https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

AI-Based Tracking of Fast-Moving Alpine Landforms Using High Frequency Monoscopic Time-Lapse Imagery

Hanne Hendrickx[1,2], Xabier Blanch[1,3], Melanie Elias[1], Reynald Delaloye[2], Anette Eltner[1]

[1]Institute of Photogrammetry and Remote Sensing, TUD Dresden University of Technology, 01062 Dresden, Germany
[2]Department of Geosciences, University of Fribourg, Fribourg, 1700, Switzerland
[3]Department of Civil and Environmental Engineering, Universitat Politècnica de Catalunya, 08034 Barcelona, Spain

Correspondence to: Hanne Hendrickx (hanne.hendrickx@tu-dresden.de)

Abstract: Active rock glaciers and landslides are critical indicators of permafrost dynamics in high mountain environments, reflecting the thermal state of permafrost and responding sensitively to climate change. Traditional monitoring methods, such as Global Navigation Satellite System (GNSS) measurements and permanent installations, face challenges in measuring the rapid movements of these landforms due to environmental constraints and limited spatial coverage. Remote sensing techniques offer improved spatial resolution but often lack the necessary temporal resolution to capture sub-seasonal variations. In this study, we introduce a novel approach utilising monoscopic time-lapse imagery and Artificial Intelligence (AI) for high-temporal-resolution velocity estimation, applied to two subsets of time-lapse datasets capturing a fast-moving landslide and rock glacier at the Grabengufer site (Swiss Alps). Specifically, we employed the Persistent Independent Particle tracking (PIPs++) model for tracking and the AI-based LightGlue matching algorithm to transfer 2D image data into 3D object space and further into 4D velocity data. This methodology was validated against GNSS surveys, demonstrating its capability to provide spatially and temporally detailed velocity information. Our findings highlight the potential of image-driven methodologies to enhance the understanding of dynamic landform processes, revealing spatio-temporal patterns previously unattainable with conventional monitoring techniques. By leveraging existing time-lapse data, our method offers a cost-effective solution for monitoring various geohazards, from rock glaciers to landslides, with implications for enhancing alpine safety and informing climate change impacts on permafrost dynamics. This study marks the pioneering application of AI-based methodologies in environmental monitoring using time-lapse image data, promising advancements in both research and practical applications within geomorphic studies.

1 Introduction

Active rock glaciers are creeping permafrost features (Barsch, 1996; Kääb and Reichmuth, 2005), serving as crucial indicators of permafrost distribution in high mountain environments (Marcer et al., 2017; RGIK, 2023). Their velocity results from various parameters such as topographic conditions, rock glacier material, and internal structure, and it reflects the thermal state of the permafrost (Cicoira et al., 2021; Delaloye et al., 2010). Higher flow speeds typically occur near the permafrost limits, where mean annual air temperatures approach 0°C (Frauenfelder et al., 2003). They efficiently transport sediment (Delaloye et al., 2010; Kummert and Delaloye, 2018), and the acceleration of these processes becomes more pronounced as rock glacier

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

creep rates increase in a warming climate (Delaloye et al., 2013; Pellet et al., 2023). The same is true for permafrost-related creep features such as landslides (Delaloye and Staub, 2016). This can pose significant geohazards when direct connections to downslope infrastructure exist. Precise monitoring of these fast-moving high-alpine landforms is thus essential for future alpine safety (Hermle et al., 2022), as it provides information about the impact of climate change on creeping mountain permafrost and its thermal state (RGIK, 2023).

The monitoring of fast creeping rock glaciers (> 3 my^{-1}, Marcer et al. 2021) or landslides is particularly challenging. Traditional techniques that require frequent field access, such as differential Global Navigation Satellite System (dGNSS) measurements, face environmental and logistical obstacles. Permanent GNSS installations can offer displacement observations with millimetre accuracy at a continuous temporal resolution, but they may not have the desired longevity on fast-moving landforms due to extreme cases of block sliding, rotation, and rockfall, necessitating re-levelling or instrument replacement (Cicoira et al., 2022). Both dGNSS and permanent installations only measure discrete points, resulting in a limited spatial distribution. However, spatial heterogeneity of landform movement can be expected depending on internal rock glacier structure and terrain characteristics (RGIK, 2023). Improved spatial coverage can be achieved using remote sensing data, such as 3D point clouds derived from Uncrewed Aerial Vehicles (UAV) and Terrestrial or Airborne Laser Scanning (TLS and ALS). These techniques minimize the need for extensive field access, enabling operators to avoid in-person exposure to the fastest-moving areas while still capturing detailed data from these regions. However, these methods often lack the temporal resolution necessary to capture sub-seasonal variations of the landform to its environmental drivers, essential to increase process understanding. Time-lapse imagery or webcam data have the capability to capture the kinematics of alpine landforms with, e.g., hourly resolution. Fixed photogrammetric or stereoscopic camera systems, which are increasingly implemented, involve multiple time-lapse cameras to reconstruct high-resolution 3D point clouds similar to those from UAVs or TLS, and at a fraction of the cost (Blanch et al., 2023; Eltner et al., 2017; Ioli et al., 2024). Nonetheless, deploying multiple cameras in dynamic alpine environments can be challenging, often requiring wide baselines that complicate point cloud generation (Ioli et al., 2024). Recent studies have demonstrated that metric measurements can be obtained from 2D images alone when a 3D model is available, reducing the need for extensive camera arrays (Altmann et al., 2020; Elias et al., 2023; Wegner et al., 2023). Furthermore, monoscopic camera data, which is more readily available than stereo images, often spans a decade or more in the European Alps (Kummert et al., 2018). Generally installed in stable terrain, these cameras tend to have greater longevity than permanent GNSS installations.

Measuring the velocity of boulders within the camera's field of view is possible by establishing correspondences between pairs of points in an image pair, a common challenge in computer vision using optical flow-based methods and feature matching (Fortun et al., 2015). Feature tracking is essentially an optimization problem, where the location of highest similarity between a reference template and a template in the destination image is considered a match (e.g., Eltner et al., 2022). Common similarity measures include Normalized Cross-Correlation (NCC) (Heid and Kääb, 2012) and least square matching (Schwalbe and Maas, 2017). Traditional motion estimation methods face challenges with large displacement, strong illumination changes, and occlusions. Image pre-selection is crucial to minimize shadowing variation (e.g., selecting images from the same time of

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

day), thereby limiting temporal resolution (How et al., 2020). Additionally, high landform velocities can cause decorrelation and noise when using typical area-based matching algorithms, such as NCC (e.g., CIAS by Kääb and Vollmer, 2000), for tracking features between consecutive surveys.

In the era of Artificial Intelligence (AI), traditional feature tracking can leverage the power of Deep Learning. Convolutional Neural Networks (CNNs), widely employed as feature extractors, replace handcrafted features or the use of image intensities or gradients (Hur and Roth, 2020). This approach offers the advantage of representing each pixel with a high-dimensional feature vector, blending distinctiveness and invariance to, for example, appearance changes, thereby enhancing feature robustness over time. This enhances image matching algorithms, as seen with models like SuperGlue (Sarlin et al., 2020), LightGlue (Lindenberger et al., 2023) and LoFTR (Sun et al., 2021).

The goal of this research is to obtain high-quality landform velocities from monoscopic time-lapse cameras, verified through permanent- and dGNSS surveys. To improve motion estimation in mountain time-lapse image sequences, we tested the deep learning model called Persistent Independent Particle tracking (PIPs++) (Harley et al., 2022; Zheng et al., 2023) to track landslide and rock glacier movements and the AI matching algorithm LightGlue (Lindenberger et al., 2023) to convert 2D image measurements to 3D information (Elias et al., 2023). By analysing a small subset of a large time-lapse image dataset, we automatically extracted velocity information as a proof-of-concept. We demonstrate the potential of new methodologies for managing big data and improving the spatio-temporal understanding of landform movements. Two fast-moving alpine landforms (a landslide and rock glacier at Grabengufer site, Switzerland) were selected as a pilot study area. The ultimate aim is to derive reliable velocity values at high temporal resolution from low-cost monoscopic time-lapse cameras, which have not been previously used quantitatively. To achieve this, we developed a workflow applicable to other research sites where validation data is scarce or where traditional remote sensing techniques fall short due to high landform displacements. This approach not only enhances analysis capabilities but also enables the use of basic systems, such as monoscopic cameras, for automatic and cost-effective monitoring and quantification.

2. Study area and dataset description

The Grabengufer study area features two fast-moving alpine landforms (Fig. 1a). The upper section (2700 – 2880 m a.s.l.) consists of a landslide moving up to 1.5 meters a year. The intermediate part of this landslide, referred to as the "feeding section," is monitored by a permanently installed webcam (hereinafter named Wcam04, Fig. 1c). This landslide feeds into a rock glacier (2400 – 2600 m a.s.l.), which moves at a rate of 0.25 to 0.69 meters per day as measured in situ in 2023 (Wcam05, Fig. 1b). The rock glacier eventually flows into a gully prone to debris flows, which is observed by a separate webcam not included in this study. The site is extensively monitored, with bi-annual dGNSS surveys, several permanent GNSS installation from PermaSense (Cicoira et al., 2022) and the Canton of Valais, ten fixed reflectors for repeated theodolite measurements, three time-lapse cameras since 2010/2013 (Wcam02, 04, and 05), and three ground surface temperature loggers.

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

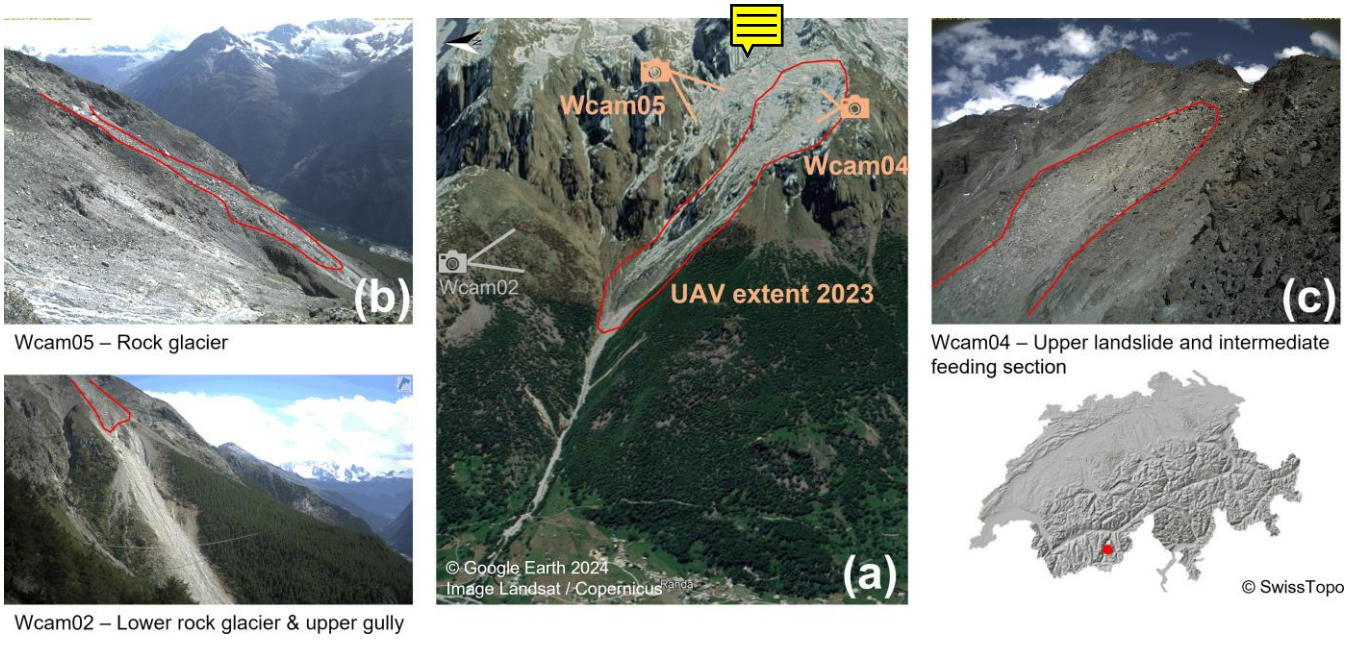


Figure 1: a) Overview of the Grabengufer study area in the Mattertal, Swiss Alps, showing the locations of the webcams used in this study, b) Example image from webcam Wcam05, capturing the rock glacier, and c) example image from webcam Wcam04, focused on the intermediate section of the landslide feeding into the rock glacier—referred to as the feeding section in this study.

The webcam characteristics are detailed in Table 1. The data used in this study was acquired by two Mobotix (M12, M15) webcams, each powered by a solar panel. The webcams capture images at hourly intervals during daylight conditions with a resolution of 3.15 megapixels (2048 x 1536 pixels). The images are sent via a GSM (Global System for Mobile Communication) internet connection to a server for storage and accessibility. For scaling the acquired images, we make use of a high-resolution point cloud (4.89 cm/px, 419 points/m^2), derived from an UAV flight performed on August 3, 2023 with a DJI Phantom 4 RTK at a flight height of 70 m above the terrain and subsequently processed in Agisoft Metashape. Four ground control points were used as check points, demonstrating a planar accuracy of 1 cm and a height accuracy of 0.3 cm. For the validation of the velocities derived from the time-lapse image sequence, we make use of bi-annual dGNSS surveys using real-time kinematics (RTK) collected at the beginning (30.06) and end (13.10) of the summer of 2022 and monthly theodolite measurements of fixed reflectors. Additionally, a permanent GNSS system installed by the Canton of Valais on a large boulder within the feeding section provides high-temporal validation data for one specific point. This system uses a fixed GNSS station on stable terrain to achieve a more accurate differential position.

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

Table 1: Test datasets used in this study, as a selection of the entire datasets

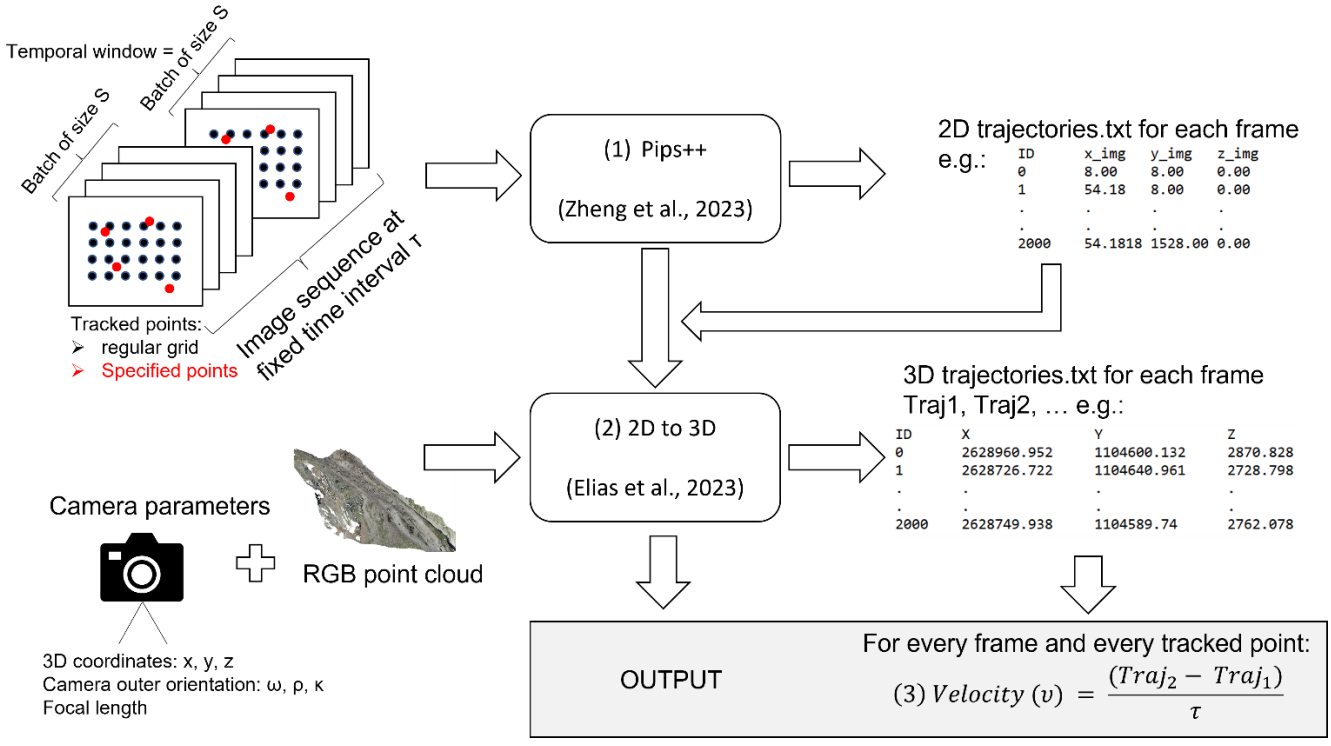
Name - Landform	Location in *CH1903+	Distance to the landform (m)	Date of first installation	Camera properties Focal length (f) Viewing direction	Total number of images (start-2023)	Subsample for this pilot study
Wcam04 – Feeding section	x: 2628628, y: 1104523, z: 2790 m a.s.l.	100	01.09.2013	f: 1.3 mm 100°	48 118	Weekly images for summer 2022
Wcam05 – Rock glacier	x: 2629114, y: 1105319, z: 2870 m a.s.l.	700	17.01.2011	f: 1.3 mm 235°	51 919	Weekly images for summer 2022

*CH1903+ is a commonly used Swiss coordinate system

3. Methods

The overall analysis was conducted step by step with the goal of: 1) getting the translation vectors or trajectories (in pixels) from predefined points to be tracked (regular grid or specified pixel coordinates) from the time-lapse image sequence in a fast and robust way (section 3.1), and 2) scaling the trajectories from pixel coordinates to real-world coordinates to calculate velocities (in m per day) (section 3.2). Both steps apply existing AI frameworks and have limited input requirements, reducing pre-processing time. Step 1 requires time-lapse images taken at fixed intervals (e.g., weekly, daily, or hourly), preferably with a datetime stamp in the filename. No pre-filtering for fog, snow, or blur is needed at this stage, and prior stabilization to account for camera movements is unnecessary, as the camera orientation parameters will be calculated during the registration of image and object data in step 2. This requires a high-resolution 3D point cloud with RGB information of the area of interests and initial estimates of the camera's intrinsic parameters, i.e., focal length, as well as the camera pose, i.e., position and rotation in object space (Fig. 2). Given the short duration of this pilot study, we assume the cameras remain stable over time.

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024



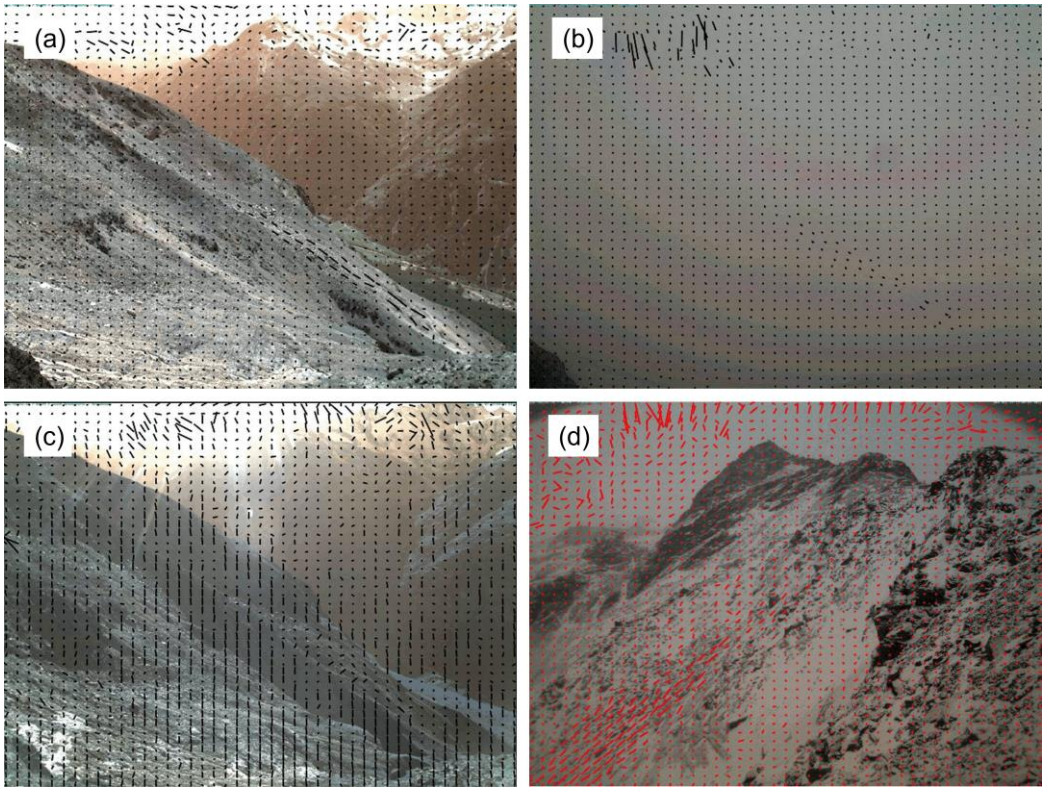
135 **Figure 2: Schematic overview of the workflow presented in this study.**

3.1 Application of the Persistent Independent Particle tracker (PIPs++)

The first step in our workflow (Fig. 2) involves feeding the image sequence, taken at fixed time intervals (e.g., weekly, daily, or hourly), into the Persistent Independent Particle tracker (PIPs++). This AI-based tracker, developed by Harley et al. (2022) and improved by Zheng et al. (2023), operates without retraining on our specific data. PIPs++ operates as a low-level tracker, relying on appearance-matching cues just like traditional methods, and temporal priors. It employs a 2D residual Convolutional Neural Network (CNN) for feature extraction in the initial step (He et al., 2016), generating a feature map for each frame, independent of temporal prior. Following feature extraction, the algorithm feeds into a deep 1D Convolutional Network (ResNet) with fixed-length kernels applied to arbitrary temporal spans (Zheng et al., 2023). Based on the feature maps, PIPs++ calculates local similarity and position of the features in an iterative step using spatial pyramids, similar to RAFT (Teed and Deng, 2020). Just like traditional methods, correlation matrices are calculated to match the feature templates. The main difference from traditional methods is that this approach operates within a multi-frame temporal context, considering all frames in a temporal window or batch of size S to search for the target. After a match is found, the trajectory of the feature within the batch is updated and new correlation pyramids are calculated in an iterative way. If tracking fails for one frame due to occlusion (e.g., fog), the trajectory of the feature can still be estimated as long as it is found in other frames within the batch (as long as

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

the fog does not last longer than the batch size S) (Fig. 3b). Moreover, the similarity is computed using multiple templates per feature. This means that the model does not only rely on the initial appearance of the target in the first frame, but also on its appearance along the trajectory, i.e., throughout the sequence. For each additional template, new correlations are computed and if tracking is successful it is included in the multi-template tracking. This approach allows for tracking the same features
155 despite changes in light conditions, snow cover (Fig. 3d), or self-occlusion (e.g., rolling boulders), by handling appearance changes and considering the trajectory to extract multiple templates per feature. For further information about the model architecture, we refer to Harley et al., (2022) and Zheng et al., (2023). The output of the model is a simple trajectory .txt file per frame (Fig. 2), with the pixel coordinates of each tracked point. This file can be visualized on the respective image for visual inspection (Fig. 3). If a regular grid is used, it is reset each time a new batch starts. When camera movements occur, a
160 homogeneous and mono-directional translation or shift of the regular grid in the stable areas can be observed (Fig. 3c). This shift can be corrected by applying a simple adjustment based on its vertical and horizontal components, similar to the method used by Kenner et al. (2018).



165 **Figure 3: Some examples of velocity vectors as output from the PIPs++ model, a) view of Wcam05, clearly identifying the moving rock glacier body although its suboptimal viewing angle, b) same view occluded by fog but tracking is interfered from the temporal prior, c) same view displaying a systematic shift in the stable areas due to camera movement, d) view of Wcam04 proving PIPs++ good performance in varying snow conditions.**

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

Depending on computational power and image resolution (in our case 3MP), a certain number of points (N) can be tracked within a certain batch size (S). In our case, we used an NVIDIA RTX A6000 GPU, with 48 GB of memory and were able to track 2000 points in a batch size of 19 frames. This means that for a temporal window of 19 frames, the same 2000 particles are tracked before resetting the grid and starting anew. Processing 400 images in this set-up took around 2 minutes. For validation, specified points, such as boulders with available GNSS data, can be tracked in addition to the grid points entered into PIPs++, demonstrating the system's accuracy and reliability. These points are inserted in the script as pixel coordinates.

3.2 From 2D to 3D: image-to-geometry scaling

The second step in the workflow (Fig. 2) is based on the methodology first introduced in Elias et al. (2019) and later updated in Elias et al. (2023). The approach is called image-to-geometry registration and enables the transformation from 2D image data into 3D object space by matching a camera image to a synthetic image rendered from a coloured 3D point cloud preserving location and depth information. This synthetic image should match the image geometry of the real camera as much as possible, which is why an initial guess on the true camera pose, i.e., position and viewing direction, as well as the cameras focal length and pixel size are required for synthetic image rendering. The 2D-3D matches are used to perform space resection to determine the camera pose and, assumed well-distributed key points in view of the camera image, cameras intrinsic parameters and lens distortion (Fig. 4). For matching camera and synthetic image, image-to-geometry registration was adapted using the AI-based matching approach LightGlue (Lindenberger et al., 2023). LightGlue is a high performance, open-source adaptation of SuperGlue (Sarlin et al., 2020) and is excellent for registering images to each other, even if they differ greatly in geometry and radiometry. After determining the camera pose and internal camera geometry in the reference system of the 3D point cloud, image information can be mapped in 3D space by back projection (called “scaling”). Initial camera properties of the used time-lapse cameras (Table 1) are available and we can access the detailed topography of the scene through a single UAV measurement, described above.

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

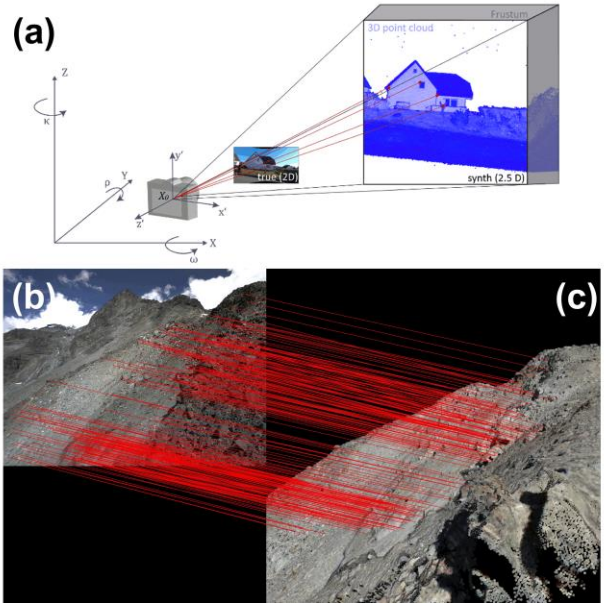


Figure 4: a) Visualisation of the image-to-geometry scaling (Elias et al., 2023), b) real 2D image from Wcam04 being matched to c) the rendered synthetic image derived from the UAV data, displaying a good distribution.

The accuracy of the determined camera calibration and thus the accuracy of the scaled 3D data is directly linked to the accuracy and quality of the 3D data. The point cloud data needs to include colour information to ensure matching to succeed. This information could either be RGB, which is common when originating from photogrammetry, or intensity information that could be used for colouring. Note, the use of 3D point clouds with intensity values from laser scanning in image-to-geometry registration requires objects that can be clearly distinguished from each other in the intensity image by different properties with regard to their reflectance. This is the case, for example, when used in urban areas where building features can be excellently distinguished in the intensity-coloured synthetic image due to different materials and consequent reflections (see Elias et al., 2023). In natural settings, like the Grabengufer site, the intensity-image will not provide enough unique features to be aligned with the camera image which is why we use an RGB-coloured 3D point cloud from UAV photogrammetry for scaling.

Image-to-geometry registration is usually done iteratively where the initial guess of the camera pose and interior camera geometry is gradually improved and used to render the synthetic image being more and more close to the cameras perspective. As a rule of thumb, the more similar the perspectives of camera and synthetic image, the more matches are likely to be found and the better the parameters of the pose and interior camera geometry of the camera image can be determined and used for scaling the image information. In this study, we parametrized image-to-geometry registration in a way, that three iterations are performed. Details on parameters are given in Elias et al. (2019) and Elias et al. (2023).

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

After the results are scaled, absolute distances and velocities are calculated for each tracked point (Fig. 2 (3)), within the chosen time interval, in this case every week for the summer of 2022. A spatial aggregation is performed to split the results for the active moving part (the feeding area or the rock glacier) and the points in the presumed stable area (dotted line in Fig. 5). An accuracy analyses is performed on the points in the stable areas, where the calculated velocities are expected to be close to zero.

4. Results and validation

As a pilot study, we processed a small subset of two time-lapse image datasets (Table 1), of a permafrost-affected landslide and its intermediate section feeding into a rock glacier, with the workflow outlined above. The results for both areas for the entire summer period (06.06 to 10.10.2022) are shown in Figure 5a.

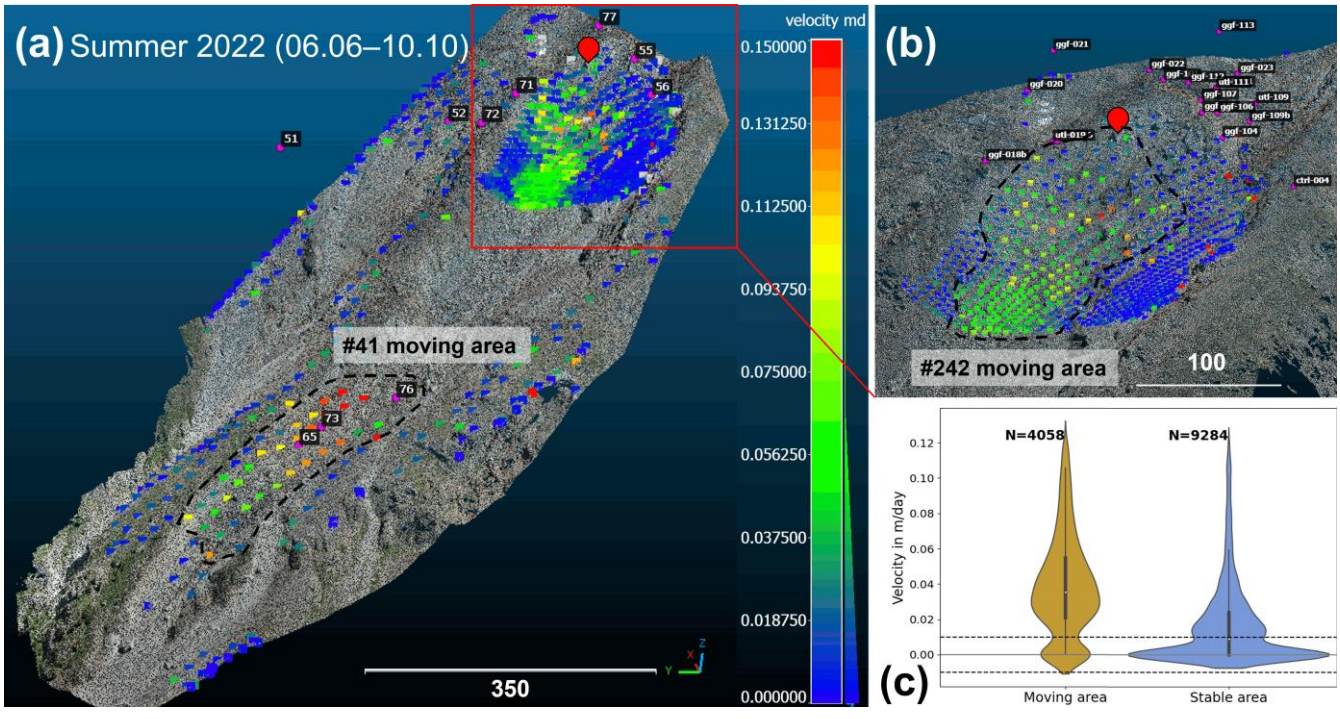


Figure 5: a) Accumulated velocity results (in meters per day) over the entire summer period (127 days) for both the feeding section (upper area, observed by Wcam04) and the rock glacier (lower area, observed by Wcam05), displayed on UAV point cloud data, including theodolite-measured points and the permanent GNSS installation (marked with a red location pin), b) close-up of the intermediate feeding section (Wcam04) including dGNSS measured points, c) plot of the weekly velocity distribution for both the moving and the stable areas (data from Wcam04), showing a clear noise level between -0.01 and 0.01 m/day in both regions. Dotted lines indicate the borders used to differentiate between moving and stable areas in this study.

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

The results are coherent with what we expect from the landform displacements, showing a largely homogenous velocity flow field with faster velocities in the upper area of the rock glacier and in the feeding area of the landslide. Points in blue display stable areas for the considered timeframe of 127 days. Especially in the feeding section, there are single large boulders that move faster than their surroundings. For validation, permanent GNSS, dGNSS, and theodolite data acquired at the beginning and end of the summer season were used. Unfortunately, not all measured points were suitable for validation, as some were outside the camera's field of view and thus could not be included. Additionally, due to logistical and safety constraints, dGNSS points could not be measured in the fast-moving areas of the intermediate feeding section, although they are available in the upper landslide area (Fig. 5b). As a result, only a few velocities derived from theodolite measurements could validate the higher velocities observed in the time-lapse sequence (Fig. 5a). More interestingly, validation was also conducted by examining detected motion in presumed stable areas (Fig. 5c). Here, we observed a clear abundance of points around zero, as expected for stable areas, with noise of about 1 cm/day. When considering weekly images, our method and camera setup can thus detect real movement if it surpasses 7 cm. An average movement of 4 cm per day (28 cm per week) was derived for the feeding section, while in stable areas, detected movement was mostly below 4 cm per day (Fig. 5c). In more detailed observations of the weekly velocities of the feeding area (Wcam04), we noted a distinct peak in velocity in the second half of July (Fig. 6). These observations align with the measured velocities from a permanent GNSS device installed on a large boulder (Fig. 5, red location pin). The uppermost landslide area and the lower rock glacier are presumably driven primarily by permafrost creep. However, the driving mechanisms of the intermediate feeding section observed by Wcam04 remain unknown. To better understand the factors controlling velocity changes and the overall behaviour of these landforms, further investigation into environmental drivers such as rainfall and temperature is necessary—especially considering that the summer of 2022 was exceptionally warm and dry, with consecutive heat waves. However, this falls beyond the scope of the current pilot study.

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

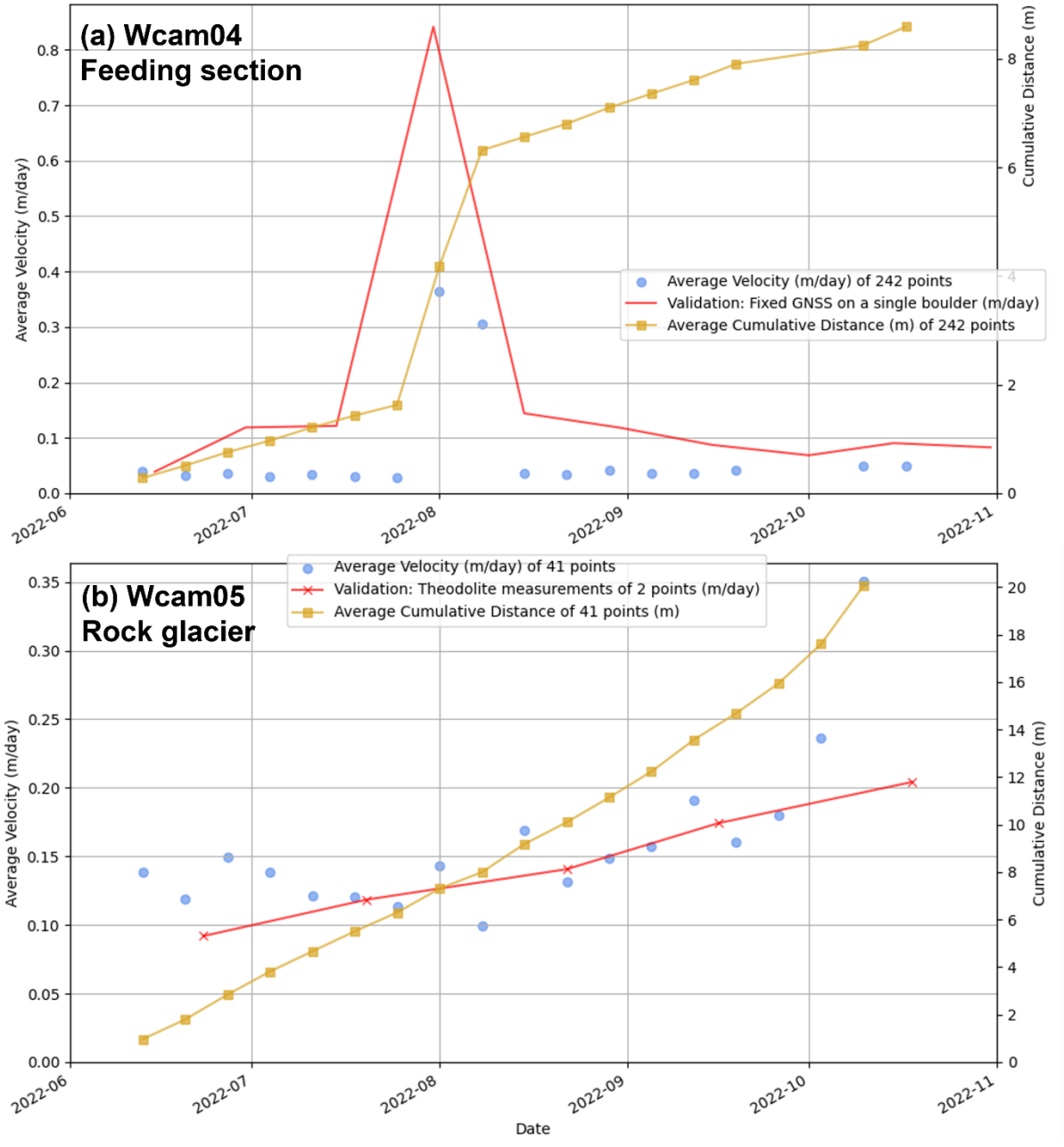


Figure 6: Graphs showing weekly velocities and cumulative distances. Panel a) displays data for the feeding section (Wcam04) and panel b) for the rock glacier (Wcam05) during the summer of 2022, derived from time-lapse imagery. Note that the data from October 26 is excluded from graph a) due to foggy conditions and was interpolated by the model. Validation points, marked in red, include measurements from the permanent GNSS installed on a large boulder in the upper feeding area (a) and two theodolite points on the rock glacier (b).

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

4. Discussion, limitations and outlook

4.1 General performance of the proposed workflow

The above-described methodology has the potential to tackle entire time-lapse image datasets to determine landform velocities, in this case creeping permafrost landforms. The results of our workflow show a good agreement with dGNSS, theodolite and permanent GNSS measurements, proving our method to be reliable, robust and fast for creating a better spatial (Fig. 5) and temporal coverage (Fig. 6) of the landform's displacement. Our pilot study demonstrates that significant new velocity information can be extracted from a basic and cost-effective device like a single webcam, greatly enhancing temporal acquisition frequency without the need for an array of cameras. This approach reveals sub-seasonal patterns, such as the short-term acceleration observed in Fig. 6, which reflects either actual motion related to permafrost creep or specific movements in the active layer, typically following significant water input from snowmelt or rain (RGIK, 2023). Spatial heterogeneity of landform movement (Fig. 5) is evident as well, including instances where larger boulders move faster, seemingly 'surfing' on the main landslide body. This phenomenon is illustrated in Fig. 6, where fixed GNSS data indicates faster overall movement compared to the entire landform, albeit following a similar temporal pattern. Although our proof-of-concept did not achieve the millimetre accuracy of fixed GNSS measurements, we were still able to detect displacements of 7 cm and 14 cm between consecutive frames using the setups of Wcam04 and Wcam05, respectively (Table 1). The measurements at the rock glacier (Wcam05) have a significantly lower spatial resolution (41 vs. 242 points tracked in the feeding area) due to the greater distance of the webcam from the area of interest and its suboptimal viewing angle (Fig. 1b). This increased distance also made the measurements more sensitive to slight camera movements, which, along with the lower level of detection, likely contributed to some of the higher variation in measured rock glacier velocity (Fig. 5b). The output of our workflow in a sense yields a sparse point cloud for every image frame in the sequence with absolute distance and velocity information as a scalar field, as visualised in Figure 5. This provides a great amount of spatial and temporal data in a manageable file format suitable for big data.

The superior performance of the PIPs++ model stems from tracking multiple timesteps jointly instead of frame-by-frame, enhancing temporal smoothness, coherency, and improving flow estimation accuracy (Hur and Roth, 2020). This makes PIPs++ especially suitable for environmental applications, where for example changes in light conditions are a common problem. Moreover, PIPs++ is trained on a very large and diverse artificial dataset PointOdyssey (Zheng et al., 2023), including rendered dynamic fog to account for (partial) occlusion and realistic in- and outdoor scenes. This entirely absent from other synthetic datasets like the FlyingChairs dataset, which was utilised to train models such as FlowFormer (Huang et al., 2022) or GMFlow (Xu et al., 2022), i.e. transformer-based models, and justifies our use of PIPs++ as a tracker in this proof-of-concept without the need of re-training the model with a sample of our own data. While previous research using monoscopic images to track a landslide (Travelletti et al., 2012) and a rock glacier (Kenner et al., 2018) was prone to mismatches because of its frame-by-frame strategy, our approach surpasses this limitation and at the same time makes it possible to process longer time periods and handle big data collected by hourly webcams more adequately. By also tracking points in stable areas, shifts

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

in camera position, a common problem in long time-lapse imagery sequences, can be corrected and used for stabilising the image sequence, similar as in Kenner et al. (2018). A main limitation of the method developed in the latter research is the requirement that the creep directions and the geometry of the mass wasting area remains unchanged due to the use of fixed azimuths and masks for georeferencing. This is solved in our methodology, since PIPs++ does not assume any movement direction (Zheng et al., 2023) and no masks are needed when scaling the data using our image-to-geometry approach (Elias et al., 2023).

4.2 Limitations

As PIPs++ still relies on appearance-matching cues similar to traditional methods, it remains sensitive to abrupt changes in appearance, such as substantial variations in snow cover or alterations in surface morphology due to e.g., extensive rockfalls. This issue was also highlighted as a major limitation by Kenner et al. (2018). However, because the algorithm can quickly process large amounts of data, we can leverage the full temporal resolution, allowing tracking to succeed as long as snow cover changes gradually, as shown in Figure 3d.

Another important limitation of our approach is that PIPs++ does not perform well in detecting movements for every pixel in the image due to the use of a regular grid compared to other frame-to-frame approaches. Currently, a regular grid size of 2000 points per image is used, limited by computational power. While PIPs++ works well with the low-resolution images in this study, higher-resolution images quickly reach the limits of our available computational resources. One major limitation is the need for specific and expensive computational setups (e.g., an NVIDIA RTX A6000 GPU with 48GB). Even with these resources, increasing the temporal window, image resolution, and tracked points can quickly hit the limits. Considering appropriate GPU and CPU computational power, around 30 min were needed to process the example dataset of this paper in an end-to-end fashion, from raw time-lapse images to velocity graphs. The majority of this time is needed to scale the output of the PIPs++ model, which is highly dependent on the size of the 3D point cloud.

Despite these limitations, we are confident that ongoing developments in AI processing will resolve these issues in the near future, making it possible to run AI frameworks on CPU only. Unlike PIPs++, the image-to-geometry approach does not require a GPU and can be performed on a laptop. However, to scale the time-lapse image data accurately, at least one UAV data survey is essential, and the quality of this scaling relies heavily on the quality of the UAV data. Additionally, a suboptimal oblique viewing angle, such as with Wcam05 (Fig. 1b), can complicate the matching process, as the rendered view will be limited to a narrow strip. This can affect the distribution of matches and compromise the estimation of the camera's interior orientation and lens distortions (Elias et al., 2019). Furthermore, our current assumption that the cameras remain stable over time justifies using a single image to determine the interior and exterior camera properties. However, for future work, it would be more accurate to perform this estimation for each time-lapse image individually, as can be done with a stereoscopic camera system (Ioli et al., 2024), to account for temperature-induced variations (Elias et al., 2020).

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

4.3 Outlook

There is still potential to optimize the proposed workflow to achieve full automation for processing entire image sequences. This includes automatically detecting moving and stable areas, which currently relies on user defined criteria and thus depends on prior site knowledge. Further enhancements to the image-to-geometry algorithm are necessary to increase its speed and computational efficiency for scaling several hundred timestamps simultaneously. Additionally, we need to conduct further investigations into the in-field calibration of intrinsic camera parameters and their stability over time (Elias et al., 2020). This analysis aims to determine if it significantly enhances the quality of scaled results, similar to the benefits that are expected when using multiple epochs of UAV data for updating the camera calibration. Improvements in 2D to 3D tools (image-to-geometry registration) could also benefit other disciplines, such as the quantitative analysis of historical terrestrial photographs for mapping historical rockfall (Wegner et al., 2023) and changes in glacier forefields (Altmann et al., 2020).

Moreover, additional research is necessary to evaluate whether AI-based tracking algorithms like PIPs++ could effectively track landform velocities using aerial datasets such as consecutive UAV orthophotos or satellite imagery. Furthermore, the developed workflow could prove valuable for analysing monoscopic time-lapse image sequences of other dynamic processes, such as lava flows (James and Robson, 2014), solifluction and gelifluction movements (Matsuoka, 2014), and flow velocities in rivers (Eltner et al., 2020; Stumpf et al., 2016). Given the prevalence of time-lapse camera data collection, a rapid and efficient method for automatically processing such extensive datasets holds significant scientific relevance. Furthermore, the fast and robust processing of the time-lapse imagery makes it possible to function as an early-warning system when processing can be carried out in near real-time, as indicated by Kenner et al. (2018).

5. Conclusion

This proof-of-concept demonstrates the potential of AI-based algorithms for tracking and matching points to improve motion estimations in time-lapse image sequences of a mountain landscape. Two fast-moving alpine landforms—a landslide and a rock glacier at the Grabengufer site in Switzerland—were selected as the pilot study area. The initial results presented in this paper show that robust and reliable velocity information can be quickly derived with minimal input data and user intervention. Our pilot study opens the door to processing entire image datasets to reveal spatio-temporal patterns that traditional monitoring methods have previously overlooked due to their limited spatial or temporal resolution.

The PIPs++ model, used for tracking features in image sequences, excels in widening the temporal window and includes a template-update mechanism that allows for changes in feature appearance. Its key advantage is its ability to accurately estimate occluded trajectories within the temporal frame, avoiding suboptimal matches and enhancing tracking accuracy, making it especially robust for environmental applications and eliminates the need for filtering blurry or foggy images as a pre-processing step. Additionally, the model's rapid performance—processing 400 images in two minutes to track features through a temporal window of 19 frames on an NVIDIA RTX A6000—is promising for handling large datasets and developing early-warning systems.

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

The image-to-geometry approach provides an accurate way to scale the 2D image data into 3D object space, even under suboptimal camera viewing angles and distance to the area of interest. Future research will focus on enhancing the image-to-geometry algorithm for efficient scaling of multiple timestamps, in-field camera calibration, and evaluating the potential benefits of using multiple epochs of UAV data.

This paper represents an important step forward in using monoscopic cameras and leveraging previously captured data that have not been processed automatically with metric values before. By significantly enhancing temporal acquisition frequency using basic time-lapse imagery with zero cost, we can achieve a level of data resolution that would be expensive with GNSS or UAV methods. Validated by discrete theodolite, GNSS, and continuous permanent GNSS measurements, our approach provides a spatially continuous understanding of landform movement. It allows data acquisition in areas where in-situ measurements are impractical due to logistical and safety constraints, and where other remote sensing techniques fail due to high landform displacements. Furthermore, depending on the image resolution, distance to the landform, and its velocity, our approach can achieve a sub-seasonal resolution of velocity information with an accuracy of several cm.

This study introduces a new tool for scientists to automatically extract velocity information from existing webcam datasets, extending its applicability to various phenomena, including lava flows, water flows, and creeping landforms such as landslides, rock glaciers, and solifluction lobes.

Code availability

Part of this work is based on existing algorithms, available under https://github.com/aharley/pips2 for PIPs++ tracker and https://github.com/cvg/LightGlue for LightGlue matching algorithm. The code used for this proof-of-concept is partly available under https://github.com/hannehendrickx/pips_env. Please follow this repository to receive further updates. All code is available upon reasonable request to the corresponding author. We encourage people with similar monoscopic time-lapse image datasets to reach out to the authors.

Data availability

The sample images and corresponding trajectory files used in this study are available at: https://github.com/hannehendrickx/pips_env. For an up-to-date overview of the sites, including the latest webcam images, please visit the following pages: Grabengufer Landslide and Grabengufer Rock Glacier.

Author contribution

HH, XB and AE designed the research. HH, XB and ME created the Python and C++ codes used in this paper. HH prepared and processed data, made all the figures and wrote the initial version of the paper. ME processed data with the image-to-

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

geometry code and provided input for the writing of those respective sections. RD conducted field investigations and provided field specific knowledge to aid the writing process. All authors improved and revised the initial manuscript.

Competing interests

The authors declare no conflicts of interests.

Acknowledgments

The authors thank Sebastian Pfaehler (University of Fribourg, CH) for capturing and processing the UAV data, and Sebastian Summermatter and Martin Volken (Geomatik AG / Planax AG) for the collection and processing of the theodolite data. We acknowledge the support of the municipality of Randa and the Canton of Valais (Service for Natural Hazards) in facilitating the field investigations. Geosat SA is acknowledged for setting up the permanent GNSS devices and handling data processing, and we thank the Canton of Valais for providing access to the permanent GNSS data. We also acknowledge Idelec SA for their assistance with webcam maintenance. We would like to thank the numerous collaborators at the University of Fribourg for their contributions to the bi-annual GNSS data acquisition, data processing, webcam maintenance, and image transfer and archiving. Air Zermatt is also thanked for ensuring safe transport to and from the site.
ChatGPT was utilised as a tool in this study to assist with coding and to enhance the English language quality of the manuscript through text input. It was not employed to generate original text during the research process.

References

Altmann, M., Piermattei, L., Haas, F., Heckmann, T., Fleischer, F., Rom, J., Betz-Nutz, S., Knoflach, B., Müller, S., Ramskogler, K., Pfeiffer, M., Hofmeister, F., Ressl, C., and Becht, M.: Long-Term Changes of Morphodynamics on Little Ice Age Lateral Moraines and the Resulting Sediment Transfer into Mountain Streams in the Upper Kauner Valley, Austria, Water 2020, Vol. 12, Page 3375, 12, 3375, https://doi.org/10.3390/W12123375, 2020.

Barsch, D.: Rockglaciers, Springer Berlin Heidelberg, Berlin, Heidelberg, https://doi.org/10.1007/978-3-642-80093-1, 1996.

Blanch, X., Guinau, M., Eltner, A., and Abellan, A.: Fixed photogrammetric systems for natural hazard monitoring with high spatio-temporal resolution, Natural Hazards and Earth System Sciences, 23, 3285–3303, https://doi.org/10.5194/nhess-23-3285-2023, 2023.

Cicoira, A., Marcer, M., Gärtner-Roer, I., Bodin, X., Arenson, L. U., and Vieli, A.: A general theory of rock glacier creep based on in-situ and remote sensing observations, Permafrost and Periglacial Processes, 32, 139–153, https://doi.org/10.1002/ppp.2090, 2021.

Cicoira, A., Weber, S., Biri, A., Buchli, B., Delaloye, R., Da Forno, R., Gärtner-Roer, I., Gruber, S., Gsell, T., Hasler, A., Lim, R., Limpach, P., Mayoraz, R., Meyer, M., Noetzli, J., Phillips, M., Pointner, E., Raetzo, H., Scapozza, C., Strozzi, T., Thiele,

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

L., Vieli, A., Vonder Mühll, D., Wirz, V., and Beutel, J.: In situ observations of the Swiss periglacial environment using GNSS instruments, Earth System Science Data, 14, 5061–5091, https://doi.org/10.5194/essd-14-5061-2022, 2022.

Delaloye, R. and Staub, B.: Seasonal variations of rock glacier creep: Time series observations from the Western Swiss Alps, in: XI. International Conference On Permafrost – Book of Abstracts, Bibliothek Wissenschaftspark Albert Einstein, 24 June 2016, Potsdam, Germany, 2016.

Delaloye, R., Lambiel, C., and Gärtner-Roer, I.: Overview of rock glacier kinematics research in the Swiss Alps: seasonal rhythm, interannual variations and trends over several decades, Geographica Helvetica, 65, 135–145, https://doi.org/10.5194/gh-65-135-2010, 2010.

Delaloye, R., Morard, S., Barboux, C., Abbet, D., Gruber, V., Riedo, M., and Gachet, S.: Rapidly moving rock glaciers in Mattertal, Jahrestagung der Schweizerischen Geomorphologischen Gesellschaft, 21–31, 2013.

Elias, M., Kehl, C., and Schneider, D.: Photogrammetric water level determination using smartphone technology, The Photogrammetric Record, 34, 198–223, https://doi.org/10.1111/phor.12280, 2019.

Elias, M., Eltner, A., Liebold, F., and Maas, H.-G.: Assessing the Influence of Temperature Changes on the Geometric Stability of Smartphone- and Raspberry Pi Cameras, Sensors, 20, 643, https://doi.org/10.3390/s20030643, 2020.

Elias, M., Weitkamp, A., and Eltner, A.: Multi-modal image matching to colorize a SLAM based point cloud with arbitrary data from a thermal camera, ISPRS Open Journal of Photogrammetry and Remote Sensing, 9, 100041, https://doi.org/10.1016/j.ophoto.2023.100041, 2023.

Eltner, A., Kaiser, A., Abellan, A., and Schindewolf, M.: Time lapse structure-from-motion photogrammetry for continuous geomorphic monitoring, Earth Surface Processes and Landforms, 42, 2240–2253, https://doi.org/10.1002/esp.4178, 2017.

Eltner, A., Sardemann, H., and Grundmann, J.: Technical Note: Flow velocity and discharge measurement in rivers using terrestrial and unmanned-aerial-vehicle imagery, Hydrology and Earth System Sciences, 24, 1429–1445, https://doi.org/10.5194/hess-24-1429-2020, 2020.

Eltner, A., Hoffmeiste, D., Kaiser, A., Karrasch, P., Klingbeil, L., Stöcker, C., and Rovere, A.: UAVs for the Environmental Sciences, 494 pp., 2022.

Fortun, D., Bouthemy, P., and Kervrann, C.: Optical flow modeling and computation: A survey, Computer Vision and Image Understanding, 134, 1–21, https://doi.org/10.1016/j.cviu.2015.02.008, 2015.

Frauenfelder, R., Haeberli, W., and Hoelzle, M.: Rockglacier occurrence and related terrain parameters in a study area of the Eastern Swiss Alps, Permafrost`, 253–258, 2003.

Harley, A. W., Fang, Z., and Fragkiadaki, K.: Particle Video Revisited: Tracking Through Occlusions Using Point Trajectories, 2022.

He, K., Zhang, X., Ren, S., and Sun, J.: Deep Residual Learning for Image Recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 770–778, https://doi.org/10.1109/CVPR.2016.90, 2016.

Heid, T. and Kääb, A.: Evaluation of existing image matching methods for deriving glacier surface displacements globally from optical satellite imagery, Remote Sensing of Environment, 118, 339–355, https://doi.org/10.1016/j.rse.2011.11.024, 2012.

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

Hermle, D., Gaeta, M., Krautblatter, M., Mazzanti, P., and Keuschnig, M.: Performance Testing of Optical Flow Time Series Analyses Based on a Fast, High-Alpine Landslide, Remote Sensing, 14, https://doi.org/10.3390/rs14030455, 2022.

How, P., Hulton, N. R. J., Buie, L., and Benn, D. I.: PyTrx: A Python-Based Monoscopic Terrestrial Photogrammetry Toolset for Glaciology, Frontiers in Earth Science, 8, 2020.

Huang, Z., Shi, X., Zhang, C., Wang, Q., Cheung, K. C., Qin, H., Dai, J., and Li, H.: FlowFormer: A Transformer Architecture for Optical Flow, in: Computer Vision – ECCV 2022, Cham, 668–685, https://doi.org/10.1007/978-3-031-19790-1_40, 2022.

Hur, J. and Roth, S.: Optical Flow Estimation in the Deep Learning Age, in: Modelling Human Motion: From Human Perception to Robot Design, edited by: Noceti, N., Sciutti, A., and Rea, F., Springer International Publishing, Cham, 119–140, https://doi.org/10.1007/978-3-030-46732-6_7, 2020.

Ioli, F., Dematteis, N., Giordan, D., Nex, F., and Pinto, L.: Deep Learning Low-cost Photogrammetry for 4D Short-term Glacier Dynamics Monitoring, PFG, https://doi.org/10.1007/s41064-023-00272-w, 2024.

James, M. R. and Robson, S.: Sequential digital elevation models of active lava flows from ground-based stereo time-lapse imagery, ISPRS Journal of Photogrammetry and Remote Sensing, 97, 160–170, https://doi.org/10.1016/j.isprsjprs.2014.08.011, 2014.

Kääb, A. and Reichmuth, T.: Advance mechanisms of rock glaciers, Permafrost and Periglacial Processes, 16, 187–193, https://doi.org/10.1002/ppp.507, 2005.

Kääb, A. and Vollmer, M.: Surface Geometry, Thickness Changes and Flow Fields on Creeping Mountain Permafrost: Automatic Extraction by Digital Image Analysis, Permafrost and Periglacial Processes, 11, 315–326, https://doi.org/10.1002/1099-1530(200012)11:4<315::AID-PPP365>3.0.CO;2-J, 2000.

Kenner, R., Phillips, M., Limpach, P., Beutel, J., and Hiller, M.: Monitoring mass movements using georeferenced time-lapse photography: Ritigraben rock glacier, western Swiss Alps, Cold Regions Science and Technology, 145, 127–134, https://doi.org/10.1016/j.coldregions.2017.10.018, 2018.

Kummert, M. and Delaloye, R.: Mapping and quantifying sediment transfer between the front of rapidly moving rock glaciers and torrential gullies, Geomorphology, 309, 60–76, https://doi.org/10.1016/j.geomorph.2018.02.021, 2018.

Kummert, M., Delaloye, R., and Braillard, L.: Erosion and sediment transfer processes at the front of rapidly moving rock glaciers: Systematic observations with automatic cameras in the western Swiss Alps, Permafrost and Periglacial Processes, 29, 21–33, https://doi.org/10.1002/ppp.1960, 2018.

Lindenberger, P., Sarlin, P.-E., and Pollefeys, M.: LightGlue: Local Feature Matching at Light Speed, in: ICCV, 2023.

Marcer, M., Bodin, X., Brenning, A., Schoeneich, P., Charvet, R., and Gottardi, F.: Permafrost Favorability Index: Spatial Modeling in the French Alps Using a Rock Glacier Inventory, Frontiers in Earth Science, 5, 105, https://doi.org/10.3389/feart.2017.00105, 2017.

Marcer, M., Cicoira, A., Cusicanqui, D., Bodin, X., Echelard, T., Obregon, R., and Schoeneich, P.: Rock glaciers throughout the French Alps accelerated and destabilised since 1990 as air temperatures increased, Communications Earth & Environment, 2, 1–11, https://doi.org/10.1038/s43247-021-00150-6, 2021.

Matsuoka, N.: Combining Time-Lapse Photography and Multisensor Monitoring to Understand Frost Creep Dynamics in the Japanese Alps, Permafrost and Periglacial Processes, 25, 94–106, https://doi.org/10.1002/ppp.1806, 2014.

https://doi.org/10.5194/egusphere-2024-2570
Preprint. Discussion started: 22 August 2024

Pellet, C., Bodin, X., Cusicanqui, D., Delaloye, R., Kaab, A., Kaufmann, V., Noetzli, J., Thibert, E., Vivero, S., and Kellerer-Pirklbauer, A.: State of the climate in 2022: Rock glaciers velocity, Bulletin of the American Meteorological Society, 104, 41–42, https://doi.org/10.1175/2023BAMSStateoftheClimate.1, 2023.

RGIK, R. glacier inventories and kinematics: Guidelines for inventorying rock glaciers: baseline and practical concepts (Version 1.0), https://doi.org/10.51363/unifr.srr.2023.002, 2023.

Sarlin, P. E., Detone, D., Malisiewicz, T., and Rabinovich, A.: SuperGlue: Learning Feature Matching with Graph Neural Networks, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 4937–4946, https://doi.org/10.1109/CVPR42600.2020.00499, 2020.

Schwalbe, E. and Maas, H.-G.: The determination of high-resolution spatio-temporal glacier motion fields from time-lapse sequences, Earth Surface Dynamics, 5, 861–879, https://doi.org/10.5194/esurf-5-861-2017, 2017.

Stumpf, A., Augereau, E., Delacourt, C., and Bonnier, J.: Photogrammetric discharge monitoring of small tropical mountain rivers: A case study at Rivière des Pluies, Réunion Island, Water Resources Research, 52, 4550–4570, https://doi.org/10.1002/2015WR018292, 2016.

Sun, J., Shen, Z., Wang, Y., Bao, H., and Zhou, X.: LoFTR: Detector-Free Local Feature Matching with Transformers, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 4, 8918–8927, https://doi.org/10.1109/CVPR46437.2021.00881, 2021.

Teed, Z. and Deng, J.: RAFT: Recurrent All-Pairs Field Transforms for Optical Flow, in: Computer Vision – ECCV 2020, Cham, 402–419, https://doi.org/10.1007/978-3-030-58536-5_24, 2020.

Travelletti, J., Delacourt, C., Allemand, P., Malet, J.-P., Schmittbuhl, J., Toussaint, R., and Bastard, M.: Correlation of multi-temporal ground-based optical images for landslide monitoring: Application, potential and limitations, ISPRS Journal of Photogrammetry and Remote Sensing, 70, 39–55, https://doi.org/10.1016/j.isprsjprs.2012.03.007, 2012.

Wegner, K., Stark, M., Haas, F., and Becht, M.: Suitability of terrestrial archival imagery for SfM-MVS based surface reconstruction of steep rock walls for the detection of rockfalls, Journal of Geomorphology, https://doi.org/10.1127/jgeomorphology/2023/0775, 2023.

Xu, H., Zhang, J., Cai, J., Rezatofighi, H., and Tao, D.: GMFlow: Learning Optical Flow via Global Matching, Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2022-June, 8111–8120, https://doi.org/10.1109/CVPR52688.2022.00795, 2022.

Zheng, Y., Harley, A. W., Shen, B., Wetzstein, G., and Guibas, L. J.: PointOdyssey: A Large-Scale Synthetic Dataset for Long-Term Point Tracking, 2023.