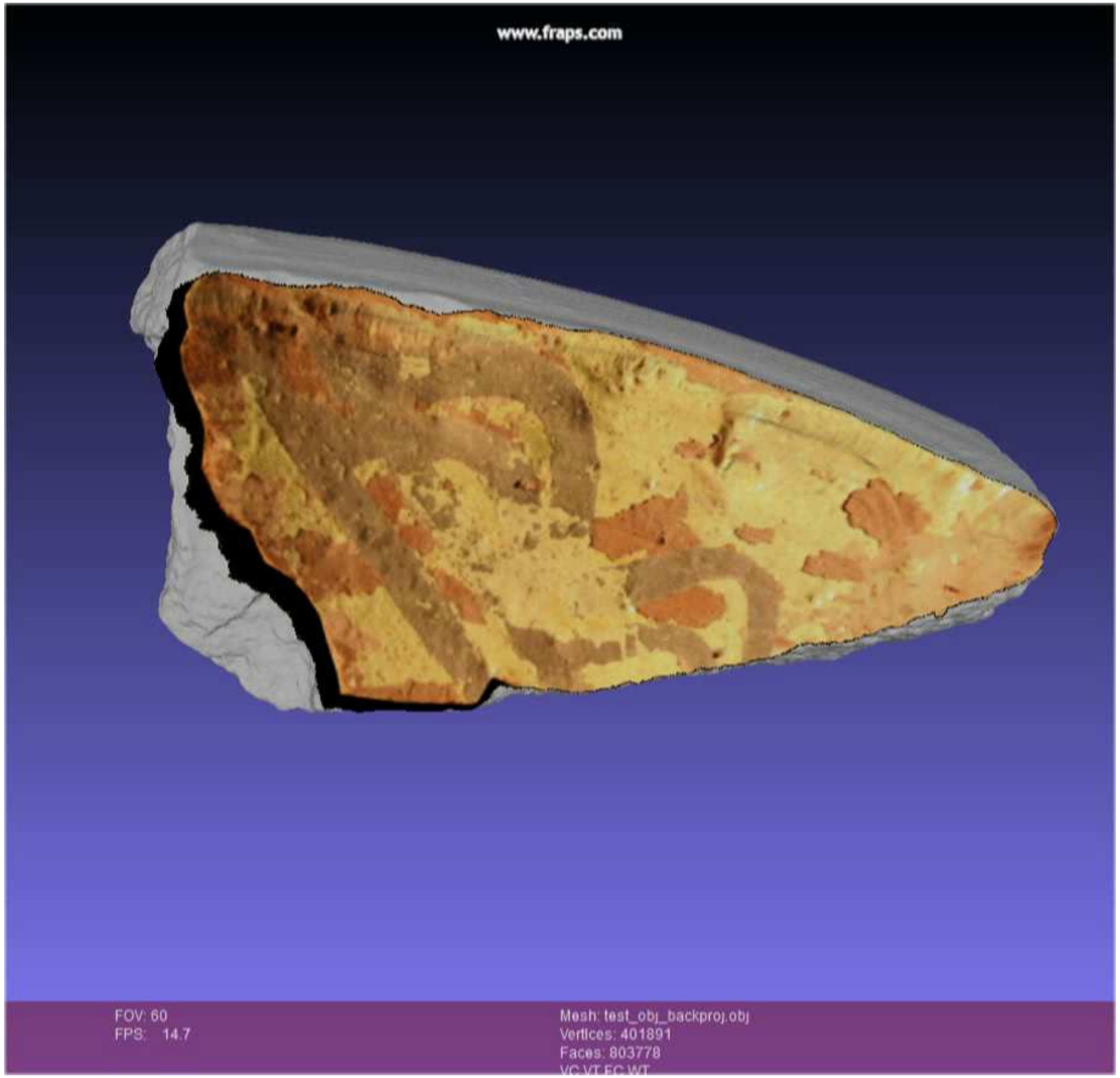
Data fusion

The final step of the workflow, given the estimated camera parameters, is to fuse the 3D raw data and the color information coming from one or multiple RGB images.

For a more pleasing visual result we used **texture mapping**, so the lower resolution of a point cloud is not affecting the visual details that much, as in the case of a vertex-colorized model.

The fusion results shown on ceramics are using only **one image each**.

Data fusion results



Data fusion

To use **multiple images**, we have to decide for each vertex point which image should it receive color from. In our workflow we solved this by calculating the surface normal in the given point, and choosing the camera that has its optical axis pointed most perpendicular onto the surface. More sophisticated decisions can be implemented here as well.

As a result we get a good quality colorized 3D pointcloud. Texture mapping in this case is not possible without "baking" a new texture image, which we didn't do. Some issues still can be seen on our results, for example:

- Some regions are dark, the sides of the warrior are shaded, since the images were not taken using a perfect homogeneous lighting setup.
- The input images were not color calibrated, differences are visible over regions that got colorized from different camera images.
- The red zones on the edges show the small error of the projection parameters, that usually accentuates towards the edges.

Data fusion results

Left: result from a single image (textured). Right: result from 4 views (pointcloud).

Conclusion

COLOR & SPACE IN CULTURAL HERITAGE

Final tips:

There is no best method one can recommend for documenting CH.

- If the detail level that a handheld small scanner can produce is sufficient, we don’t need color calibrated results, and the cost and availability on site is also an important question, than these type of devices are very useful.
- If we need higher color resolution, or even color calibration, then the use of a separate professional camera is advisable, that can be color calibrated with standard algorithms, but in this case, the 3D model should miss geometric imperfections to be able to fuse the data correctly. Of course this involves higher costs, more expensive scanning equipment and usually conditions that are hardly met on the field.

Conclusion

COLOR & SPACE IN CULTURAL HERITAGE

The workflow that we've presented has been applied on real test cases, showing promising results.

- This approach in general may have an advantage over classical solutions, in case of textureless objects, that have no characteristic paint on them, but have some usable geometric properties (some holes, dents, edges, etc.), or in case of repetitive texture as well.
- It may not be possible to use in case of objects with a rough surface, where no small region can be selected without self-occlusion of the data, or in case of objects with a continuous smooth surface, without any slight edges, cracks, holes or imperfections.

We are open to test our method on any other objects and also to make all the 3D data available for those interested, but this may not be possible with all the objects. Please contact the authors for more details.

Conclusion

- In our work we have addressed the *PT3* and *PT5* COSCH *Primary Tasks*, as we have investigated a special processing chain, and showed its potential impact on quality and information content; further more in our conference paper we have presented the practical framework for multisensory data acquisition and processing with its implementation and evaluation.
- We have addressed the subtasks *st3.1* and *st3.2* of *WG 3*.