Application of ‘Structure from Motion’ (SfM) technique in physical hydraulic modelling

Sanat Kumar Karmacharya¹,², Meg Bishwakarma², Ujjwal Shrestha¹ and Nils Rüther *¹
¹Department of Civil and Environmental Engineering, NTNU, Trondheim, Norway
²Hydro Lab Private Limited, Lalitpur, Nepal
* Corresponding author (nils.ruther@ntnu.no)

Abstract. There are many methods available for measurement of bed morphology in physical hydraulic model studies considering mobile bed sediment. Among which, there are sophisticated instrumentations which provide quality results in shorter time but are vastly expensive and requires special training for execution. Whereas the conventional surveying methodology, which is simple and inexpensive, requires plenty of time for the measurement and processing of the data. That is why the recent developments in ‘Structure from Motion’ (SfM) technique have made it a potential candidate for an inexpensive and efficient tool for measurement of bed morphology in physical hydraulic model studies. SfM method allows to simultaneously determine both the parameters of the camera and the 3D structure of a scene by combining 2D images taken from multiple viewpoints. SfM tools can create a dense point cloud out of a set of partially overlapping photographs taken even by a budget friendly digital camera. The SfM method have already been used as an alternative for topographic surveying to create high-resolution digital elevation models (DEM). Some researchers had also used it for measurement of bed morphology in laboratory experiments. In this study, different freely available SfM tools were used to create a dense point cloud from a set of photographs representing a short reach in a river model in the hydraulic laboratory at Hydro Lab. The selected tools were compared with each other and against a commercial software, based on the methodologies used, processing time and quality of the output. Then the results from SfM method were compared with actual measurements in the physical model done with a conventional surveying technique using a theodolite and a level machine. The results showed that free SfM tools can also produce efficient results compared to commercial tools and SfM method can be used as an inexpensive and efficient alternative for bed morphology measurements in physical hydraulic models.

1. Introduction
With aggressive advancement in technology, currently various methodologies are available, over conventional ground surveying techniques, for creating a high-resolution digital elevation models (DEM) of a topography. Aerial photogrammetry, airborne lidar and ground based terrestrial laser scanners (TLS) are some advanced technologies, which have revolutionized the quality of DEMs extending their spatial extent, resolution and accuracy [1]. Recently, easy access to unmanned aerial vehicles (UAV) and drones has made the aerial imaging surveys more convenient and inexpensive.

Besides the development in large-scale terrestrial surveying, there has also been a huge advancement in data acquisition and processing technologies for hydraulic laboratories. Producing DEMs of physical hydraulic models and/or recording bed morphologies in fluvial sediment transport studies in hydraulic laboratories can be carried out more accurately by using laser scanning or acoustic sounding systems. These systems can be tailored for semi or fully automatic data acquisition curtailing...
the experiment time. Despite these sophisticated instrumentations are useful in producing high quality DEMs in shorter time, they require high logistical cost and specialized user expertise. Therefore, many hydraulic laboratories still use conventional measurement techniques that are inexpensive and simple though more time demanding. In conventional techniques, measurements are taken at selected points/cross-sections and those data are interpolated in-between. The accuracy of such measurements can be improved by increasing the density of measured points especially at desired details to be captured but it will ultimately increase the time required for measurement and data processing. On the contrary, the laser and acoustic scanning can record highly dense point cloud in shorter time ensuring higher resolution DEM.

In this study, an advanced yet inexpensive and easy to perform photogrammetric method, called ‘Structure from Motion’ (SfM), was used to produce high resolution DEM of a physical hydraulic river model. Basic principle governing SfM method is similar to stereoscopic photogrammetry in which a 3D structure is developed from a series of overlapping 2D images [1]. Unlike conventional photogrammetry, SfM method utilizes advanced algorithms by which it automatically solves the relative camera positions, orientation and geometry of the target object based on the features extracted from the set of images [2]. Possibility of using low-cost consumer level digital cameras and availability of free and open-source processing tools has given SfM method boundless potentials. Nowadays, SfM has already been widely used in various fields like archaeology, geosciences, robotics, terrestrial surveying, real state, film and entertainment, sports etc. The possibility of using SfM method in an inexpensive and simple way to record 3D information from laboratory models was assessed in this study.

2. Methodology

2.1. Structure from motion (SfM)

There are various algorithms available for application of SfM but the general principle remains the same and has been described by [1] [3] and [4]. SfM requires a set of overlapping images, capturing the object from multiple viewpoints, as an input. From the images, common feature points across the image set called key point descriptors are identified using a scale invariant feature transform (SIFT) algorithm. With these feature points and camera parameters extracted from the images, camera location, orientation and position of the feature points are simultaneously resolved in a relative 3D coordinate system. Once the spatial positions of the images are established, a sparse bundle adjustment (BA) algorithm is used to create 3D points covering the area of interest. Then, dense point cloud is produced by intensifying the sparse point cloud with multi view stereo (MVS) techniques. Once the dense point cloud is obtained, it can be used for further processing like developing a DEM, mesh generation, creating a 3D model etc. as per requirement.

2.2. Software

Nowadays, due to the application of SfM method in diverse fields, there are various software available for its implementation. Following six software were selected for application of SfM method in this study. First two are commercial software whereas remaining four are free to use.

2.2.1. PhotoScan. PhotoScan (now available as Metashape) is commercial software developed by Agisoft LLC, Russia. It is an easy to use stand-alone software which offers all the capabilities from processing images to 3D model generation and texturing and includes additional pre-processing and post processing features. Professional edition of the software is priced USD 3,499 and the standard edition is available for USD 179. However, standard edition itself is enough for generation of 3D models and lacks only other extra features that professional edition offers.

2.2.2. ReCap. ReCap (named after abbreviation of Reality Capture) is developed by Autodesk Inc USA and is a cloud-based service tailored for generation of 3D models from photographs or laser scans. ReCap Photo, which is included within ReCap pro, is specifically targeted for support UAV and drone photo capture workflows. ReCap Photo can be used to create photo textured meshes, photo-
based point clouds with geo-location and high-resolution orthographic views with elevation maps. It can also be used in object mode to create 3D models of objects. It is a commercial software and ReCap pro subscription is needed to use the service. The subscription costs USD 40 for a month, USD 305 for a year and USD 915 for three years. The service uses cloud credits for each project and additional cloud credits can be purchased separately once subscribed. Currently, ReCap pro subscription allows you to process up to 1,000 images for one project in aerial mode while the limit is 300 photos for one project in object mode. The software can be used freely under academic license but has the limitation of 100 photos for one project and academic users may also have to wait in long queue for processing. Since it is a cloud-based service, no expensive hardware is required for the processing and the images can be uploaded via smartphone app too. The disadvantage could be it works like a black box and users have very limited control over the quality of the output.

2.2.3. VisualSfM. VisualSfM, developed by Dr. Changchang Wu, is a GUI application for 3D reconstruction using structure from motion (SFM). The reconstruction system integrates different algorithms like SIFT on GPU (SiftGPU), Multicore Bundle Adjustment and Towards Linear-time Incremental Structure from Motion [5–7]. VisualSfM runs fast by exploiting multicore parallelism for feature detection, feature matching, and bundle adjustment. It is one of the first free photogrammetry program to utilize the power of graphical processing unit (GPU). It can process up to sparse reconstruction and utilizes Yasutaka Furukawa's PMVS/CMVS tool for dense reconstruction, which has to be integrated and can be run from the VisualSfM's GUI. However, the sparse point cloud from VisualSfM can be processed in other dense reconstruction tools as well to produce dense point cloud. VisualSfM is free and open source but the licensing restricts its commercial use.

2.2.4. Regard3d. Regard3d is a free and open-sourced structure from motion software which is developed by Roman Hiestand. It offers complete SFM processing up to dense point cloud generation. It can also generate surface, from the dense point cloud, either with colored vertices or with a texture. It has integrated several algorithms and users have a bunch of options to control the quality of the output.

2.2.5. Meshroom. Meshroom is a free and open-source 3D reconstruction software based on the Alice Vision framework. It can perform complete 3D reconstruction up to textured surface creation. It has a node-based workflow that gives its users a lot of control on setting the processing and once its set whole processing can be completed in one click. The advantages of this software are it supports augmented reconstruction i.e. more pictures can be added for better detailing while the processing is on-going. Additionally, it can also perform live reconstruction. The disadvantage of Meshroom could be that it requires CUDA enabled GPU with at least computing capability of 2.0. It should be noted that it does not warn users if they lack the mandatory GPU requirement and the processing seems to be frozen without any notifications. Despite of its promising features and capabilities, there is not any official user manual or guide available yet.

2.2.6. COLMAP. COLMAP is a free and open-sourced Structure-from-Motion (SfM) and Multi-View Stereo (MVS) pipeline with a graphical and command-line interface. It offers a wide range of features for reconstruction of ordered and unordered image collections [8]. It uses MVS technique for dense point cloud reconstruction and uses screened Poisson surface reconstruction algorithm to recover 3D surface geometry from the dense point cloud. However, COLMAP requires CUDA-enabled GPU (at least CUDA version 7.x) to perform dense reconstruction and surface creation processes. But processes up to sparse reconstruction can be carried out without CUDA-enabled GPU and the output can be exported to do dense reconstruction with other tools. For beginners, COLMAP has an automatic reconstruction tool that simply takes a folder of input images and produces a sparse and dense reconstruction in a workspace folder.

2.2.7. Meshlab. Meshlab is the open source software for processing and editing 3D point clouds and 3D triangular meshes and creating 3D models [9]. In this study, Meshlab was used for transformation
of 3D dense point cloud in arbitrary coordinates created by SfM software to 3D dense point cloud in actual coordinates.

2.3. Hardware

2.3.1. Camera. Sony a6300 (model ILCE-6300) camera was used for taking pictures for this study. It was a mirrorless digital camera with 24-megapixel Exmor RS sensor, and 425 phase detection autofocus points. The camera setting has a significant impact on the image quality which ultimately affects the quality of the output. After various trials with different settings, following camera settings were used for this study:

- **Shooting mode**: Shutter Priority (S)
- **Shutter speed**: 1/80 sec
- **ISO**: 640
- **Aperture**: F9-F11
- **Focal length**: 16 mm
- **Image format**: JPEG

2.3.2. Workstation. The specification of the workstation determines the total processing time except for the cloud-based processing as with ReCap. For fair comparison, same workstation with following specifications were used for processing with all the selected software:

- **Operating System**: Microsoft Windows 7 Ultimate 64-bit
- **Processor**: Intel® Core™ i7-4790 CPU @3.60GHz
- **Installed memory (RAM)**: 32 GB
- **GPU**: NVIDIA GeForce GTX 750 Ti -18 GB available memory

2.4. Ground control points (GCPs)

With SIFT algorithm, the 3D structure is created based on relative spatial relationships between the original image locations in an arbitrary 3D coordinate system [10]. That means the structure (shape) of the object is recovered but the size is scaled to some arbitrary scale factor. Hence, the final output, either the dense point cloud or 3D surface model, must be transformed using rotation, translation and scaling. To perform this transformation, several ground control points (pre-defined set of points with known coordinates) are needed in the study area. For large scale terrestrial surveying, in order to obtain high accuracy of final output (RMSE<1) 1GCP per 200m² is needed to be placed in the interest area [11]. But, at least 5 control points are needed to acquire a precise 3D point cloud [12]. Total 8 control points were used in this study and these control points were marked such that it covered all the concerned area. Image acquisition process was started after setting the control points. Also, it is possible to get a scaled final product without placing manual markers if the images from a GPS enabled camera is used for processing. However, the manual marker placement with pre-defined coordinate system is usually more accurate allowing more precise geo-referencing.

3. Case Study

3.1. Study area

A small reach of a physical hydraulic model of a river, in the hydraulic laboratory of Hydro lab at Kathmandu, was taken as the object for this study. The model was built in 1:40 scale. The length of the study area is about 3 m which represents 120 m long river reach in prototype. Eight ground control points covering all the region was marked on the model as shown in the Figure 1.
3.2. Image acquisition
Several images were taken from varying camera position and angle covering the whole study area. Few close-up images of the boulders were also taken to obtain better detailing. It was ensured that there was enough light in the model with no direct sunlight and sharp shadows. Each feature and control points were captured in at least three images from different viewpoints.

3.3. Image processing and dense reconstruction
At first, the images were filtered removing very bright, dark, blurred and shadowed images if any. 46 images were selected for further processing. The set of images were processed using all the selected software individually and 3D dense point clouds were created. Meshroom and ReCap created 3D mesh as output, which were later converted to 3D dense point cloud using Meshlab. Total time for feature extraction/feature matching, sparse reconstruction and dense reconstruction by each software was recorded and compared as one of the major performance parameters. The processing time up to dense reconstruction for the selected software is shown in Table 1. The processing was done mostly with default settings of the software to assess the possibility of using them by a complete beginner.

| Tools      | Feature extraction and Feature matching | Sparse reconstruction | Dense reconstruction | Total processing time |
|------------|----------------------------------------|-----------------------|---------------------|-----------------------|
| Colmap     | 1 hr 51 min                            | 31 min                | 8 hr 4 min          | 10 hr 26 min          |
| Meshroom   | 2 min                                  | 3 min                 | 1 hr 41 min         | 1 hr 46 min           |
| Photoscan  | 5 min                                  | -                     | 3 hr 34 min         | 3 hr 39 min           |
| ReCap      | -                                      | -                     | -                   | 2 hrs                 |
| Regard3d   | 1 hr 38 min                            | 29 min                | 35 min              | 2 hr 42 min           |
| VisualSfM  | 3 min                                  | 31 sec                | 24 min              | 58 min                |

3.4. Post processing
The 3D dense point clouds created by the selected software were transformed from arbitrary 3D coordinates to the real prototype coordinates (not model coordinates) using rotation, translation and scaling in reference to the coordinates of GCPs. It was done in Meshlab using ‘roto-translation’ with ‘uniform scaling’ in geo-referencing tool. During the transformation, GCPs with error (RMSE) greater than 1 were eliminated to ensure better accuracy [13]. Thus, the georeferenced point cloud is generated. The point cloud data was exported to create a contour map with 1m X 1m grid resolution as shown in Figure 2.
3.5. Manual measurements

For assessing the accuracy of outputs by selected software, 4 cross-sections in the river model were manually measured using a level machine and bar scale. These cross sections were compared with the corresponding cross-section profiles extracted from the contour plots produced by using dense point clouds from the selected software.

![Figure 2. Contour Map plotted from 3D point cloud created by PhotoScan](image)

4. Results and discussion

It was observed that SfM method created high density point cloud thus captured better details compared to the manual measurement. The total number of points in the final 3D dense point clouds vastly varies from 2.03 million for Meshroom to 29.2 million for PhotoScan. Although point cloud by Meshroom contained minimum number of vertices, the quality of the output was not lesser. Hence, bigger number of vertices may not necessarily mean better quality of the output. The final output by each software and the total number of vertices forming the 3D output are shown in Figure 3. Finally, 4 cross-sections profiles within the study area were extracted from the contour map plotted using 3D dense point cloud from each software. Each of the cross-section profiles were plotted together with that obtained from manual measurements and compared with each other as shown in Figure 4 where comparison for two cross-section profiles are shown. The plots show that the cross-section profiles generated with selected SfM tools are in close agreement with each other and with the manually measured cross section profile.

To quantify the capability of the SfM tools to predict the vertical dimension, elevation of the 150 points, where manual measurements were done, were extracted from the point clouds generated by each of the selected software. Those elevations were compared with measured elevations and mean absolute error (MAE), root mean square error (RMSE) and coefficient of determination ($R^2$) were calculated as shown in Table 2. It is to be noted that MAE and RMSE values shown in Table 2 are in meters as the analysed results were in prototype scale. The results show that each of the selected software is good at predicting the vertical dimension with MAE below 0.24 m and RMSE below 0.30 m in prototype scale which represents MAE below 6 mm and RMSE below 7.5 mm in model scale. Here, Photoscan stood out as ‘the best among equals’ by scoring lowest MAE and RMSE, and highest $R^2$ value. However, the quality of results are dependent on various factors like choice of different algorithms and respective parameter values within SfM technique, resolution of DEM generated from dense point cloud and accuracy of GCPs used for geo-referencing. Also the acceptable limit for error or discrepancies varies with the purpose of the model study, scale factor and measurement techniques. Therefore, there is an immense possibility to obtain better results with free and open source software by tweaking various parameters whereas commercial software like Photoscan and ReCap work as a
black box model and give their users lesser control. On the other hand, Photoscan and ReCap are user friendly, easy to use and can deliver quality results with minimum involvement of the user.

![Figure 3. Outputs by (a) COLMAP, (b) Regard3d, (c) PhotoScan, (d) Meshroom, (e) VisualSfM and (f) ReCap and total number of vertices in respective outputs](image)

![Figure 4. Cross section profiles (a) L9-R11 and (b) L9-R10](image)

**Table 2.** Analysis of the error in predicting elevations by the selected software.

| Tools      | MAE   | RMSE  | $R^2$  |
|------------|-------|-------|--------|
| Colmap     | 0.161 m | 0.216 m | 0.9958 |
| Meshroom   | 0.176 m | 0.232 m | 0.9956 |
| Photoscan  | 0.155 m | 0.195 m | 0.9965 |
| ReCap      | 0.165 m | 0.222 m | 0.9962 |
| Regard3d   | 0.232 m | 0.294 m | 0.9946 |
| VisualSfM  | 0.175 m | 0.228 m | 0.9956 |
5. Conclusion
From the results of this study, it can be said that SfM method can be used in hydraulic laboratories to efficiently capture 3D geometry from a physical hydraulic model in shorter time yet with better details and within acceptable accuracy. It is also concluded that the free and open source software are also capable of producing good results as compared to results from commercial software. Moreover, free and open sourced software offers more control to the users and hence have huge potential for researchers to produce even better-quality results. But for commercial purposes, where quality results are required in shorter time with minimum involvement of the user, commercial software are recommended. For example, Photoscan and ReCap can perform all the processes including scaling of the final model output and DEM generation within one platform and hence provide the complete solution. Whereas, free and open sourced software requires additional third-party software to perform different processes e.g. Meshlab for geo-referencing the output model. Lastly, it should be noted that the first criterion to achieve better quality output is to take better quality photographs.

References
[1] Westoby M J, Brasington J, Glasser N F, Hambrey M J and Reynolds J M 2012 “Structure-from-Motion” photogrammetry: A low-cost, effective tool for geoscience applications Geomorphology 179 300–14
[2] Snavely N, Seitz S M and Szeliski R 2008 Modeling the World from Internet Photo Collections Int. J. Comput. Vis. 80 189–210
[3] James M R and Robson S 2012 Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application J. Geophys. Res. Earth Surf. 117
[4] Micheletti N, Chandler J H and Lane S N 2015 Investigating the geomorphological potential of freely available and accessible structure-from-motion photogrammetry using a smartphone Earth Surf. Process. Landf. 40 473–86
[5] Wu C 2013 Towards Linear-Time Incremental Structure from Motion 2013 International Conference on 3D Vision - 3DV 2013 2013 International Conference on 3D Vision - 3DV 2013 pp 127–34
[6] Wu C VisualSFM: A Visual Structure from Motion System http://ccwu.me/vsfm/, 2011
[7] Wu C, Agarwal S, Curless B and Seitz S M 2011 Multicore bundle adjustment CVPR 2011 CVPR 2011 pp 3057–64
[8] Schönberger J L and Frahm J M 2016 Structure-from-Motion Revisited 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) pp 4104–13
[9] Cignoni P, Callieri M, Corsini M, Dellepiane M, Ganovelli F and Ranzuglia G 2008 MeshLab: An Open-Source Mesh Processing Tool (The Eurographics Association)
[10] Micheletti N, Chandler J H and Lane S N 2015 Structure from motion (SFM) photogrammetry
[11] Oniga V-E, Breaban A-I and Statescu F 2018 Determining the Optimum Number of Ground Control Points for Obtaining High Precision Results Based on UAS Images Proceedings 2 352
[12] Goldstein E B, Oliver A R, deVries E, Moore L J and Jass T 2015 Ground control point requirements for structure-from-motion derived topography in low-slope coastal environments (PeerJ PrePrints)
[13] Alfredsen K, Haas C, Tuhtan J A and Zinke P 2018 Brief Communication: Mapping river ice using drones and structure from motion The Cryosphere 12 627–33