

## PHOTOGRAMMETRY AS A PARTICIPATORY RECOVERY TOOL AFTER DISASTERS: A GROUNDED FRAMEWORK FOR FUTURE GUIDELINES

C. Pezzica<sup>1,3\*</sup>, A. Piemonte<sup>2</sup>, C. Bleil de Souza<sup>3</sup>, V. Cutini<sup>1</sup>

<sup>1</sup> DESTeC, Department of Energy, Systems Territory and Construction Engineering, University of Pisa, Italy - [valerio.cutini@ing.unipi.it](mailto:valerio.cutini@ing.unipi.it)

<sup>2</sup> DICl, Department of Civil and Industrial Engineering, University of Pisa, Italy - [andrea.piemonte@unipi.it](mailto:andrea.piemonte@unipi.it)

<sup>3</sup> WSA, School of Architecture, Cardiff University, King Edward VII Avenue, Cardiff, UK - ([pezzicac](mailto:pezzicac@cardiff.ac.uk), [bleildesouzac](mailto:bleildesouzac@cardiff.ac.uk))@cardiff.ac.uk

### Commission II, WG II/8

**KEY WORDS:** Collaborative Mapping, Street-level data, Photogrammetry, Disaster Recovery, Guidelines, Crowdsourcing, SfM

### ABSTRACT:

This paper identifies the application domain, context of use, processes and goals of low-cost street-level photogrammetry after urban disasters. The proposal seeks a synergy between top-down and bottom-up initiatives carried out by different actors during the humanitarian response phase in data scarce contexts. By focusing on the self-organisation capacities of local people, this paper suggests using collaborative photogrammetry to empower communities hit by disasters and foster their active participation in recovery and reconstruction planning. It shows that this task may prove technically challenging depending on the specifics of the collected imagery and develops a grounded framework to produce user-centred image acquisition guidelines and fit-for-purpose photogrammetric reconstruction workflows, useful in future post-disaster scenarios. To this end, it presents an in-depth analysis of a collaborative photographic mapping initiative undergone by a group of citizen-scientists after the 2016 Central Italy earthquake, followed by the explorative processing of some sample datasets. Specifically, the paper firstly presents a visual ethnographic study of the photographic material uploaded by participants from September 2016 to November 2018 in the two Italian municipalities of Arquata del Tronto and Norcia. Secondly, it illustrates from a technical point of view issues concerning the processing of crowdsourced data (e.g. image filtering, selection, quality, semantic content and 3D model scaling) and discusses the viability of using it to enrich the pool of geo-information available to stakeholders and decision-makers. Final considerations are discussed as part of a grounded framework for future guidelines tailored to multiple goals and data processing scenarios.

### 1. INTRODUCTION

The growing exposure of contemporary cities to natural and man-made hazards calls for a better management of urban contingencies (Borsekova & Nijkamp, 2019). As these are becoming an integral part of urban planning practice, instruments such as Geographic Information Systems (GIS) and (Historic) Building Information Modelling (BIM and HBIM), are now a crucial component of the toolbox of planners and disaster managers seeking sustainable urban futures (Billen et al., 2014). They allow managing complex collaborative projects involving several actors and can be used to guide decision-making through scenario-building techniques and simulations at all scales (Breunig et al., 2017). Despite GIS and BIM require access to a rather large amount of geographic data, which could be hardly available after a disaster, novel opportunities for their use in data scarce contexts are emerging as a result of rapid technological advances in the 3D documentation of the built environment. An example are increasingly accessible scanning techniques, low-cost photogrammetric methods and flexible software solutions (Wilson et al., 2009), supported by the continuous development of powerful computer vision algorithms for the semantic mining of 3D models (e.g. deep segmentation). Another opportunity is represented by the ubiquitous availability of low-cost imaging devices.

Past studies have demonstrated that in the aftermath of a disaster, high-resolution satellite images can be insufficient to fully record buildings and urban aggregates or may be simply unavailable due to weather-related factors, besides lacking in accuracy and coverage of vertical plans (Toschi et al., 2017). Considering that many authors suggest compensating for these shortcomings by

using UAV imagery (Erdelj et al., 2017; Petrides et al. 2017; Copernicus EMS, 2016), this paper argues that crowdsourced street-level imagery can be an additional valuable resource to complete missing information useful to urban designers and planners in circumstances where other data were either unavailable, too poor or excessively delayed. The UN Action Guidelines (UNISDR, 2017) highlight the potential benefits of using crowdsourcing to collect geographic information: beyond the rapid gathering of data at a large scale, these include building community resilience by indirectly educating participants about risks in their areas. Nonetheless, enabling a safe participation of the general public in data capture operations after disasters remains an open challenge. Some advantages and limits are analysed by Poblet et al. (2014).

Within this framework, the paper claims that photogrammetry can play a key role in *building back better* (UNDRR, 2015) by effectively bridging the efforts of citizen scientists (members of the public collecting data about the built environment) and professionals from the humanitarian sector. It suggests that, a synergy between top-down responses by relief actors and bottom-up initiatives by local communities can be achieved in the context of a temporary - yet often enduring - urban crisis, by positively focusing on the resourcefulness and self-organisation capacities of local activists and by supporting an open flow of data and knowledge produced by citizen-scientists (UN environment, 2018). To this end, it discusses the technical and methodological challenges of using collaborative, low-cost, street-level photogrammetry as a tool to orient strategic decisions, empower the affected communities and foster their active participation in post-disaster planning. I.e., for the crowdsourced data to be efficiently processed with state-of-the-art tools, it can be useful

to instruct people on how to collect images that match the current technical needs and constraints of photogrammetry. This paper develops a framework to produce guidelines for non-technicians aiming at enhancing the value of collaborative photographic mapping efforts as well as some processing tips so the data can turn useful to the many, by tracing a shared recovery path towards resilience.

## 1.1 Background and related work

The last two decades have been characterised by a revolution in 3D surveying techniques due to the impact of the computer vision community in the field of photogrammetry and the web, which have pushed forward the computational possibilities for 3D data acquisition and processing as well as the ease of the associated methods. Initially, only trained professionals could perform photogrammetric surveys. They were required to capture pictures with a metric camera moving through a certain scene at a given time and subsequently process those using stereo pairs. Thus, the interval between the observations used to have a certain temporal consistency in line with the required overlap (usually 60% in one direction and 20% on the perpendicular axis), homogeneous lighting conditions, and fixed internal camera parameters. While this pipeline allowed a better control on the data acquisition process, it was dependent upon both expensive equipment and professional expertise. Therefore, terrestrial laser scanners (TLS) became a highly popular alternative even if past technologies were at times impractical (e.g. survey of narrow staircases or high buildings). Despite their recent diffusion and improvement, TLS remain rather expensive. Therefore, for many applications such as heritage conservation, geography studies and medical scanning, there is now an interest in producing 3D data using much cheaper and portable technologies such as photogrammetry from smartphones (Barbero-García et al. 2017; Kirchhöfer et al., 2011; Micheletti et al., 2015; Zhang et al., 2016), which became viable when traditional survey requirements were relaxed thanks to the integration in the processing pipeline of analytical approaches and automated procedures developed by the machine vision community at the beginning of the 2000s. The ease with which some algorithms model geometrical distortions associated with image acquisition has allowed the range of usable sensors to increase significantly. The automated tracking of points across multiple images using SIFT and SURF for feature design, and more recently deep learning methods, has supported the handling of large image datasets. Furthermore, an increased flexibility in photogrammetric 3D modelling is achieved by using in sequence Structure-from-Motion (SfM) and Multi-View Stereo (MVS) algorithms. In the SfM case the camera is not initially calibrated and the goals are to estimate the extrinsic and the intrinsic (these are possibly unknown or perhaps constrained) parameters and reconstruct the sparse 3D scene. SfM starts by identifying correspondent feature points across many images whose position, orientation and distortion, are unknown. These parameters are then inductively inferred by the algorithm, which proceeds in an iterative manner by sampling an object multiple times. Hence, SfM needs numerous images to reconstruct the sparse 3D point cloud. In the MVS case, the camera is calibrated, thus the internal parameters (intrinsic) of each camera as well as its position and pose (extrinsic) are known and the goal is to estimate the dense 3D scene. Then, a scaled 3D model is obtained using ground control points (GCPs). This pipeline differs from traditional ones, where to estimate a 3D shape just a couple of stereo pictures are required together with data about the position, orientation, and geometrical distortions of a specific camera sensor. Differently from SfM methods, classic photogrammetry uses the stereo matching to identify correspondent points, whose 3D coordinates are then directly determined.

Early studies such as that of Gruen, Remondino, & Zhang (2004) demonstrate an early interest in exploiting images found on the internet and/or collected by tourists for use in photogrammetry applications in contexts where the collection of data presents difficult challenges. However, proposed 3D reconstruction methods were still difficult to apply and required the intervention of trained specialists. One of the first key papers paving the way for the combined use of SfM and unordered image collections by exploiting the power of bundle adjustment and self-calibration was written by Snavely, Seitz, & Szeliski in 2007. Later, SfM-MSV have allowed the development of many photogrammetric projects based on crowdsourcing such as BigSfM (Snavely et al., n.d.). These include the 3D modelling of city scale monuments using tourists' pictures (see the case study of Rome in Agarwal et al., 2011) as well as the reconstruction of heritage buildings and sites destroyed by terrorism such as Palmira (Wahbeh, Nebiker, & Fangi, 2016). Others address disastrous incidents such as the recent fire of Notre Dame de Paris, for which some experimental photogrammetric models have started appearing on ©Sketchfab (Bandera, 2019). Another relevant advance in the direction of fast, automated and low-cost photogrammetry from smartphone-acquired images is represented by the diffusion of cloud-based processing services such as ©Recap Photo or free mobile Apps for 3D scanning such as ©3DSizeME, ©3D Creator or ©Bevel, just to name a few. These services allow uploading images (in the case of ©Recap up to 300) on external servers for processing, and then download a textured 3D model from the cloud. Using a similar system, an international research team is now developing a mobile application with on-the-fly feedback for collaborative mapping projects as part of the European Replicate project (Nocerino et al., 2017). Arguably, these new technologies will help putting further down costs and need for expert supervision and, possibly, processing time as well.

This paper will not repeat a demonstration of the potential of the SfM approach for crowdsourcing applications. Rather, it seeks to highlight the specific needs of SfM-MSV methods regarding image acquisition to enable the maximum exploitation of novel photogrammetry pipelines in post-disaster contexts. We want to leverage empirical data to set the ground to produce user-centred image acquisition guidelines and fit-for-purpose reconstruction workflows that would enhance the results of future collective mapping efforts. The paper clarifies the technical challenges and the practical constraints of using crowdsourced image data to produce qualitative and/or quantitative data, but a detailed assessment of the results falls out of its scope. A relevant precedent in the literature is represented by the work of Griffiths et al. (2015) on the Heritage Together project, which emphasise the importance of generating novel research processes to maximise the outcomes of public archaeology initiatives. However, our paper differs from it in the analysis approach, as we study the behaviour of citizen scientists post factum, that is, without exchanging knowledge during the process. Operatively a visual ethnographic analysis (Bryman, 2016) is used to examine the way people documented their reality in a collaborative photographic mapping initiative carried out between the towns of Arquata del Tronto and Norcia after the 2016 central Italy earthquake. The study is conducted by considering the technical needs of SfM-MSV processes and results are tested by post-processing the imagery, illustrating three different examples: a successful, an acceptable and an unsuccessful one.

## 2. METHODS AND ANALYSIS

The analysis of the data collected starts by examining the quality of it in relation to the technical components involved in reconstructing a 3D model, specifically: (i) Image coverage and overlap, (ii) Camera sensor and image collection mode, and (iii)

Camera pose. A visual ethnographic analysis is used to illustrate how the behaviour of citizen-scientists during image acquisition affected the quality of the photographs so realistic guidelines to collaborative mapping can be outlined. This data is then post-processed using standard photogrammetry techniques and assessed in relation to image selection, 3D model reconstruction and scaling, followed by examples illustrating three different quality of results and related issues. As a preamble for the analysis, we present an overview of how a group of citizen-scientists got involved in this image acquisition process, by explaining their motivations within the Italian Humanitarian context. This contextualization is important to develop a framework for the provision of successful guidelines for collaborative mapping because it shows they are grounded in a user-centred perspective aimed at fostering best practices in the ubiquitous deployment of photogrammetry witnessed to date.

## 2.1 The motivation within a humanitarian context

As images reflect specific habits and personal narratives, including an interpretation of history, space and truth (Pink, 2006), photos have long been used as a research tool in visual ethnographic studies tackling certain aspects of culture and society (Schwartz, 1989). Setting the context in which images are produced provides comprehensive understanding of what is possible to achieve from a user-centred perspective. An analysis of an image dataset depicting the aftermath of the 2016 Central Italy earthquake is undertaken using photos gathered in two municipalities, in conjunction with additional information such as geographic maps, time data, official reports and other relevant documents from the web, and via its photogrammetric processing. Motivations, main drivers and promoters of the initiative in the Italian humanitarian context are outlined from the first call for action, made soon after the disaster by the civic hacking website Terremoto Centro Italia (2016), which opened only a few days after the first seismic wave. The needs behind the request of open photographic contributions were: (i) storytelling by reconstructing the post-earthquake narrative to communicate with people outside the affected area who cannot visit the territory; (ii) freedom to choose the subjects, including less popular areas and remote villages; (iii) better coverage, by mapping off-road paths where global mapping services do not go (e.g. rural/mountainous territories); (iv) frequent updates with possibility to compare situations at different times and with data from other services such as @Google Street View (GSV); (v) freeing data from everyone's photos to transfer on cartographic maps as Points of Interest (POIs); (vi) obtain a 3D reconstruction of the photographed subjects; (vii) strengthen civic activism and community participation.

Mapillary, a crowdsourcing platform able to manage and extract data from images using computer vision techniques (Mapillary, 2018), was suggested as a suitable tool because it authorises the open use of the images uploaded (CC BY-SA licence). No limitations were given in terms of survey organisation and objective or camera type (smartphones, action cameras and SLR ones were equally welcomed). One year later Act!Onaid Italia organised in Arquata del Tronto a Monitoring and Civic Action School (SMAC) involving a collaborative mapping challenge aimed at empowering local people through civic activism. There, a group of 39 people surveyed 18 km of streets (716 images) during a 1-hour walk (green areas in fig.1). Some guidance was given to participants to read before the survey (see table 3): mainly a summary of those in the Mapillary website. The use of @Flickr was suggested as a further alternative to capture pictures with geotags using GPS data. Geotags were recommended to label records at specific times and places to allow cross-comparisons with data captured during official inspections.

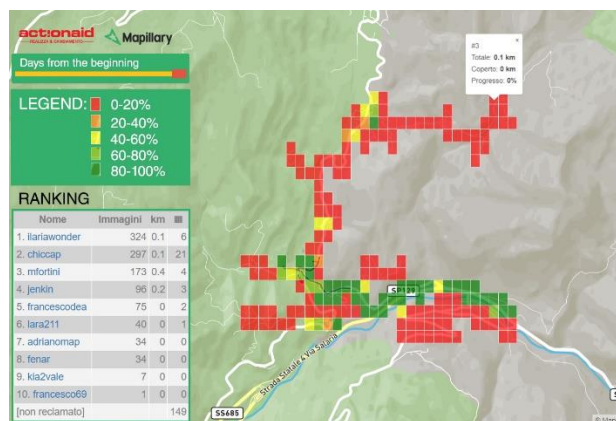


Figure 1. Areas to map, Mapillary challenge Arquata del Tronto

## 2.2 The technical components of reconstructing a 3D model

These humanitarian initiatives illustrate an ambition for broader coverage and more frequent update of photographic mapping data. In order to compare crowdsourced data with commercial data (tab. 1), we first map and overlay data about the spatial and temporal distribution of imagery collected by global mapping services such as Google, and the people (fig. 2). Then, image density, orientation and content are analysed in order to account for concentration of information in certain points or on certain subjects (fig. 3). The number of contributors and the timing of data capture for each photographed area is contextually recorded in GIS. Finally, a classification of sensor types, image resolution, exposure levels as well as frame typologies (wide angle, 360 and standard lens) is done to account for the level of variation of image quality in the dataset (tab 2). This first part of the study is largely conducted using QGIS (an open GIS platform) as it allows to parallelly connect to Google and Open Street Map web layers and Mapillary's image repository (with the possibility to filter data according to shooting time or contributors' names) and overlay the different information layers so to get a comprehensive account of the described phenomena. By enabling the semantic enrichment of map-data through tags, QGIS also serves as a base for the analysis of image content enabling the verification of the level of compliance to the guidelines provided to participants in the SMAC (see table 3).

### 2.2.1 Coverage and overlap

The greatest part of Google Street View surveys in both Arquata del Tronto and Norcia is dated 2011, with very little official later updates referring to the area of Norcia only. In Norcia, some citizen-scientists have decided to contribute directly to the GSV platform rather than to Mapillary exclusively, so the latter presents a poorer coverage and a lesser number of contributors if compared with the case of Arquata del Tronto (see table 1).

	Norcia	Arquata
GSV km	277	87
Mapp.km	15	18
OSM km	758	253
GSV coverage	36,54%	34,43%
GSV update number	4	3
GSV year update	2011; 2017; 2018	2011
Map. coverage	1,98%	13,82%
Map update number	3	17
Map year update	2016; 2017	2016; 2017

Table 1 coverage and update frequency in Norcia and Arquata



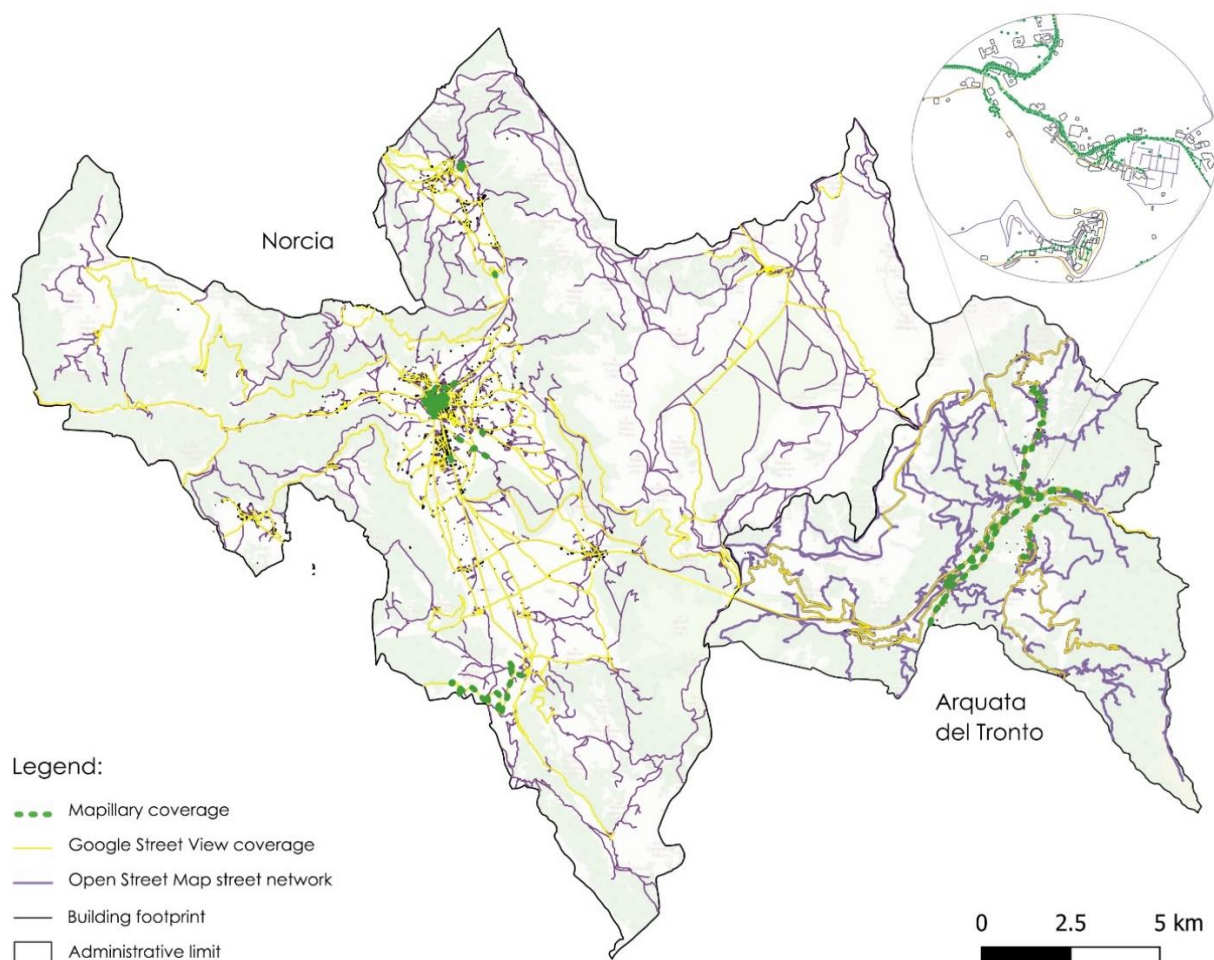


Figure 2. GSV, OSM and Mapillary coverage (in the zoom a dot represents a single pictures) of Norcia and Arquata del Tronto, Italy

Despite the coverage rate of Arquata del Tronto seems encouraging (even more so if we consider just urbanised area), from the perspective of SfM-MVS methods, a more fine grain analysis of mapping density, overlap and content is needed (Carbonneau et al., 2017). As most pictures were taken without a 360 camera but rather while walking around with a smartphone or a tablet, information about their orientation and content (photographed subject and distance from it) become relevant to determine their suitability for the photogrammetric reconstruction of an artefact.



Figure 3. Image orientation and density in Borgo 1, Arquata

To illustrate the issue we take the example of Borgo 1, a site selected by the municipality of Arquata to build a Temporary Housing (TH) camp, where ~ 210 pictures were taken with different orientations (see blue and red arrows in figure 3), but mostly from the same location: the main road passing on the Northern side of the temporary settlement. Along this path (~ 265 m) the mapping density is ~ 0.58 images per meter. However this value decreases dramatically (half of the initial value in the best cases) if we filter the data by date (some images refer to the emergency phase with the tents, others to the urbanisation of the site or to the finished TH camp), or by photographed subject (some images focus on the road, others on the TH site or on the temporary service structures built on the other side of the road). Moreover, in some pictures the presence of external objects such as cars and street signals in the foreground obstructs the view and hence diminishes the visual information available as well as the actual overlap between pictures. Even if the image density was higher, the presence of elements in the foreground could hinder the capacity of the software to correctly detect and match corresponding feature points. In similar cases, a proper image selection, effective masking of noisy elements or the introduction of mark points may be critical to boost the 3D reconstruction. Despite the TH site being an area rather than an isolated object, we can still consider it a case of close-range photogrammetry as the maximum distance of the subjects from the camera sensor is always below 200 meters. However, because the range of distances varies continuously from 200 to 20 m, the 3D digital reconstruction task becomes highly challenging due to the noise introduced by this multi-scale distancing.

### 2.2.2 Sensors and image collection modes

From table 2 we can see that a variety of sensors were used, with outcomes showing that all camera types were able to generate pictures with an acceptable resolution and within a relatively little variation range. Apart from one 360 camera, which was probably operated by a citizen-professional (e.g. a photographer or a surveyor), all the images collected in the two areas have standard frames (St.). Additionally, it is possible to infer that all images come from a single snapshot (S) instead of from video feeds, as there is little regularity among the frames. Most of the images are taken by manual (M) shooting while walking (W) but in a few cases an automatic shooting mode (A) was preferred, mostly when driving a car (D). At times, the same person switches between data collection modes. Walking was preferred within urban areas, where pedestrian routes are generally safer even in the absence of pavement. Driving was preferred for mapping the periphery, where there is no suitable space for pedestrians to walk or rest. The fact that six citizen-scientists appear documenting both cities shows that civic activists may be available to travel and collect data across a larger territory.

CitizenS	Sensor type	resolution	mode	frame
Arquata del Tronto				
TCI	Apple iPhone7	3264 x 2448	W/M D/A	St. /S
	Ooneplus A000	4096 x 2160	W/M	St. /S
verobog	Apple iPad5	3264 x 2448	W/M D/A	St. /S
	Apple iPad3	2592 x 1936	W/M	St. /S
chiccap	Sam SM-G920F	3264 x 2448	W/M D/A	St. /S
cquintili	Apple iPad3	2592 x 1936	D/A	St. /S
kia2vale	Sam SM-N910F	5312 x 2988	W/M	St. /S
mfortini	Oneplus A5	4608 x 3456	W/M	St. /S
micchela	Apple iPhone9	4032 x 3024	W/M D/A	St. /S
Ilariaw.	Apple iPhone	4032 x 3024	W/M	St. /S
adriano	Sony D510	3264 x 2448	W/M	St. /S
jenkin	LGE LG-H870	4160 x 3120	W/M	St. /S
francesc	Apple iPhone	4032 x 3024	W/M	St. /S
fenar	Sam SM-G900F	5312 x 2988	W/M	St. /S
Franc69	Sam SM-N9005	4128 x 3096	W/M	St. /S
Lara211	HUA VTR-L09	3968 x 2976	W/M	St. /S
starchild	HUA FRD-L09	3968 x 2240	W/M	St. /S
kymolos	Apple iPhone5	3264 x 2448	D/A	St. /S
saraveg	Sam SM-G920F	3264 x 1836	W/M D/A	St. /S
consta	Canon eos 450D	4272 x 2848	W/M	St. /S
	ASUS Z00ED	3264 x 1836	W/M	St. /S
gtfabio	Apple iPhone5	3264 x 2448	W/M	St. /S
Vale.p	Ricoh Theta S	5376 x 2688	W/M	360/S
Norcia (besides 6 in common with Arquata)				
gem	Apple iPhoneSE	4032 x 3024	W/M	St. /S
blind	Sam SM-J730F	4128 x 3096	W/M	St. /S

Table 2 sensor types and shooting mode

### 2.2.3 Camera pose

Another aspect to consider is the object or area being photographed in relation the position of the camera, i.e. camera view. The dataset shows that many people decide to map main roads (very few paths are represented) frontally. Others map what is on the right or left side of a road by moving parallel to it and perpendicular to the street. It was not uncommon to see people switching from one mode to the other, thus creating gaps in the data collected and inconsistencies related to alternating oblique with frontal views (see example in fig. 6). Occasionally, instead

of moving in one direction or the other, citizen-scientists were standing in a fixed position and rotating around themselves attempting to get a panoramic view. However, the black spots in the post-processed stitched image of figure 4a, show the alignment with the terrain was not maintained all the way through, with image sets captured in this way covering angles of much less than 180 degrees (see fig. 4b), which is insufficient to recreate a full panorama (left aside a full 360 picture for which vertical rotations are required). This observation is relevant in relation to the declared aim of the mapping event to generate a proxy of the service offered by GSV. For photogrammetry purposes this represents an issue as stereo is a requirement, meaning that it is necessary to physically move the camera in one direction (translation) instead of around a point (rotation). Furthermore, the centre of rotation cannot be too close to the object being photographed (as see in fig. 4c), as this would make it inevitably prominent in the scene. Finally, since pictures are taken along the streets, the distance to the building façades varies continuously. Cameras mounted on cars, despite providing more consistency in relation to orientation, tend to: (i) present motion blur; (ii) capture other vehicles/objects in the foreground, occluding facade views completely or (iii) portrait poorly textured surfaces or views that are indistinguishable from one another, becoming useless in 3D reconstruction.



Figure 4. Stitched panoramic images, Arquata del Tronto

This analysis shows that photogrammetric processing from crowdsourced data needs to cater for: issues related to low image quality, occlusions and scale as well as differences in image sizes, proportions and orientation. Additionally, they need to address all the challenges coming from differences in time of acquisition such as missing parts (those present in one image and absent in another) and related to changes in illumination and weather (e.g. sharp shadows and/or significant glare and contrast). Computer vision algorithms may help dealing with these issues but regardless of how advanced they are many pictures are still needed to compensate for uncertainties and inconsistencies in image acquisition. If this is not possible or images are not available, as in our case studies, then the collection phase becomes critical to maximise the usefulness and completeness of the final output, making guidelines an essential instrument to drive collective efforts. Table 3 presents a summary of guidelines given to citizen-scientists participating to the SMAC including a brief account of how much on average they have been followed by participants.

	Guidelines	Respected		
		Y	N	N.A.
Camera handling	• Hold the camera straight	x		
	• Walk using the single-shot mode	x		
	• Pedal with sequential shooting			x
	• Shoot from a car in sequence mode			x
	• Landscape orientation (until in tight spaces or extra height makes a difference)	x		
Camera orientation	• Roads, cycle and pedestrian paths (in direction of the route or in the opposite)		x	
	• Buildings and buildings' entrances (front side view)		x	
	• Squares, parks and open spaces (everything of interest)	x		
Content (Don'ts)	• Avoid taking pictures when privacy is an issue (e.g many people present; risk to identify a person; close-ups of private properties, private or restricted access areas; military areas and sensitive areas)		x	
	• Avoid photos against the sun	x		
	• Check your photos in the APP before uploading		x	
Survey Management	• Plan ahead			x
	• Go when fewer cars and people are around (Sundays)			x
	• Move when the sun is out	x		
	• Organized for an adequate time in path and acquisition (consider battery life)	x		
	• In a car, make sure the windows are clean and that the car is not in the picture			x
	• Use an external Bluetooth GPS for better accuracy (especially in cars)		x	
	• Do it with friends	x		

Table 3. Overall compliance (Yes, No, Not Applicable) to guidelines provided to participants in the SMAC, Arquata del Tronto

### 2.3 Image processing

Three image subsets, representing a successful (21 photos), an acceptable (31 photos) and an unsuccessful (10 photos) example, are extracted from the Mapillary dataset and processed using @Photoscan. Images are elaborated using a standard processing pipeline in order to identify difficulties that could be better addressed during the collection and processing phases. For each image subset, the reconstruction process consists of photo orientation, bundle adjustment, point cloud and mesh generation, and finally texture mapping. Observed problems and positive outcomes are reported and discussed in the following paragraphs.

#### 2.3.1 Image filtering and selection

In this study images are selected manually, but it is anticipated that for large scale applications an automatic selection process based on their quality (e.g. in focus, not blurred) and content (e.g. quantity of new visual information, overlap) would be preferable. For instance, it may be possible to automatically filter images according to their sharpness, which clearly improves processing outcomes by allowing the extraction of good feature points. This can be done in a first instance, simply by exploiting the Estimate Image Quality algorithm of @Photoscan. However, because it mainly accounts for contrast between pixels, it is not suitable to detect motion blur, which could be treated in a second filtering step, deploying more specialised algorithms (Sieberth et al., 2016). Depending on the amount and type of data available it may be advisable to relax the sharpness requirements to compensate for poor coverage as shown by Griffiths et al., (2015). Subsequently, images can be selected according to the number of new features they present; which can be found by firstly matching feature points in image pairs (using any measure of distance) and then calculating the difference between the number of new and old ones. This information is useful also to compute image overlaps. For example, Nocerino et al. (2017) automatically select frames from video feeds presenting a minimum of 20% and a maximum of 80% overlap (to prevent inaccuracies in the SfM triangulation), which is estimated by dividing the number of new features by the amount of those already present in all

previously selected images, storing feature information and updating them at every step.

#### 2.3.2 Model scaling

Photogrammetric models are inherently dimensionless, with a number of ways available for them to be scaled, all of them requiring the retrieval of at least one known distance. One possibility is to use the official cartography to get the real-world coordinates of some key points. Then, these coordinates can be assigned to the corresponding points on the point cloud or used to compute a distance useful to scale the model. This method is limited to the resolution of the typically available cartographic data, which normally presents an error of approximately 1 meter (in maps with scale 1:5000 and even more in maps with scale 1:10:000). This limit becomes especially critical in rural areas where a finer grain cartography is often simply not available. An alternative is then to scale the model using measurable objects captured in the pictures, for instance a road signal because it has a standard dimension. If measurable objects are not present in the images, another possibility would be their deliberate inclusion while photographing (e.g. a scale bar). This option limits the capacity of citizen-scientists to collect data spontaneously. Using GPS data is not a viable alternative as accuracy is normally worse than that of coarse cartographic data and precision instruments are not commonly available to people.

#### 2.3.3 Examples: Successful, acceptable and unsuccessful

Figure 5 shows the positive results of an experiment concerning the 3D reconstruction of a fragment of Norcia's historic urban walls. The output looks promising and shows the potential of the method to provide valuable 3D information about, for instance, damage to vertical structures of urban heritage. The images were taken from the same citizen-scientist, following a path parallel to the walls from a fixed distance while driving. Pictures were automatically captured by an action camera installed on the vehicle and present an acceptable sharpness and degree of overlap. Overall, the 3D reconstruction is rather satisfying considering that heavy manual interventions during post-processing were not necessary.





Figure 5. 3D Reconstruction of Norcia historic urban walls

In other cases, photogrammetric tests have produced quite acceptable results after the introduction of a set of mark points and the subsequent use of @Photoscans' optimisation function to improve the calculation of internal orientation parameters. An example is shown in figure 6. The 3D model was obtained using data collected by a citizen-scientist walking frontally in the direction of a street and shooting pictures in a not-automatic way. This collection mode prevents an accurate reconstruction of the vertical façades as they are always recorded from an oblique angle, which causes an uneven distribution of homologous points in the scene worsened by the presence of the plain texture of the plaster finishing. The consequent loose camera calibration in turn provokes the introduction of noise in the 3D model as well as the presence of holes. Moreover, the variation of relative position between the object and the photographer causes the issues in the 3D reconstruction, due to the variability of spatial correspondence between pixels in the foreground and pixels in the background, introducing additional noise in the model. Here, the manual introduction of mark points in the post processing allows to reduce the noise and enhance the 3D output.



Figure 6. 3D reconstruction of a fragment of Norcia city centre

At times, despite the images being frontal and rich in information (see fig.7), serious problems arise in the reconstruction due to the "geometry" used for image capture, as the lack of stereo simply prevents the creation of a 3D model. In this last example 7 images out of 10 were taken with a polar symmetry (fig. 4a), 6 had to be discarded from the initial dataset, leaving just 4 images for the final processing. Excluded images were the most representative and complete but caused the result to be ultimately unusable.



Figure 7. 3D point cloud of some ruins, SMAC data, Arquata

### 3. DISCUSSION AND RESULTS

Well beyond a simple description of the disaster trajectory, this paper suggests exploiting image crowdsourcing to better deal with the complex interweaving dynamics of post-disaster recovery and reconstruction. By drafting the scope of relevant applications in relation to urban disasters this study identifies opportunities for the development of fit-for-purpose semi-automatic collaborative mapping procedures. Lessons learned constitute a grounded framework to improve future guidelines enhancing citizen-scientists' efforts and supporting the use of photogrammetry as a participatory recovery tool after disasters. Experiments highlighted that objects captured at multiple scales (pixel-wise) and difficulties in model scaling are critical issues. Results indicate that as it is unlikely that citizen-scientists will carry sophisticated external GPS devices, it is crucial to give better guidance on how to introduce measurable elements in the pictures that allow scaling the 3D models in the post-processing phase; especially in cases where complementary data such as UAV imagery and detailed cartographies are not available. Failing to do this would put into serious question the usefulness of the entire street-level photogrammetric reconstruction process. Many of the analysed datasets present poor image quality, insufficient overlapping, occlusions or are excessively close to or far from the photographed object. Compliance to guidelines in the SMAC group survey proved generally high apart from two key aspects: camera orientation and content of the images, which are closely related issues. Future guidelines should include targeted indications aimed at clarifying points that proved to be too general or unclear and add a few missing ones, such as:

- Guidance for different goals (e.g. GSV-like panoramas vs qualitative or quantitative 3D reconstruction)
- Distance from the target and scale
- Measurable references for scaling
- Awareness of obstacles and occlusions
- Use of videos to extract frames instead of single-shots
- Use of multiple sensors (e.g. GPS + accelerometers)
- How to check images (lighting exposure, blur/focus)

Additionally, external risk factors (e.g. debris, interference with other operations, exposure to hazards etc.) should be considered. When improving data collection is not possible or people simply do not comply to any guideline, some tailoring may be needed in the photogrammetric processing pipeline such as: (i) Automatic filtering of images, (ii) Retrieval of customised features (via deep learning); (iii) Smart grouping of images into subsets/chunks; (iv) Use of information from external sensors to refine results.

To implement these features, is recommendable to use open photogrammetry software with scripting possibilities such as @MicMac or @ColMap. If, on the contrary, a full automation of processing is required, it may be worth considering the deployment of cloud-based services such as @Recap Photo. A comparison between models obtained in one and in the other way, using increasingly challenging datasets (such as the 3 presented in this paper) could indeed represent an interesting line for future research. Future work should then address the problem of merging street-level data with UAV or/and high-resolution satellite imagery and evaluate the final precision of the models obtained. The approach to the analysis adopted in this paper could be extended to other fields beyond that of disaster management, to address cases where rapid collaborative mapping is desirable. The methodology is transferable to develop grounded frameworks and user-centred guidelines aimed at supporting the surveying activity of lay people in all cases where short terms action is needed (e.g. to support untrained professionals in circumstances where time is an issue such as some policemen in newly found crime scenes) or a situation in which 3D information is required (e.g. for insurance claims).

## REFERENCES

- Agarwal, S., Furukawa, Y., Snavely, N., Simon, I., Curless, B., Seitz, S. M., & Szeliski, R., 2011. Building Rome in a day. *Communications of the ACM*, 54(10), 105.
- Bandera, M., 2019. Notre Dame crowdsourced photogrammetry, Sketchfab. Retrieved May 31, 2019, from <https://sketchfab.com/3d-models/notre-dame-crowdsourced-photogrammetry-c6f03e2284e049f38f18378b012c979>
- Barbero-García, I., Lerma, J. L., Marqués-Mateu, Á., & Miranda, P., 2017. Low-Cost Smartphone-Based Photogrammetry for the Analysis of Cranial Deformation in Infants. *World Neurosurgery*, 102, 545–554.
- Billen, R., Cutting-Decelle, A.-F., Marina, O., de Almeida, J.-P., M., C., Falquet, G., ... Zlatanova, S., 2014. 3D City Models and urban information: Current issues and perspectives. *3D City Models and urban information: Current issues and perspectives – European COST Action TU0801* (pp. 1–118).
- Borsekova, K., & Nijkamp, P., 2019. *Resilience and urban disasters : surviving cities* (K. Borsekova, Ed.).
- Breunig, M., Borrmann, A., Rank, E., Hinz, S., Kolbe, T., Schilcher, M., ... Mazroobsemani, N., 2017. Collaborative multi-scale 3D city and infrastructure modeling and simulation. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-4/W4*, 341–352.
- Bryman, A. (2016). *Social research methods*.
- Carbonneau, P. E., & Dietrich, J. T., 2017. Cost-effective non-metric photogrammetry from consumer-grade sUAS: implications for direct georeferencing of structure from motion photogrammetry. *Earth Surface Processes and Landforms*, 42(3), 473–486.
- Copernicus, 2016. EMSR177: Earthquake in Central Italy. Retrieved June 6, 2019, from Copernicus EMS - Mapping website: <https://emergency.copernicus.eu/mapping/list-of-components/EMSR177>
- Erdelj, M., Natalizio, E., & Akyildiz, I. F., 2017. *Help from the Sky: Leveraging UAVs for Disaster Management*. Retrieved from [www.computer.org/pervasive](http://www.computer.org/pervasive)
- Griffiths, S., Edwards, B., Wilson, A., Karl, R., Labrosse, F., LaTrobe-Bateman, E., ... Tiddeman, B., 2015. Small Works, Big Stories. Methodological approaches to photogrammetry through crowd-sourcing experiences. *Internet Archaeology*, (40).
- Gruen, A., Remondino, F., & Zhang, L., 2004. Photogrammetric Reconstruction of the Great Buddha of Bamian, Afghanistan. *The Photogrammetric Record*, 19(107), 177–199.
- Johansson, B., Johansson, B., & Cipolla, R., 2002. A System for Automatic Pose-Estimation from a Single Image in a City Scene. In *IASTED International Conference Signal Processing, Pattern Recognition, and Application*.
- Kirchhöfer, M., Chandler, J., & Wackrow, R., 2011. Cultural Heritage Recording Utilising Low-Cost Closerange Photogrammetry. *Geoinformatics FCE CTU*, 6(0), 185–192.
- Mapillary, 2018. Mapillary's computer vision technology – Mapillary. Retrieved May 31, 2019, from <https://help.mapillary.com/hc/en-us/articles/115001770329-Mapillary-s-computer-vision-technology>
- Micheletti, N., Chandler, J. H., & Lane, S. N., 2015. Investigating the geomorphological potential of freely available and accessible structure-from-motion photogrammetry using a smartphone. *Earth Surface Processes and Landforms*, 40(4), 473–486.
- Nocerino, E., Poiesi, F., Locher, A., Tefera, Y. T., Remondino, F., Chippendale, P., & Van Gool, L., 2017. 3D Reconstruction with a Collaborative Approach Based on Smartphones and a Cloud-Based Server. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2(W8)*, 187–194.
- Petrides, P., Kolios, P., Kyrkou, C., Theocharides, T., & Panayiotou, C., 2017. *Disaster Prevention and Emergency Response Using Unmanned Aerial Systems*.
- Pink, S. (2006). *The Future of Visual Anthropology: Engaging the Senses*.
- Poblet, M., García-Cuesta, E., & Casanovas, P., 2014. *Crowdsourcing Tools for Disaster Management: A Review of Platforms and Methods*.
- Schaffalitzky, F., & Zisserman, A., 2002. Multi-view Matching for Unordered Image Sets, or "How Do I Organize My Holiday Snaps?" In *Lecture Notes in Computer Science book series (LNCS)*, vol, pp. 414–431.
- Schwartz, D., 1989. Visual ethnography: Using photography in qualitative research. *Qualitative Sociology*, 12(2), 119–154.
- Sieberth, T., Wackrow, R., & Chandler, J. H., 2016. Automatic detection of blurred images in UAV image sets. *ISPRS Journal of Photogrammetry and Remote Sensing*, 122, 1–16.
- Snavely, N., Seitz, S. M., & Szeliski, R., 2007. Modeling the World from Internet Photo Collections. *International Journal of Comput Vision*.
- Snavely, N., Crandall, D., Hauagge, D., Wilson, K., Cao, S., n.d. BigSFM: Reconstructing the World from Internet Photos. Retrieved May 31, 2019, from <http://www.cs.cornell.edu/projects/bigsfm/>
- Terremoto Centro Italia, 2016. Racconta e visita i luoghi del terremoto grazie a Mapillary. Retrieved from <https://terremotocentroitalia.info/2016-10-10-usare-mapillary-per-raccontare/>
- Toschi, I., Remondino, F., Kellenberger, T., & Streilein, A., 2017. A Survey of Geomatics Solutions for the Rapid Mapping of Natural Hazards. *Photogrammetric Engineering & Remote Sensing*, 83(12), 843–859.
- UN environment, 2018. Citizen science: crowdsourcing data to support disaster response | UN Environment. Retrieved May 31, 2019, from <https://www.unenvironment.org/news-and-stories/story/citizen-science-crowdsourcing-data-support-disaster-response>
- UNDRR, 2015. Sendai Framework for Disaster Risk Reduction - UNDRR. Retrieved May 31, 2019, from <https://www.unisdr.org/we/coordinate/sendai-framework>
- UNISDR, 2017. *Words into Action Guidelines: National Disaster Risk Assessment Special Topics H. Citizens' Participation and Crowdsourcing*. Retrieved from [www.wired.com/2006/06/crowds/](http://www.wired.com/2006/06/crowds/)
- Wahbeh, W., Nebiker, S., & Fangi, G., 2016. Combining public domain and professional panoramic imagery for the accurate and dense 3D reconstruction of the destroyed temple in Palmyra. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume III(5)*, 81–88.
- Wilson, L., Rawlinson, A., Mitchell, D. S., McGregor, H. C., & Parsons, R., 2009. The Scottish ten project: collaborative heritage documentation.
- Zhang, S., Shan, J., Zhang, Z., Yan, J., & Hou, Y., 2016. Integrating smartphone images and airborne lidar data for complete urban building modelling. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B5*, 741–747.