Type of the Paper (Article)

The example of using action cameras in the invention of perforated objects - a case study in the Royal Castle

Jakub Kuszyk [1], **Natalia Styś** [1]

[1] Warsaw University of Technology; (kuba.kuszyk@selftest.pl, natalia.stys99@gmail.com)

Corresponding Author: kuba.kuszyk@selftest.pl, Tel.: +48-501-276-582

Received date: 09/06/2024; Accepted date: 27/06/2024; Published date: 28/06/2024

Abstract: The article explores the possibility of low-cost action camera usage in the inventory of hard-to-reach objects with complex structures. The main focus of the study is to compare data and results acquired from action cameras, full-frame cameras and terrestrial laser scanning. The object of the study was timber-earth embarkment exhibited in the Royal Castle in Warsaw. It is characterised by a perforated surface with high reflectivity, cracks, and hard-to-reach spaces. The study employs the Xiaomi Mi 4K Action Camera, Canon EOS 5D Mark II and Leica RTC 360 for comparison to TLS. The measurement consisted of a few stages: setting control points and light sources and gathering two sets of photos: general images acquired by the full-frame camera and precise images acquired by the full-frame camera and action camera. The elaboration of the data was mainly done in Agisoft Metashape, with some additional work in GIMP image editor and CloudCompare used for analysis. Results show that, despite the high level of usage in the case of texture, images from action cameras are very demanding if the final product is based on data from them only. Illumination and stabilisation of the camera are crucial in this process. In conclusion, the article emphasises the high value of data acquired from the action camera and highlights essential conditions that must be met for the data to be usable.

Keywords: action cameras, close-range photogrammetry, MVS, SfM, full-frame cameras

1. Introduction

Close-range photogrammetry and Terrestrial Laser Scanning (TLS) are well-known methods of documenting the condition of cultural heritage and architecture (Bruno et al., 2022, Warchoł, 2013). Results of such inventories are applicable by constructors, architects, and conservators (Stylianidis, 2020, Galantucci et al., 2019). In light of the recent development in low-cost photography and photo acquisition tools, the growing potential of close-range photogrammetry can be seen (Markiewicz et al. 2018). It is used to invent relics and monuments, providing data for multiple uses, such as orthoimages or mesh models. The main benefits of this type of invention are complexity, mobility and the fact that this data does not deteriorate thanks to its digital form (Chatzistamatis et al., 2018, Cabrelles, 2009). Structure from Motion (SfM), a method of automated 3D reconstruction, makes the camera alignment process much more manageable. However, the final effect of this method depends on various factors, such as image quality and quantity of characteristic points (Chandler et al., 2016). It provides interior orientation elements and exterior orientation elements. Another method used in generating 3D models from images is Multi-View Stereo (MVS). It is a natural enhancement of the stereoscopic approach. Still, instead of processing two images, it generates a 3D model based on multiple images from different positions around the developed object (Furukawa et al., 2015). Thanks to those factors, photogrammetry is relatively easy and cheap, making it widely used. While in perfect conditions, it can provide perfect data to analyse the condition of monuments in a digital form, there is a low probability that conditions will meet the expectations every time. While in the case of TLS, lightning around the object is irrelevant and does not change the final product, it is the most crucial aspect of the measurement process when the data is gathered using the camera. Of course, thanks to this property, it is possible to get an object's geometry and colour instead of just geometry. Although both methods are helpful in different situations, there is difficulty in gathering high-quality data from hard-to-reach spaces, especially while measuring objects with complex structures. One of the main issues is the lack of possibility to fully visualise the entire object using one data source (Markiewicz et al., 2014). In the case of close-range photogrammetry, low-cost action cameras are possible. Thanks to its small size and the possibility of fixing it on a rod, cameras could be crucial in inventing hard-to-reach objects.

The paper compares data and products obtained by the full-frame camera (Canon EOS 5D Mark II) and Xiaomi Mi 4K Action Camera. It compares the model's geometry generated from photos and laser scanning. The sets of images were used to generate dense point cloud and mesh models in Agisoft Metashape software. Photos from each device were used to generate different sections of the object but with the same properties. The article presents an example of using an action camera with two different approaches: generating images from video frames and using a self-timer with a second interval.

2. Materials and Methods

The study presents the data acquisition methodology; the most important aspects of this process are the setting of light sources, photo overlap, and camera stability, except that it describes the data processing with the use of Python, GIMP and Agisoft Metashape software. In the end, the CloudCompare programme summarised and analysed the final product. The main focus of the study is to present the way of capturing photos of an object with a complex texture that is hard to access.

The data acquisition process started with setting up the lamps and control points. For close-range photogrammetry, this part is crucial because the final colour of the 3D model will reflect the colour of an object in the photos. Control points are essential in the image orientation process. The algorithm aligns photos more efficiently based on control points than only automatically generated common points. Control points were positioned around the object at different altitudes and distances from the object to avoid errors in image processing.

Canon EOS 5D Mark II was used to get general photos of the relic, with the use of a 24 mm lens, shown in Figure 1, as well as precise images of its structure, with the use of a 35 mm lens, shown in Figure 2 where it was possible.

Figure 1. General photo of the relic

Figure 2. Precise photo of the structure

Tripod enabled better camera stabilisation, thanks to which exposure time could be extended. The camera was used at three different heights to get better geometry for the photo network. The final count of the photos was 531, consisting of 93 general photos and 438 precise photos of the structure. The Xiaomi Mi 4K Action Camera was used mainly in the central section of the relic due to insufficient space for the tripod. The experiment consisted of two approaches: the use of video and self-timer. In this case, the camera was attached to a rod, without stabilisation, significantly impacting the final result. The result of this measurement was 31 minutes of video and 103 photos. Photos from the video were exported using the simple function from the OpenCV library in Python. The algorithm extracted every thirtieth frame, translating into one photo per second. Unfortunately, frame extraction has some limitations. Due to it being automated, the operator cannot control the outcome. Some frames are overexposed or blurry and have to be deleted. Therefore, it is possible to delete crucial photos without which alignment would be unsuccessful. After the filtering process, the final result was 1025 photos.

Image processing consists of several stages:

- Colour, contrast and brightness correction of the photos – GIMP
- Align general images with the use of control and self-calibration of the cameras – Agisoft Metashape
- Align precise photos with an automated point-based method – Agisoft Metashape
- Mesh model generation from depth maps of the general images and detailed photos
- Texture generation from a smaller sample of the precise photos (200 photos)
- Import control points' coordinates acquired in CloudCompare software for analysis purposes

3. Results

The processing stage was challenging due to problems with correct images alignment. It was caused mainly by various reasons, such as the number of images and their geometry – they were made very close to the object, so in some cases, the algorithm could not find similarities in images next to each other. Because of that fact, a small sample of general images was used to align alongside precise images.

3.1. Model generated from general images and precise images

The main issue of the model generation process was the accuracy of image alignment. There was no issue regarding precise images acquired by the Canon camera. Images were sharp and not overexposed compared to photos acquired by action camera. It resulted in less confident camera placement for the images inside the model, resulting in a lower quality of a resulting 3D model, as seen in Figures 3 and 4. The red marking indicates low confidence in surface placement, and the blue colour indicates high confidence:

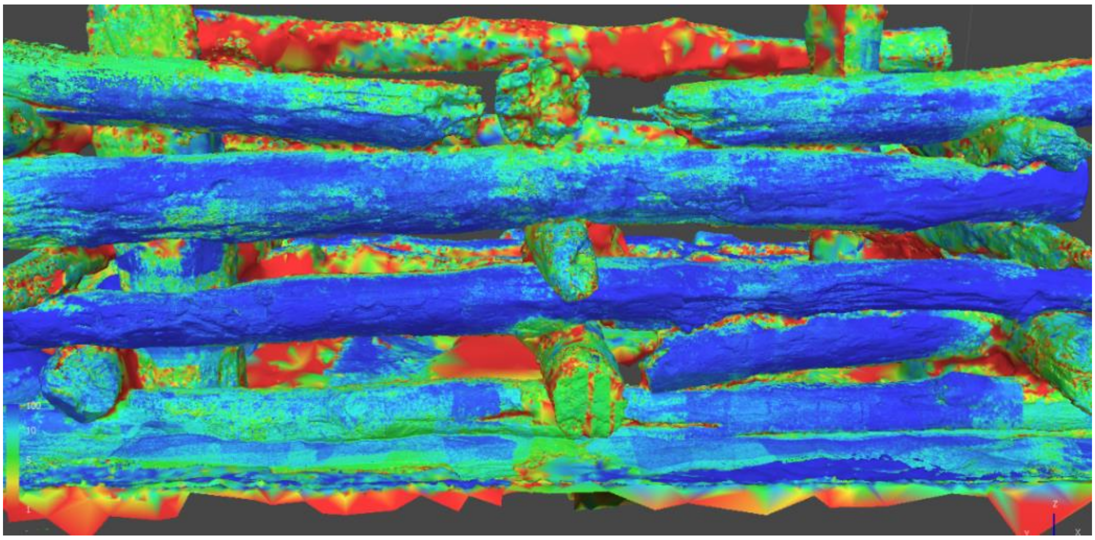

Figure 3. Outside of the 3D model displayed in "confidence" mode

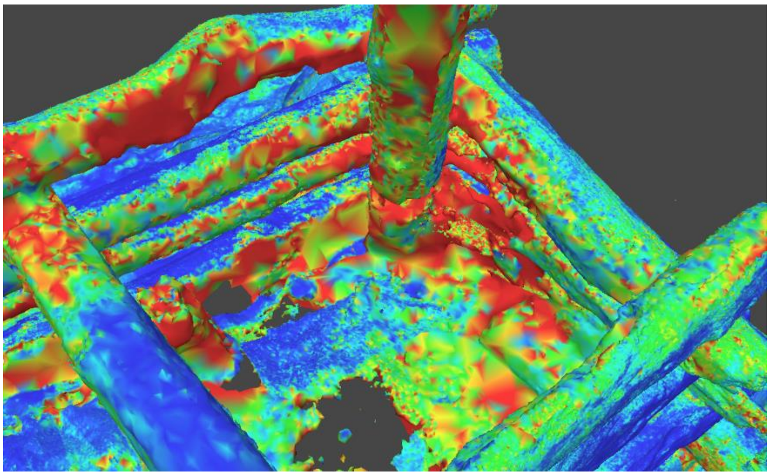

Figure 4. Inside of the 3D model displayed in "confidence" mode

It is also visible in wireframe mode, where it can be seen that the outside of the model (Figure 5.) is very detailed and dense, while the inside section is generalised (Figure 6.).

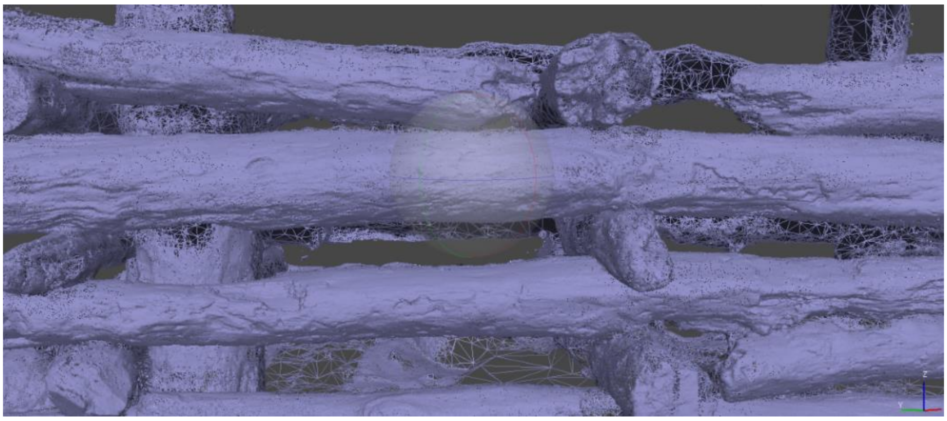

Figure 5. The outside view of the 3D model displayed in "wireframe" mode

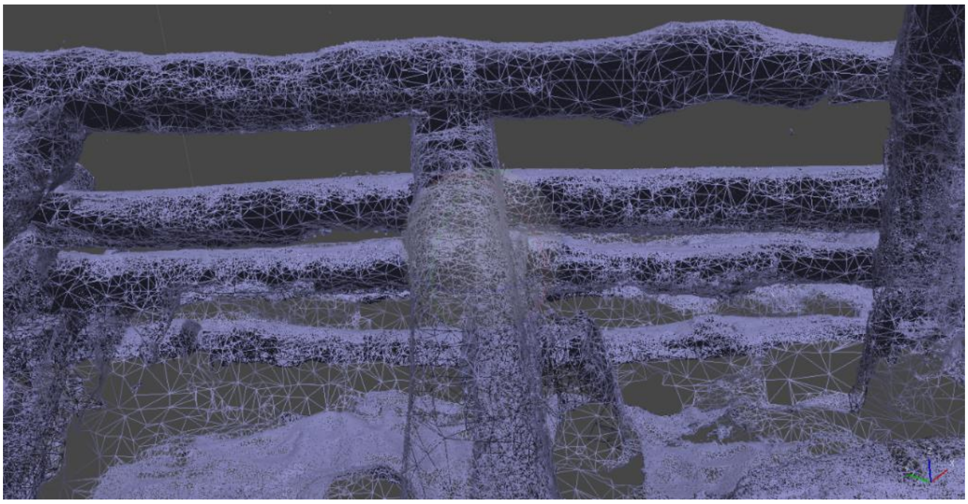

Figure 6. The inside view of the 3D model displayed in "wireframe" mode

It was mainly caused by insufficient illumination and stabilisation of the camera. It also shows that image acquisition from videos is not the appropriate method because self-timer photos were taken on the top section of the object, which is denser than the inside. Due to those mistakes, the geometry of the 3D model was unsatisfactory, but it shows the significant impact of camera stabilisation and lens quality in this type of elaboration. In the final product, 3D geometry was generated using terrestrial laser scanning.

3.2. Texture generation from original images and edited images

Problems with geometry do not change the fact that those images were crucial in texture generation. To perfectly reflect reality, some images were edited in GIMP software due to excessive illumination. As seen in Figures 7 and 8, the edited photos allowed texture generation, representing a better tone on the object than the original photos.

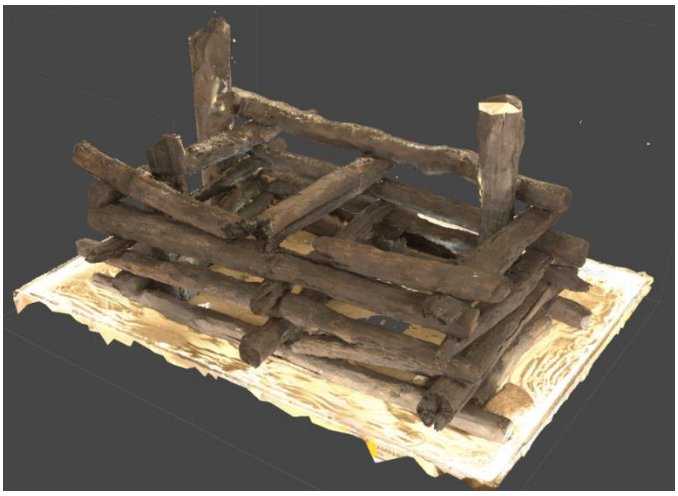

Figure 7. 3D model generated from images with texture from original photos

Figure 8. 3D model generated from TLS with texture from edited photos

Even though the texture looks better, a distortion is visible on the sand and platform of the object. For that reason, spherical images from laser scanning were used to replace these sections of the texture. In this case, parts of the photos had to be excluded from the texture generation achieved by masking those sections on the images (Figure 9.).

Figure 9. Placement of masks on edited photos

The final product of this process represents the actual state of an object in the most accurate way of all approaches (Figure 10.).

Figure 9. The final product of the project

3.3. Comparison of 3D models generated from images and TLS

CloudCompare software was used as the analysis tool in this section. The comparison was mainly based on the object's cross sections, the distance between the planes of both 3D models and their complexity in Figure 10. It is seen that models are nearly identical on the outside, and further inside, the more significant gaps can be observed.

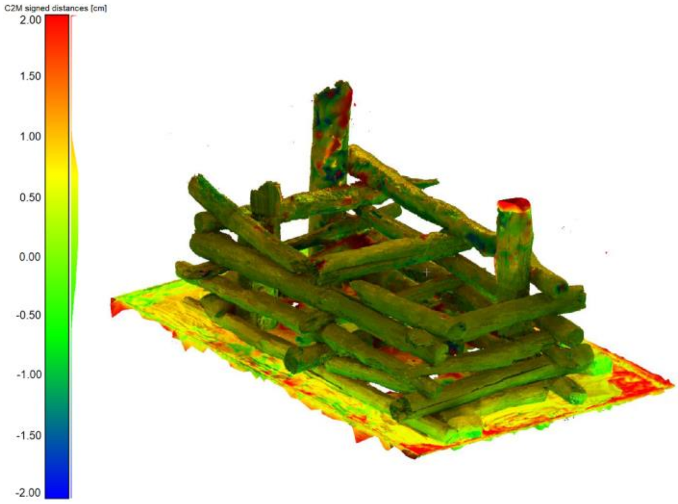

Figure 10. Distance between models

The complexity disproportions are visible in the lower parts of the monument, where illumination was insufficient. Although the geometric part of the model is imperfect, the tone extracted from the images after the edition is way better, which shows that editing images makes a huge difference in that context, in contrast to the complexity of the model. Figure 11 and Figure 12 shows that perfectly:

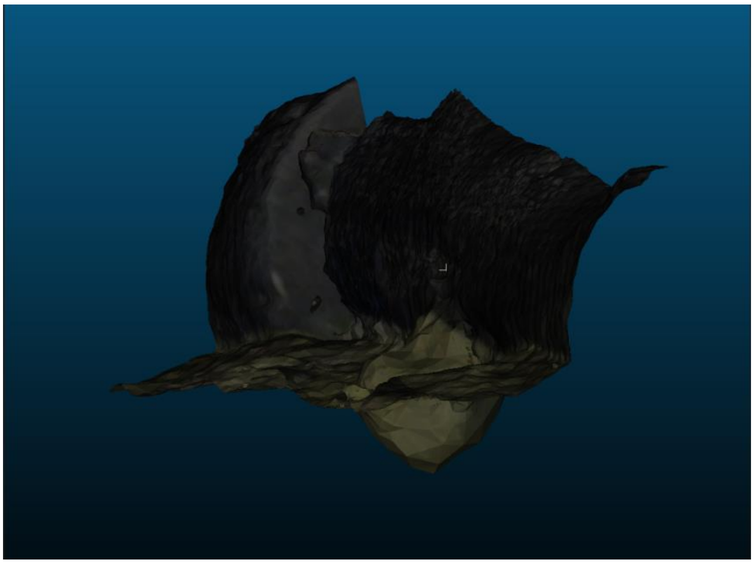

Figure 11. The 3D model with geometry acquired from TLS and texture from edited photos

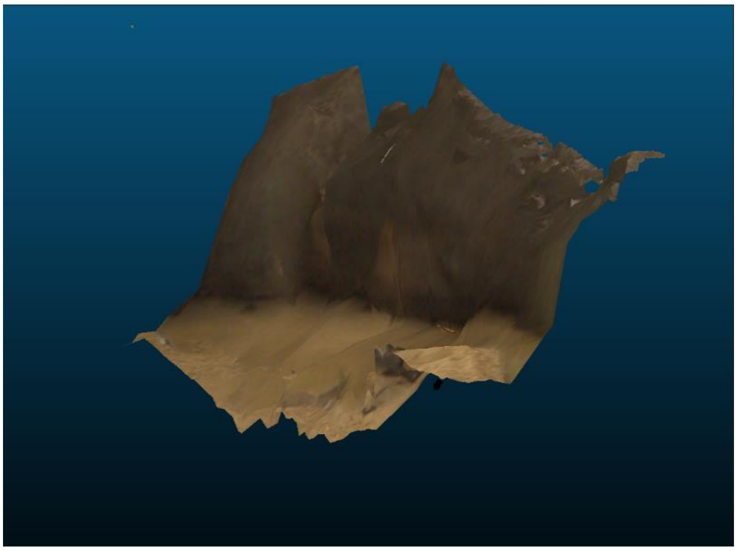

Figure 12. The 3D model with geometry and texture acquired from unedited images

It is seen that although action camera usage is cheaper, more flexible and more reliable in the texturing process, it is still more demanding than the usage of TLS. With proper illumination and stabilisation, the effects of camera usage are as good as TLS when talking about geometry acquisition.

4. Discussion and conclusion

The use of cameras and TLS in the inventory of monumental objects is at its highest level. Both methods have their pros and cons. At this stage, it can be said that the use of cameras is better than that of TLS. However, there are disparities in the preparation process for the measurement site. As the technology improves, close-range photogrammetry with action cameras can be the primary data source for such projects. Better stabilisation devices like gimbals for action cameras and more compact lamps are available on the market so that those hard-to-reach spaces can be appropriately measured.

The study compares the inventory results acquired from full-frame cameras, low-cost action cameras, and TLS. In this experiment, it is clearly seen that all of those methods have their advantages over each other, depending on the developed object's location, complexity and illumination. While, in this case, the action camera gave the worst geometry results, it must be pointed out that without this data, gathering information about the object's colour would not be possible. The full-frame camera gave data that was as useful as this from TLS, but the usage of the camera is way cheaper than TLS. Of course, TLS provided the best results on the outside and inside of the object but was not that useful in texture building.

After the experiment, it seems inevitable that the action camera is a valuable tool in close-range photogrammetry, although a couple of crucial conditions must be met. Those are stabilisation of the camera, which ensures that photos will be sharp and well interpreted by the SfM algorithm. Another condition is lightning, which can be as tricky as using a full-frame camera in hard-to-reach places. One might say that there are the same conditions with the use of full-frame cameras. However, with a bit of space for all those devices, it can be harder to implement them.

Authors Contributions: J.K. and N.S. organised the conceptualisation of the idea and the methodology employed in this paper. Following that, J.K. carried out the experimental design. J.K. and N.S. worked on the data acquisition at the Royal Castle in Warsaw. J.K. carried out the original writing and draft preparation. J.K. and N.S. undertook the data analysis. All authors have read and agreed to the published version of the manuscript.

Funding: "This research received no external funding".

Conflicts of Interest: "The authors declare no conflict of interest".

References

1. Bruno N., Mikolajewska S., Roncella R., Zerbi A., 2022, "Integrated processing of photogrammetric and laser scanning data for frescoes restoration", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 46, (pp. 105-112)

2. Warchoł A., 2013, "Aspects of photogrammetric data integration for generation 3D models of the selected objects located in the urban space", Archiwum Fotogrametrii, Kartografii i Teledetekcji, 24, (pp. 199-209)

3. Stylianidis E., 2020, "Photogrammetric Survey for the Recording and Documentation of Historic Buildings.", Cham, Switzerland: Springer.

4. Galantucci R. A., Fatiguso F., 2019, "Advanced damage detection techniques in historical buildings using digital photogrammetry and 3D surface analysis.", Journal of Cultural Heritage, 36, (pp. 51-62)

5. J. Markiewicz, S. Łapiński , M. Pilarska , R. Bieńkowski , A. Kaliszewska., 2018 "Investigation into the use of action cameras in the documentation of architectural details – the case study of a baroque chamber". The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2, ISPRS TC II Mid-term Symposium "Towards Photogrammetry 2020", 4–7 June 2018, Riva del Garda, Italy

6. Chatzistamatis S., Kalaitzis P., Chaidas K., Chatzitheodorou C., Papadopoulou E. E., Tataris G., Soulakellis N., 2018, Fusion of TLS and UAV photogrammetry data for post-earthquake 3D modeling of a cultural heritage Church. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42, (pp. 143-150)

7. Cabrelles M., Galcerá S., Navarro S., Lerma J. L., Akasheh T., Haddad N., 2009, "Integration of 3D laser scanning, photogrammetry and thermography to record architectural monuments", In Proceedings of the 22nd CIPA Symposium, Kyoto, Japan (pp. 11-15)

8. Chandler, J., & Buckley, S., 2016, "Structure from Motion (SFM) Photogrammetry vs Terrestrial Laser Scanning.", Loughborough University.

9. Furukawa Y., Hernández S., "Multi-View Stereo: A Tutorial", 2015 https://carlos-hernandez.org/papers/fnt_mvs_2015.pdf

10. J. Markiewicz, D. Zawieska, 2014, Terrestrial scanning of digital images in inventory of monumental objects? – Case study, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-5, ISPRS Technical Commission V Symposium, 23 – 25 June 2014, Riva del Garda, Italy