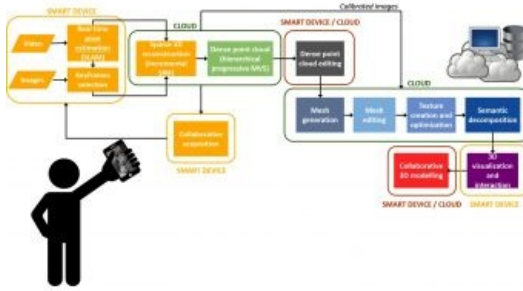


CHEAP AND RAPID 3D MODELLING

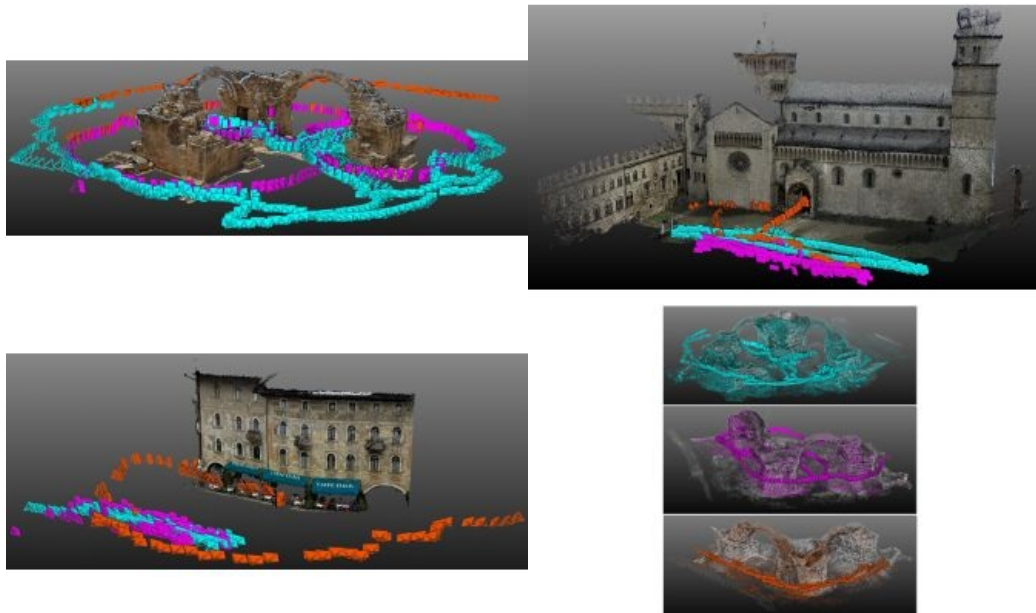
Point Clouds from Smartphones



	# videos	Smartphone	Resolution (pixel)	# keyframes	Smartphone orientation
Sancti Petrus	3	Huawei P9	1500x1080	1201546235	Landscape & Portrait
	4	OnePlus One	1500x1080	1201546235	Portrait
	3	Samsung S6	3600x1050	8450756	Landscape & Portrait
Plaza Duomo	1	LG Nexus 5X	1500x1080	84	Landscape & Portrait
	1	Samsung Galaxy Alpha	640x480	91	Landscape & Portrait
	2	SonyPS	1500x1080	36	Landscape
Cattedrale	1	LG Nexus 5X	1500x1080	175	Landscape & Portrait
	1	Samsung Galaxy Alpha	640x480	218	Landscape & Portrait
	1	SonyPS	1500x1080	107	Landscape & Portrait

Smartphones are omnipresent, and many people can no longer do without them. Smartphone cameras capture images suited for generating point clouds and 3D models. Apps running on smartphones and software running on a remote server enable easy 3D modelling from multiple images. The challenge is to train and guide laymen through a proper image capture strategy using their smartphones. The authors of this article investigated the potential use of smartphones for cheap and rapid generation of point clouds and 3D models exploiting a collaborative approach.

Today's smartphone images are higher in quality than they were a few years ago. Overlapping images taken with smartphones combined with software running on desktop computers allow the creation of dense and accurate point clouds without expensive tools. Images and videos captured by smartphone cameras can be processed by pipelines running on a remote server. This puts easy 3D modelling from multiple smartphone images within reach for applications such as 3D documentation and modelling in cultural heritage and the creative industry. Disaster management may also benefit from this approach for



documenting, monitoring and inspecting hazardous environments.

‘Replicate’ Project

The popularity of 3D modelling from images has soared in many applications over the last decade, mainly because of the availability of high-quality, affordable photographic tools and easy-to-use software enabling automatic pipelines. The creation of 3D content has become common for professionals and the general public alike because of the widespread availability of digital media, 3D printing, and extended reality. For example, apps that implement structure from motion (SfM), dense image matching (DIM) and meshing for generating 3D models (e.g. ItSeez3D, TRNIO) running on smartphones can be downloaded from app stores. Alternative solutions split the digitisation process between the smartphone and a high-performance computing (HPC) remote server. The images or videos are captured by smartphones while the heavy computation required by SfM, DIM and meshing is done on the remote server (Figure 1). Next, the 3D model is accessible through the smartphone. This idea of a hybrid 3D modelling pipeline is further scrutinised in the EU H2020 REPLICATE project. The research focuses on guiding multiple smartphone users, called ‘replicators’, through the image capture stage, the editing of the 3D model and its conversion into 3D media content.



Figure 1: REPLICATE pipeline jointly done on smartphones and the HPC remote server.

From Video Stream to 3D Model

Smartphone users can capture images individually or in groups, either at the same time or at different moments. Authorised collectors can use an app – such as Replicorder, developed by the authors – to extract and select keyframes and then upload them together with the GNSS position and acceleration data to the remote server. Keyframes are selected based on frame quality, such as sharpness, and overlap with previous frames. The overlap is estimated from the number of newly extracted and matched feature points with respect to the

previously selected frames. To avoid gaps, the time span between two keyframes, regardless of their quality, should not exceed a certain time interval. On the server, first an incremental SfM algorithm is applied on the overlapping frames to process images captured by multiple individuals at different times. The sparse point cloud, processed at the server, can be displayed on the smartphone during frame acquisition. Newly uploaded keyframes are matched with keyframes already used for reconstruction. When the frame acquisition is completed, a global bundle adjustment (BA) is executed to refine the camera parameters and sparse point cloud. Next, DIM is carried out to build a dense point cloud from the sparse cloud imposing photoconsistency constraints across pairs of frames. All users can simultaneously visualise the progress of the 3D modelling in a preview window in the smartphone app. Users can also interact with the 3D modelling session via a web page.



Table 1: Main characteristics of the images captured at the three test sites.

Tests

Six off-the-shelf Android smartphones captured single or multiple video streams (Table 1) of three cultural heritage sites. Each site was captured by three smartphones by different users to simulate a collaborative 3D documentation approach. The first site, the ‘Saranta Kolones’ monument at Paphos, Cyprus, measuring about 16m x 16m x 5m, is on the UNESCO World Heritage list. The dimensions and complexity are major challenges. The videos were taken throughout the day resulting in varying lighting conditions (Figure 2). The second site was the north-facing façade of the cathedral on ‘Piazza Duomo’ in Trento, Italy, which has a length of 100m and a height of 30m (Figure 3). The third site was the south-facing facade of a painted building measuring 30mx15m and housing ‘Caffe Italia’ in Trento. These sites were simultaneously captured by three collectors (Figures 2 to 4). Challenges of these sites were the flat geometry of the facades and the presence of people, cars, trucks and other moving objects.



Figure 2: RGB dense point cloud of Saranta Kolones; coloured pyramids indicate position and orientation of the keyframes.

Results

The metric potentialities of the pipeline were examined by comparing the dense point clouds with reference data acquired with professional digital single lens reflex (DSLR) cameras equipped with fixed focal length lenses and processed with state-of-the-art commercial software. For the ‘Saranta Kolones’ dataset, 176 images were captured; it was possible to achieve a final root mean square (RMS) error better than 5mm computed on 20 checkpoints. For the ‘Piazza Duomo’ and ‘Caffe Italia’, 359 images were captured; comparison with 18 checkpoints resulted in an RMS error better than 10mm. Before comparison, noisy elements were removed from the point clouds generated with REPLICATE and then aligned to the reference data in a local reference system, through the iterative closest point (ICP) algorithm with scale factor registration method (CloudCompare). The comparison showed that large differences between the point clouds generated by three smartphone cameras and reference data, up to 50cm, are visible at object edges. Here the poorest image quality of the smartphone cameras, along with the pixelisation effect, is most evident, which negatively affects the result of dense image matching. However, the global geometry of the structures shows smaller deviations, suitable for 3D documentation and modelling for the creative industry.

Concluding Remarks

Researchers and developers are invited to access the datasets of the three test sites, including reference data, through <http://tev.fbk.eu/collaborative3D>.

Acknowledgments

This research has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement number 687757.



Figure 3: RGB dense point cloud of Piazza Duomo; coloured pyramids indicate position and orientation of the keyframes.



Figure 4: RGB dense point cloud of Caffe Italia; coloured pyramids indicate position and orientation of the keyframes.



Figure 5: Three sparse point clouds of Saranta Kolones (coloured) generated from the three smartphones; grey indicates the entire point cloud.

Further Reading

Kolev, K., Tanskanen, P., Speciale, P. and Pollefeys, M. (2014) Turning mobile phones into 3D scanners. Proc. IEEE CVPR, pp. 3,946-3,953.

Locher, A., Perdoch, M., Riemenschneider, H. and Van Gool, L. (2016) Mobile phone and cloud—A dream team for 3D reconstruction. Proc. IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1-8.

Muratov, O., Slynko, Y., Chernov, V., Lyubimtseva, M., Shamsuarov, A. and Bucha, V. (2016) 3DCapture: 3D Reconstruction for a Smartphone. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 75-82).

Nocerino, E., Lago, F., Morabito, D., Remondino, F., Porzi, L., Poiesi, F., Bulò, S.R., Chippendale, P., Locher, A., Havlena, M. and Van Gool, L., (2017) A smartphone-based 3D pipeline for the creative industry – the REPLICATE EU project. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLII-2-W3, pp. 535-541

Nocerino, E., Poiesi, F., Locher, A., Tefera, Y.T., Remondino, F., Chippendale, P. and Van Gool, L. (2017) 3D Reconstruction with a Collaborative Approach Based on Smartphones and a Cloud-Based Server. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLII-2-W8, pp. 187-194

<https://itseez3d.com>

<http://www.trnio.com>

<http://www.cloudcompare.org>

<https://www.gim-international.com/content/article/point-clouds-from-smartphones>
