3D SURFACE RECONSTRUCTION FROM MULTI-DATE SATELLITE IMAGES

Sebastian Bullinger, Christoph Bodensteiner, Michael Arens
Fraunhofer IOSB, Ettlingen, Germany
{sebastian.bullinger, christoph.bodensteiner, michael.aren}@iosb.fraunhofer.de

KEY WORDS: Satellite Imagery, 3D Reconstruction, Photogrammetry, Surface Reconstruction, Texturing.

ABSTRACT:
The reconstruction of accurate three-dimensional environment models is one of the most fundamental goals in the field of photogrammetry. Since satellite images provide suitable properties for obtaining large-scale environment reconstructions, there exist a variety of Stereo Matching based methods to reconstruct point clouds for satellite image pairs. Recently, the first Structure from Motion (SfM) based approach has been proposed, which allows to reconstruct point clouds from multiple satellite images. In this work, we propose an extension of this SfM based pipeline that allows us to reconstruct not only point clouds but watertight meshes including texture information. We provide a detailed description of several steps that are mandatory to exploit state-of-the-art mesh reconstruction algorithms in the context of satellite imagery. This includes a decomposition of finite projective camera calibration matrices, a skew correction of corresponding depth maps and input images as well as the recovery of real-world depth maps from reparameterized depth values. The paper presents an extensive quantitative evaluation on multi-date satellite images demonstrating that the proposed pipeline combined with current meshing algorithms outperforms state-of-the-art point cloud reconstruction algorithms in terms of completeness and median error. We make the source code of our pipeline publicly available.

1. INTRODUCTION

1.1 3D Surface Reconstruction with Satellite Imagery

The creation of large-scale environment reconstructions is relevant for several application areas such as urban planning or environmental monitoring. While there is a variety of modalities and sensors to compute such models, the usage of observation satellites has recently gained significant attraction because of the progress made in the field of image-based reconstruction. In contrast to pictures of ground-level and airborne devices, satellite images cover huge areas, which allow for reconstructing large-scale environment models with less effort. In contrast to active sensors like Lidar or Radar, image data inherently represents appearance information, which allows image-based reconstruction pipelines to derive not only geometry, but also the corresponding texture information.

Stereo Matching and Structure from Motion (SfM) are currently the predominant methods to derive geometric structures from satellite images (de Franchis et al. [2014] Zhang et al. [2019]). While SfM reconstructs information of multiple images, Stereo Matching is restricted to single image pairs. Thus, SfM based approaches are inherently better suited to process large (unstructured) image sets such as multi-date satellite imagery. Since SfM obtains intrinsic and extrinsic camera parameters during reconstruction, the corresponding results allow to utilize dense and surface reconstruction methods to compute dense point clouds, to triangulate corresponding meshes and to perform mesh texturing. In comparison to point clouds, textured meshes oftentimes create a similar appearance with a lower information redundancy and inherently contain specific properties that simplify the computation of consistent visualizations or intersection tests at different scales.

1.2 Paper Overview and Contribution

We propose a pipeline that allows to compute textured meshes from multi-date satellite imagery. To apply available surface reconstruction libraries to satellite images, several preprocessing steps are required. This leads to the following main contributions:

(1) We leverage PAN-sharpening (Pohl and Genderen, 1998) to combine information of image data captured by different sensors, i.e. we compute high-resolution satellite images with color information exploiting the high resolution reconstruction of panchromatic images and the color information of low-resolution multispectral images - see Section 2. Since satellites cover huge areas the corresponding reconstructions can easily exceed available memory capacities of modern workstations. Thus, we extract and align only a specific area of interest of the panchromatic and multispectral imagery using the metadata of the respective satellite image.

(2) The projection properties of satellite imaging sensors are typically described by Rational Polynomial Camera (RPC) models (Tao and Hu, 2001). Since the reconstruction operates only on a specific (continuous) part of each satellite image, the corresponding (complex) RPC model can be substituted with a Finite Projective Camera (FPC) model that approximates the RPC model locally. Since current state-of-the-art dense and surface reconstruction libraries do not support the skew parameter present in finite FPC models, we perform a specific adaption of the reconstructed FPC models that allows us to maintain the consistency of the reconstruction results by performing a skew correction of the corresponding images and depth maps. A mathematical derivation that justifies the selected transformation is provided in Section 3.

(3) To improve the numerical stability of the depth map computation, we follow the MVS step proposed by Zhang et al. (2019). Thus, the converted (and undistorted) depth maps are defined w.r.t. to a reparameterized space. Since Zhang et al. (2019) do not provide information on how to obtain the actual depth values, we provide a concise derivation in Section 4 that demonstrates how the inverse projection matrix allows to recover the actual depth values.

(4) Following the steps above we are able to utilize current
In the following, we describe the proposed pipeline for reconstructing textured surface models of multi-date satellite imagery. The dependencies of the different pipeline components and their relationships are depicted in Fig. 1. Processing multi-date images of observation satellites with current state-of-the-art image-based reconstruction libraries easily exceeds memory capacities of modern workstations. To control memory requirements and processing times our pipeline determines an area of interest (AOI) in a given set of panchromatic and multispectral image pairs using a single bounding box defined in UTM coordinates. The panchromatic band is agnostic to color information, but shows a lower ground sample distance than the color bands captured by the multispectral sensor.

Since many state-of-the-art feature descriptors such as Root-SIFT (Arandjelovic, 2012) exploit image gradients instead of plain color information, we utilize panchromatic images to reconstruct the scene geometry, i.e., to perform SIM, MVS and meshing computations.

The SIM component exploits the approach by Zhang et al. (2019) to locally approximate the RPC model (Tao and Hu, 2001) of each panchromatic image with an FPC model describing only the projection properties of the area of interest. The FPC model has 11 degrees of freedom - the corresponding calibration parameters are estimated using PnP (Zhang et al., 2019) + Zhang et al. (2019) + Ummenhofer and Brox (2017). The SIM component exploits the approach by Zhang et al. (2019) + Zhang et al. (2019) + Schönhuber et al. (2016) + Goesele et al. (2007) + Zhang et al. (2019) + Schönhuber et al. (2016) to locally approximate the RPC model (Tao and Hu, 2001) of each panchromatic image with an FPC model describing only the projection properties of the area of interest. The FPC model has 11 degrees of freedom - the corresponding calibration parameters are estimated using PnP (Zhang et al., 2019) + Zhang et al. (2019) + Ummenhofer and Brox (2017).

In the following, we describe the proposed pipeline for reconstructing textured surface models of multi-date satellite imagery. The dependencies of the different pipeline components and their relationships are depicted in Fig. 1. Processing multi-date images of observation satellites with current state-of-the-art image-based reconstruction libraries easily exceeds memory capacities of modern workstations. To control memory requirements and processing times our pipeline determines an area of interest (AOI) in a given set of panchromatic and multispectral image pairs using a single bounding box defined in UTM coordinates. The panchromatic band is agnostic to color information, but shows a lower ground sample distance than the color bands captured by the multispectral sensor.

Since many state-of-the-art feature descriptors such as Root-SIFT (Arandjelovic, 2012) exploit image gradients instead of plain color information, we utilize panchromatic images to reconstruct the scene geometry, i.e., to perform SIM, MVS and meshing computations. The SIM component exploits the approach by Zhang et al. (2019) to locally approximate the RPC model (Tao and Hu, 2001) of each panchromatic image with an FPC model describing only the projection properties of the area of interest. The FPC model has 11 degrees of freedom - the corresponding cal-

1 Source code is available at https://github.com/hidden/for_review
Panchromatic Images  

Area of Interest (AOI)  

Multispectral Images  

AOI Extraction, Tonemapping  

SfM  

PAN-Sharpening  

Skew-correction of Cameras, Images and Depth Maps  

MVS  

Skew-correction of Images  

Meshing  

Texturing  

Figure 1. Pipeline overview including the different components, the corresponding results and their computational dependencies. As input the pipeline receives a set of panchromatic and multispectral images as well as a bounding box defining the area of interest.

The calibration matrix is shown in (1). In order to represent the obtained SfM reconstruction with camera models containing no skew parameters (i.e., 10 degrees of freedom), we apply an intrinsic parameter decomposition of the registered cameras as well as a skew correction of the corresponding input images and depth maps. Further details are provided in Section 4.1 and Section 4.2. This representation allows us to utilize current state-of-the-art dense and surface reconstruction libraries to compute dense point clouds and non-textured meshes.

In order to texture the reconstructed surfaces, we leverage PAN-Sharpening (Pohl and Genderen, 1998) to compute high-resolution satellite images with color information exploiting the high resolution resolution of panchromatic images and the color information of low-resolution multispectral images. Analogously to the panchromatic images, we apply the skew correction step to the PAN-Sharpened images to utilize current texturing libraries.

We considered several state-of-the-art SfM pipelines to reconstruct satellite images as shown in Table 1. The evaluation conducted by Knäpsch et al. (2017) and Stathopoulou et al. (2019) suggest that Colmap (Schönberger and Frahm, 2016) is currently one of the top performing SfM pipelines. Conveniently, there exists already VisSat (Zhang et al., 2019), a satellite specific adaption of Colmap, that allows us to perform a sparse reconstruction of multi-date satellite images.

We considered several methods (Kazhdan and Hoppe, 2013; Schönberger et al., 2016; Fuhrmann and Goesele, 2014; Ummenhofer and Brox, 2017; Jancosek and Pajdla, 2011) to reconstruct a watertight surface of the environment. Our quantitative evaluation in Section 6 suggests that the algorithm by Ummenhofer and Brox (2017) is a reasonable candidate to perform the surface reconstruction in the context of satellite images.

As shown in Table 1, our pipelines use the method by Waechter et al. (2014) in the texturing step.

3. SATELLITE IMAGE PREPROCESSING

3.1 Area of Interest Extraction

As previously mentioned, earth observation satellites typically use panchromatic and multispectral sensors for image acquisition to capture different resolutions and wavelength spectra. The panchromatic band is agnostic to color information, but shows a lower ground sample distance than the color bands captured by the multispectral sensor. We use the provided metadata of the corresponding multispectral image to extract the blue,
green and red color band. Further, we filter all panchromatic and multispectral images with high cloud coverage (i.e., > 50%) to reduce the processing time using the corresponding metadata. To compute the area of interest for each panchromatic or multispectral image, we use the corresponding RPC model to project a geo-registered rectangle (provided by the user) to pixel locations.

3.2 Tone Mapping

Since the satellite images are defined as high dynamic range images with long-tailed intensity distributions, a naive scaling results in images with low contrast. To tackle this issue we perform the tone mapping procedure proposed by Zhang et al. (2019) for panchromatic images that includes a gamma correction and a filtering of intensity values below the 0.5 percentile and above the 99.5 percentile. The gamma correction is defined by \( I_p = I_p^{1/2.2} \), where \( I_p \) denotes the gamma corrected image and \( I \) the original satellite image. To extend this approach to multispectral images, we apply the percentile filtering for each channel individually. This is necessary to avoid that a single color band dominates the filtering process, which potentially results in unintended color shifts.

3.3 PAN-Sharpening

Since the RPC models are geo-registered with a set of image specific ground control points, we are able to project the area of interest to each panchromatic and multispectral image pair. By extracting the resulting sub-parts, we obtain an aligned image pair that represents the same area. Because of the alignment we are able to exploit PAN-Sharpening (Pohl and Gendener 1998) to leverage the high resolution of the panchromatic image and the color information of the multispectral image to compute a high-resolution image including color information.

Recently published PAN-Sharpening approaches (Fu et al. 2019; Deng et al., 2019) evaluate the algorithms only on images with low resolution. By conducting several experiments we observed that processing high-resolution imagery with these methods is infeasible because of their high computational requirements. For instance, the implementation by Fu et al. (2019) required on our workstation more than 30s to process a \( 100 \times 100 \) multispectral and a \( 400 \times 400 \) panchromatic image. Instead we use a weighted variant (GDAL/OGR contributors 2020) of the widely used Brovey transform (Pohl and Gendener 1998). While the algorithm produces overall good results, it occasionally shows artifacts for individual pixels or small pixel regions - caused by strong reflecting surfaces such as vehicles. Since our pipeline computes the final mesh textures with the algorithm proposed by Waechter et al. (2014), a photo-consistency check ensures that such artifacts do usually not hamper the final appearance.

4. SUBSTITUTION OF RECONSTRUCTION RESULTS

We leverage the approach by Zhang et al. (2019) in order to locally approximate each pushbroom camera model with a finite projective camera (FPC) model (Hartley and Zisserman 2004). Current state-of-the-art libraries for surface reconstruction and texturing such as MVE (Fuhrmann et al., 2014) or OpenMVS (Cernia 2021) do not support the skew parameter contained in finite projective camera models. We address this issue by substituting the reconstructed FPC models with simpler calibration matrices containing no skew factors. To maintain the consistency of the reconstruction we perform a skew correction of the corresponding images and depth maps.

In the following, we derive a decomposition of the FPC calibration matrix that justifies the selected camera model substitution and the conducted image and depth map transformation.

4.1 Decomposition of Finite Projective Camera Models

The calibration matrix of the FPC model is shown in (1).

\[
K_p = \begin{bmatrix} f_x & s & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix}
\]

where \( f_x \) and \( f_y \) represent the focal length of the camera in terms of pixel dimensions in the \( x \) and \( y \) direction. Similarly, \( p_x \) and \( p_y \) define the principal point with respect to the two pixel dimensions. The value \( s \) reflects the skew of the camera. If \( s = 0 \) this means that the pixels are skew such that the \( x \) and \( y \) axes are not perpendicular.

We decompose the calibration matrix \( K_p \) of the FPC into a simplified calibration matrix \( K_{p'} \) and an affine transformation \( T_{p' \rightarrow p} \) with \( K_p = T_{p' \rightarrow p}K_{p'} \). In general, the transformation \( T_{p' \rightarrow p} \) is given by (2).

\[
T_{p' \rightarrow p} = \begin{bmatrix} f_x & s & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{T_x} & 0 & -\frac{p_x}{T_x} \\ 0 & \frac{1}{T_y} & -\frac{p_y}{T_y} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{f_x}{T_x} & -\frac{s}{T_x} \frac{p_x}{T_x} + \frac{p_y}{T_y} \\ 0 & \frac{f_y}{T_y} & -\frac{p_y}{T_y} \frac{p_y}{T_y} + p_y \\ 0 & 0 & 1 \end{bmatrix}
\]

The projection of a point in camera coordinates \( p_c \) onto an homogeneous image point \( \tilde{p}_t \) may now be reformulated according to (3).

\[
\tilde{p}_t = K_{p}p_c \iff \tilde{p}_t = T_{p' \rightarrow p}K_{p'}p_c \iff T_{p' \rightarrow p}\tilde{p}_t = K_{p'}p_c.
\]

Equation (3) shows that the original calibration matrix \( K_p \) can be substituted with an arbitrary finite projective camera model \( K_{p'} \) by transforming the image projections with \( T_{p' \rightarrow p} \).

4.2 Camera Substitution and Skew Correction

Many state-of-the-art MVS, surface reconstruction and texturing libraries including the (official) implementations used in our evaluation (see Section 5) support only skew-free camera models \( K_s \), i.e., \( s' \) must be 0. The choice for \( f_x, f_y, p_x \) and \( p_y \) is not restricted. By defining \( (f_x', f_y', p_x', p_y', s') \) according to \( (f_x, f_y, p_x, p_y, s) \) according to \( (f_x', f_y', p_x', p_y', s') := (f_x, f_y, p_x, p_y, 0) \) we obtain the transformation \( T_{x \rightarrow p} \) shown in (4).

\[
K_p = \begin{bmatrix} f_x & s & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{f_x'} & 0 & -\frac{p_x'}{f_x'} \\ 0 & \frac{1}{f_y'} & -\frac{p_y'}{f_y'} \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & \frac{s}{f_x'} & \frac{-s p_x'}{f_x'} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]

The matrix \( T_{x \rightarrow p} \) in (4) contains not only a skew correction factor (i.e., \( \frac{s}{f_x'} \)), but also a translation component (i.e., \( \frac{-s p_x'}{f_x'} \)).
Thus, a part of the pixels will be translated outside of the image boundaries and important information will be lost. Instead, we define $K$, according to \( f_x, f_y, p_x, p_y, s \) := \( f_x, f_y, p_x - \frac{s p_x}{Z}, p_y, 0 \) to obtain a translation-free skew-correcting matrix $T_{s\rightarrow p}$ as shown in (5).

\[
\begin{bmatrix}
    f_x & s & p_x \\
    f_y & p_y \\
    0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    \frac{1}{f_x} & 0 & \frac{-p_x}{f_x} + \frac{s p_y}{f_x} \\
    \frac{1}{f_y} & 0 & \frac{-p_y}{f_y} \\
    0 & 0 & 1
\end{bmatrix}
= \begin{bmatrix}
    1 & \frac{1}{v} & 0 \\
    0 & 1 & 0 \\
    0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    x' \\
    y' \\
    1
\end{bmatrix}
\]

(5)

For each registered image $I_p$ (corresponding to $K_p$) we compute a skew-free image $I_s$ and a skew-free depth map $D_s$ (corresponding to $K_s$) using $I_s = T_{s\rightarrow p}I_p$ and $D_s = T_{s\rightarrow p}D_p$. While removing the skew factors we applied a cubic interpolation to compensate artifacts caused by the discrete nature of pixel based image representations. See Fig. 3 for a comparison of the original and the corresponding skew-free image.

5. REAL DEPTH MAP RECOVERY

Because the distances of the registered cameras and the reconstructed point cloud (in the context of satellite images) are extremely large, the values in the corresponding depth maps are compressed to an interval that is far away from the origin. To improve the numerical stability of the depth map reconstruction, it is reasonable to perform the computations in a reparameterized space. We follow the reparameterization proposed by Zhang et al. (2019), which uses plane-plus-parallax (Irani et al. 1998).

Since Zhang et al. (2019) do not describe how to obtain the real depth map values, we provide a concise explanation in the following. The reparameterization is defined according to (6).

\[
\begin{bmatrix}
    P_{11} & P_{12} & P_{13} & P_{14} \\
    P_{21} & P_{22} & P_{23} & P_{24} \\
    P_{31} & P_{32} & P_{33} & P_{34} \\
    0 & 0 & Z_\star & -Z_d
\end{bmatrix}
\begin{bmatrix}
    x' \\
    y' \\
    z
\end{bmatrix}
\approx
\begin{bmatrix}
    uZ \\
    vZ \\
    Z_\star
\end{bmatrix}
\]

(6)

\[
\begin{bmatrix}
    u & v \\
    v & Z_\star
\end{bmatrix}
\begin{bmatrix}
    Z_\star \\
    Z_d
\end{bmatrix}
= [uZ, vZ, Z, Z_\star/2]T
\]

(7)

To increase numerical stability, Zhang et al. (2019) normalize $P$ and $P^{-1}$. Let $n_{p^{-1}}$ denote the normalization factor of $P^{-1}$. We obtain the real depth values using $n_{p^{-1}} \cdot Z$.

By computing $K_s$, $I_s$ (see Section 3.1) and $D_s$ for each registered image, we are able to utilize established libraries for MVS and texturing in the context of satellite images such as the algorithms proposed by Fuhrmann et al. (2014) and Cernea (2021).

6. EXPERIMENTS AND EVALUATION

For our experiments we use multi-date satellite imagery of the IARPA MVS3DM dataset (Bosch et al. 2016). The ground sample distance of the panchromatic and the multispectral images is approximately 0.3 m and 1.3 m.

6.1 Quantitative Evaluation

To compare the proposed pipeline with previously reported results we utilize the benchmark presented by Zhang et al. (2019), which is based upon the IARPA MVS3DM dataset (Bosch et al. 2016). The corresponding ground truth consists of a set of geo-registered airborne lidar scans covering different urban areas near San Fernando, Argentina. Zhang et al. (2019) follow the evaluation protocol by Bosch et al. (2016) to perform a metric comparison of the MVS and the ground truth point cloud. To this end, both point clouds are mapped to a predefined geo-registered grid representing a height map - with cell sizes of 0.5 x 0.5 m². The value for each grid cell is defined by the maximum height value of all corresponding points. Since the height map comparison is subject to the initial geo-registration, the evaluation protocol refines the alignment before performing the actual evaluation. Bosch et al. (2016) define two metric scores that represent the median error (ME) and the completeness (CP) of the reconstructed height map. Concretely, CP denotes the percentage of cells with an error less than 1 m.

In order to exploit this benchmark for mesh evaluation, we extract a set of points from the mesh surface. Since the reconstructed vertex positions usually show a non-uniform (and texture dependent) distribution, they are only partly suited to represent the full surface geometry. Thus, we utilize poisson disk sampling (Cook 1986) to extract a set of points following a random distribution while ensuring that the distance between any pair of samples is larger than a certain predefined threshold. We follow previous practice (Zhang et al. 2019) and fill small holes with adjacent values.

In order to determine a state-of-the-art surface reconstruction approach for multi-date satellite imagery, we quantitatively evaluate the proposed pipeline with several modern meshing algorithms. Table 2 reports the corresponding ME and CP metrics based on surface points extracted with poisson disk sampling. Supplementary, Table 3 provides an analysis using only the reconstructed mesh vertices as surface representation. This allows us to compare the reconstructions with previous results.
Table 2. Quantitative evaluation of the proposed pipeline with different state-of-the-art surface reconstruction algorithms. The surface points used to compute the metrics are extracted with poisson disk sampling. The evaluated surface reconstruction methods include Poisson (Kazhdan and Hoppe, 2013), Colmap (Schönberger et al., 2016; Labatut et al., 2009), OpenMVS (Cernea, 2021; Jancosek and Pajdla 2014), FSSR (Fuhrmann and Goesele, 2014) and GDMR (Ummenhofer and Brox, 2017). FSSR and GDMR require scale information for each point in the fused point cloud to triangulate the corresponding mesh. In this case we substitute the depth map fusion approach of VisSat (Zhang et al., 2019) with the equivalent processing step of MVE (Goesele et al., 2007). The metrics completeness (CP) ↑ and median error (ME) ↓ follow the definition by Bosch et al. (2010). The top three results are highlighted in green (first), blue (second) and red (third).

| Pipeline    | Depth Map    | Meshing Algorithm | Point Sampling | Site 1 | Site 2 | Site 3 |
|-------------|--------------|-------------------|---------------|--------|--------|--------|
| Proposed    | VisSat       | Poisson           | Poisson Disk  | 72.1   | 0.307  | 66.5   |
| Proposed    | VisSat       | Colmap            | Poisson Disk  | 73.7   | 0.305  | 67.9   |
| Proposed    | VisSat       | OpenMVS           | Poisson Disk  | 73.8   | 0.317  | 67.9   |
| Proposed    | MVE          | FSSR              | Poisson Disk  | 71.1   | 0.329  | 67.0   |
| Proposed    | MVE          | GDMR              | Poisson Disk  | 73.4   | 0.307  | 68.0   |

Table 3. Quantitative evaluation of the proposed pipeline with different state-of-the-art surface reconstruction algorithms. The reconstructed mesh vertices are used as surface representation to compute the metrics. See Table 2 for a description of the different algorithms and metrics. As baseline for our comparison we use the vertices of the fused point clouds computed with VisSat (Zhang et al., 2019) and MVE (Goesele et al., 2007). The top three results are highlighted with green (first), blue (second) and red (third).

| Pipeline    | Depth Map    | Meshing Algorithm | Point Sampling | Site 1 | Site 2 | Site 3 |
|-------------|--------------|-------------------|---------------|--------|--------|--------|
| Zhang et al. 2019 | VisSat       | -                 | Vertex        | 71.7   | 0.340  | 66.8   |
| Proposed    | MVE          | -                 | Vertex        | 63.3   | 0.699  | 55.9   |
| Proposed    | VisSat       | Poisson           | Vertex        | 73.4   | 0.305  | 66.7   |
| Proposed    | VisSat       | Colmap            | Vertex        | 73.3   | 0.284  | 67.7   |
| Proposed    | VisSat       | OpenMVS           | Vertex        | 71.9   | 0.307  | 67.3   |
| Proposed    | MVE          | FSSR              | Vertex        | 72.3   | 0.313  | 66.8   |
| Proposed    | MVE          | GDMR              | Vertex        | 73.0   | 0.319  | 68.0   |

Table 4. Parameter configuration of the meshing algorithms used for the quantitative evaluations.

| Meshing Algorithm | Modified Parameter | Value (Default Value) |
|-------------------|--------------------|-----------------------|
| Poisson           | max_proj_dist      | 0 (20)                |
| OpenMVS           | min_point_distance | 0 (2.5)               |
| FSSR              | smoothing_iterations | 0 (2)               |
| GDMR              |                    | -                     |

Table 5. Number of vertices reconstructed with the different point fusion and mesh triangulation algorithms for each site.

| Depth Map Fusion | Meshing Algorithm | Site 1 | Site 2 | Site 3 |
|------------------|-------------------|--------|--------|--------|
| VisSat           | -                 | 6,289k | 1,880k | 1,780k |
| MVE              | -                 | 199,264k | 53,872k | 55,618k |
| VisSat           | Poisson           | 11,023k | 4,549k | 4,228k |
| VisSat           | Colmap            | 4,861k  | 1,424k | 1,381k |
| VisSat           | OpenMVS           | 3,964k  | 1,210k | 1,078k |
| MVE              | FSSR              | 3,850k  | 773k   | 825k   |
| MVE              | GDMR              | 3,353k  | 859k   | 1,675k |

6.2 Qualitative Evaluation

Fig. 2 shows several reconstructed meshes using the proposed approach with GDMR for meshing, which represents the best configuration of our pipeline. The different reconstructions represent the sites used for the quantitative evaluation. Fig. 3 visualizes the plain geometry of different meshes obtained with Poisson, Colmap and GDMR. We observe that the reconstructed surfaces exhibit different characteristics (e.g. varying levels presented by Zhang et al., 2019). If viable, we utilize the default mesh reconstruction parameters whenever possible - an overview of (non-)deviating parameters is presented in Table 4.
Figure 2. Qualitative evaluation of the proposed pipeline using GDMR (Ummenhofer and Brox, 2017) for three sites contained in the multi-date satellite image dataset provided by Bosch et al. (2016). of completeness, smoothness or mesh granularity), which are consistent to the results reported in Section 6.1.

7. CONCLUSION

This paper presents an approach to reconstruct textured water-tight meshes for multi-date satellite imagery including a detailed description of different pipeline steps and their dependencies. In contrast to previous work, our method leverages current dense and surface reconstruction algorithms to obtain a representation of the scene. We present a decomposition of FPC calibration matrices that allows us to substitute the camera models in the SfM reconstruction with simpler camera models containing no skew factors. To ensure consistent reconstruction results we apply corresponding skew corrections to input images and depth maps. These adjustments enable us to utilize state-of-the-art mesh reconstruction algorithms in the context of satellite images. The paper presents an extensive quantitative evaluation of the pipeline on multi-date satellite images. By sampling points from the reconstructed mesh surfaces we compare our surface reconstructions with point clouds obtained with previously proposed SfM and MVS methods. Our evaluation shows that current state-of-the-art meshing algorithms outperform previous results in terms of completeness and median error. To foster future research we make the source code of the proposed pipeline publicly available.

References

AliceVision, 2018. Meshroom: A 3D reconstruction software. https://github.com/alicevision/meshroom [Accessed January 2021].

Arandjelovic, R., 2012. Three things everyone should know to improve object retrieval. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2911–2918.

Barnes, C., Shechtman, E., Finkelstein, A., Goldman, D. B., 2009. PatchMatch: A Randomized Correspondence Algorithm for Structural Image Editing. ACM Transactions on Graphics, 28(3).
