Georeferencing of Fire Front Aerial Images using Structure from motion and Iterative Closest Point

Francisco Sargento
francisco.sargento@tecnico.ulisboa.pt

Instituto Superior Técnico, Universidade de Lisboa, Portugal

October 2021

Abstract

This work proposes the use of Structure-from-motion (Sfm) and Iterative Closest Point (ICP) as a forest fire georeferencing algorithm to be used with images captured by an aerial vehicle. Sfm+ICP uses the real time video captured by an aircraft's camera, as well as its IMU and GPS measurements to reconstruct a dense 3D point cloud of the disaster area captured by the camera. The Sfm reconstruction is divided in two steps: a sparse reconstruction step using Speeded up robust features (SURF) for camera pose estimation, and a dense reconstruction step relying on a Kanade–Lucas–Tomasi (KLT) feature tracker initialized using the minimum eigenvalue algorithm. This dense 3D reconstruction is then registered to a real Digital Elevation Model (DEM) of the surrounding area, thus refining the point cloud to better match the terrain. The reconstruction is then used as the basis of the georeferencing estimates, as any target's location can be estimated by averaging the 3D coordinates corresponding to its nearby pixels. The algorithm was validated with a real forest fire video. The results demonstrate that Sfm+ICP can perform accurate 3D reconstructions while also georefering several targets in a forest fire environment. The results also show the algorithm is robust to high IMU and GPS errors, making it a far better option than optic-ray-based georeferencing for UAVs with unreliable telemetry.
Keywords: forest fire, UAV, structure from motion, georeferencing, DEM, ICP

1. Introduction

Forest fires have reached unprecedented figures in Portugal. In the 1980's, 75000ha of Portuguese land burned in forest fires. That number increased to 100000 ha in the 1990's and 150000ha in the 2000's, and that trend shows no signs of slowing down, with a series of socioeconomic and climate change related effects contributing to an ever increasing fire risk [5].

While downward economic cycles, budget scarcity, rural depopulation and forest land mismanagement have exacerbated Portugal's fire risk, other southern European nations have also been affected by forest fires. In the 2010's, over 3 million hectares of European land burned as a result of forest fires. Whilst the main victims were Portugal and the Mediterranean countries of Spain, Italy, France and Greece, climate change will continue to cause rising temperatures and decreasing rainfall, which will likely increase the length and severity of the fire season [3]. Additionally, high emissions climate models predict that by 2070 most of Central Europe may also experience regular and lengthy fire seasons [2].

UAVs are relevant components of modern firefighting operations, which will help us fight the next waves of fire seasons. UAVs possess rapid maneuverability, extended operational range, improved personal safety and cost efficiency, when compared to other remote sensing solutions, making them particularly useful in fire monitoring and detection, given their ability to perform fire search, confirmation and observation [20]. Georeferencing algorithms are a critical aspect of these remote sensing systems: by locating a fire quickly and accurately, fire monitoring systems can rely on quality data to be used by fire propagation models and firefighting authorities, saving lives and property.

Georeferencing is the process of assigning locations to geographical objects within a geographic frame of reference. In practice this means associating a set of 2D pixel coordinates from one image to a set of 3D world coordinates, most often expressed in latitude, longitude and altitude. There are two main methods to solve the georeferencing problem: direct and indirect georeferencing.

On the one hand, direct georeferencing uses navigational information and the camera to determine a target's geographic coordinates. Most direct georeferencing methods try to solve a single-ray backprojection problem: the process of determining the

ground coordinates of pixels in a single aerial image with the support of a DEM. Direct georeferencing is the most common method used in medium altitude georeferencing (between 500m and 1000m). Santana [16] presents a novel direct georeferencing algorithm to be used in forest fire scenarios. He proposes the use of Iterative Ray Tracing (IRT), which was initially proposed by Xiang in [18], in conjunction with an unscented transform to both estimate the target's position, and its uncertainty. This approach is well suited for medium altitude georeferencing, but its accuracy degrades dramatically when relying on error prone GPS and IMU measurements, since it has no way to correct large telemetry errors. Xu proposes an electro-optical stabilization and tracking platform which integrates the camera with the UAV's navigational data [19], in order to stabilize the camera's orientation. Xu also proposes the use of a Cubature Kalman filter. These techniques greatly improve the georeferencing accuracy of the direct algorithm, however, they also increase its implementation cost, complexity and computation time.

On the other hand, indirect georeferencing methods do not rely on navigational information, instead, they register photographs to georeferenced data, such as a DEM or satellite images. This registration can be done by directly comparing a photograph to a satellite image, but most state of the art algorithms process the image before the registration. For example, Structure from motion (Sfm) constructs a 3D point cloud from several sequential images, which can then registered to a DEM. The two most common medium altitude indirect georeferencing methods are Image registration and Sfm.

Image registration is the process of overlaying two images of the same scene taken at different times, from different viewpoints and with different sensors. The papers mentioned below use this process to register images taken by an UAV with georeferenced satellite/aerial images, and use the result of that registration to perform georeferencing. There are two types of image registration methods: correlation-based and pattern based. On the one hand, correlation-based methods place the sensed image at every location in the reference image, and adopt a similarity criteria to decide the most accurate location. On the other hand, pattern-based methods do not directly use image intensity levels, instead, they match patterns in both the sensed and reference image to find the best fit. Conte [7] proposes a correlation-based image registration approach. Correlation-based methods are efficient and can be applied to areas with no obvious landmarks, such as forests. However, correlation-based methods are not as accurate as pattern-based ones, being very sensitive to scene changes, such as smoke, fire or tree canopy height differences. Zhuo [21] uses Scale-Invariant Feature Transform (SIFT) in a pattern-based image registration approach, matching UAV images with satellite images. This approach achieves a decimeter level accuracy, however, it is not practical for medium altitude tilted camera platforms, especially when operating in dynamic environments, such as wild forest fires. Lindsten et al. [13] use environmental classification to classify and match superpixels. This method is more robust to orientation estimation errors, but not to dynamic environment changes. However, environmental classification outperforms classic feature detectors in the accuracy of pattern-based image registration in rural scenes. Hamidi [8] uses database matching techniques to refine the coarse initial altitude and position parameters of the camera derived from its navigational data. Image registration is used as a first refinement step, while a direct IRT method is used to perform the georeferencing step itself, taking advantage of the refined camera poses.

Structure from motion (Sfm), which was initiated by the computer vision research community, has now been widely used for automated triangulation of overlapping UAV-based frame imagery. Sfm is a set of computer vision algorithms that facilitate the photogrammetric reconstruction of 3D scenes from images alone. It has gained popularity in recent years due to its ability to deal with sets of unordered and heterogeneous images without prior or accurate knowledge of the camera's intrinsic and extrinsic parameters. Sfm is used for forest remote sensing in [4] and [14]. These papers prove that Sfm can be used for medium altitude forest remote sensing with meter level accuracy and LIDAR equivalent point cloud density. They also prove that it can be run in real time, and perform georeferencing and data collection at the same time. Notwithstanding Sfm faces some obstacles in forest environments, as is highlighted by Iglhaut et al. [10]: feature extractors perform poorly in forests, where the scene is often dynamic, and sudden illumination changes are commonplace.

The present work proposes a Sfm+ICP georeferencing method. Here, Sfm is used to reconstruct a 3D model of the fire area using techniques most similar to [4] and [12]. The reconstruction is then improved by matching it to a known DEM of the terrain around the aircraft, similarly to [17] but with a more flexible registration algorithm that matches the downsampled reconstructed point cloud and the upsampled DEM. This registration is in essence similar to DEM matching, but instead of matching DEMs, the algorithm matches high density point clouds, which is more computationally intensive, but yields better registration results. Using Sfm allows the algorithm to densely reconstruct the op-

erational area and consequently georefer dozens of targets at the same time with no increase in computation time. Using it in real time is also challenging due to some of its time intensive routines, thus the algorithm needs to use a simple and quick kind of Sfm that still guarantees high georeferencing accuracy. Moreover, using the ICP makes this algorithm robust to telemetry errors, even when they are exceedingly large.

2. Background

The Sfm+ICP algorithm is naturally comprised of two main blocks: Structure from motion and Iterative Closest Point. Both these subroutines are introduced in the following subsections.

2.1. Structure from motion

Structure from motion is the process of estimating the 3D structure of a scene from a set of images. It can produce high quality, dense, 3D point clouds of a landform for minimal financial cost [11]. Sfm is already well established in fields such as archaeology and cultural heritage [15], and its use as a topographic survey technique has surged in recent years.

The Sfm problem can be formulated as:

Given: m images of n fixed 3D points.
Problem: Estimate the m projection matrices, P_i, and n 3D points, X_j, from the mn correspondences, x_{ij}, assuming the following camera model holds true:

$$x_{ij} = P_i X_j, \quad i = 1, ..., m \ , \ \ j = 1, ..., n, \qquad (1)$$

where P_i are 3x4 matrices and X_j and x_{ij} are, respectively, 3D and 2D points in homogeneous coordinates.

Figure 1 shows a simple Sfm example with only 3 images. In real applications, not all the 3D points, X_j, are present in every single image, and the point correspondences, x_{ij}, contain outlier matches, however, the basic idea behind Sfm still holds true.

Sfm is not a single technique, it is a workflow, employing multiple algorithms developed from 3D computer vision, traditional photogrammetry and more conventional survey techniques. Most published Sfm implementations follow the same 5 step process [11]: The **first** step is to detect keypoints in all the m images. Each keypoint is a specific pixel that contains a distinct feature. **Then**, the keypoints are matched across all images, in order to find the correspondences x_{ij}. **Thirdly** Multi-view triangulation (MVT) is used to initially estimate the 3D points, X_j, and the projection matrices, P_i. **Next,** Bundle adjustment (BA) is performed to refine the initial X_j and P_i estimates provided by the previous step. Some Sfm applications iterate MVT and BA to converge on a better 3D reconstructions, while others do this just once. **Finally,** a Linear similarity transformation is applied: the X_j estimates provided by the BA are expressed in an arbitrary reference frame relative to the camera. Hence, Sfm's final step is to scale, translate and rotate the 3D points, X_j, to get a reconstruction expressed in N-E-D reference frame.

Sfm's principal output is a 3D point cloud of the scene captured by the images it processed. The reconstruction's point density depends mostly on the feature extractor used and the type of scene being surveyed: the more distinct features a scene has, the more dense its reconstruction will be. On the other hand, the reconstruction's georeferencing accuracy hinges on the accuracy of the camera model, GPS and IMU. Sfm has no way to minimize the effect of these errors, hence why ICP is used.

2.2. Iterative Closest Point

ICP is a point cloud registration framework used to find a local transformation to align two sets of points. Each step of the ICP algorithm attempts to align the two point clouds closer and closer, converging on a rotation matrix and translation vector that minimize the average distance between the point two clouds [6].

The ICP algorithm is based on a simple four step iterative process, designed to sequentially bridge the gap between the fixed and moving point clouds:

1. Select points on one point cloud.
2. Find the closest points on the other (fixed) point cloud.
3. Minimize the sum of the distances between the two sets of points. This produces a new point cloud more similar to the fixed one than in the previous iteration.
4. Repeat the previous 3 steps until the the two point clouds are similar enough.

In most geoscientifical applications, the moving point cloud is recovered from the environment using some point cloud extraction sensor, such as LIDAR, RADAR or Sfm. On the other hand, the fixed point cloud is often a Digital Elevation Model (DEM) of the environment being worked on.

DEMs were traditionally created from existing topographic maps and aerial photos, however, most recent DEMs are generated using synthetic-aperture radar usually mounted on a satellite, or from even newer techniques such as LIDAR. Europe's most accurate DEM is the EU-DEM v1.1 [1] which was made by merging elevation data from the Shuttle Radar Topography Mission (SRTM) and the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) DEMs. SRTM

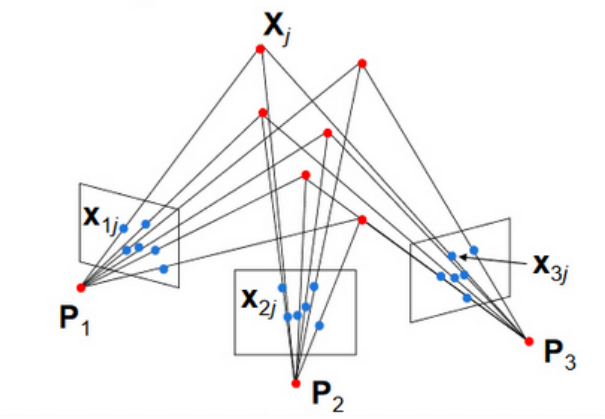


Figure 1: Structure from motion problem [9]

was generated using synthetic-aperture radar while ASTER was made from satellite images. EU-DEM v1.1 combines the perks of both ASTER and SRTM, resulting in a 30m resolution DEM with a 7m RMSE of most of Europe's countries.

3. Implementation

Sfm+ICP's main input is a set of sequential photographs/video frames taken by an aircraft's camera, which allows the Sfm block to densely reconstruct the fire region. However, without additional knowledge of the aircraft's orientation and localization, that reconstruction is not grounded to Earth coordinates. Hence, the Sfm+ICP algorithm also uses the aircraft's GPS and the camera's IMU in order to transform the reconstructed point cloud from the camera's reference frame to an inertial N-E-D reference, centered on a point with known coordinates. That transformation is determined by registering the point cloud to a DEM of the surrounding area. Finally, the algorithm receives a set of pixels corresponding to the targets that the user intends to georefer.

With those inputs, the algorithm produces two outputs: the targets' geographic coordinates and a dense 3D point cloud of the fire area.

The SFM+ICP algorithm has four sequential main blocks: motion estimation, dense reconstruction, point cloud registration and target georeferencing.

Motion estimation estimates the camera's pose at the time each photograph was taken. This step's algorithm can be seen in Algorithm 1. The camera poses are estimated using a simple and quick Sfm, often called "sparse Sfm" in contrast to the "dense Sfm" used in the next step. Sparse Sfm is a Sfm algorithm that uses sparse features, allowing it to be run several times faster than dense Sfm. The attribute sparse comes from the fact that this Sfm step extracts only the best quality features, rejecting lower quality ones, which allows it to accurately estimate the camera's pose in each of the images. These poses are this block's main output, and will be used in the next step to produce the dense point cloud, which is then used as the basis for the georeferencing estimates.

Algorithm 1 Motion estimation algorithm

Inputs: images, IP.
Outputs: camera poses.

1: Initialize the first camera pose as the global reference frame's origin.
2: Extract SURF features from the first image.
3: **for all** remaining images **do**
4: Extract SURF features from the image.
5: Match these features to the last image's features.
6: Estimate the pose of the camera relative to the previous camera, using the matched features.
7: Convert this relative pose to a global pose.
8: **end for**
9: Estimate the 3D location of all the features matched between all the images, using Multi-view triangulation.
10: Refine the camera poses and feature 3D locations using Bundle adjustment.

Dense reconstruction generates a dense point cloud of the area captured in the images. The dense reconstruction algorithm can be seen in Algorithm 2 .It starts by detecting features in each image and matching them across all the images. Multi-view triangulation is then used to coarsely estimate the 3D location of these features, which is then refined using a single bundle adjustment. These three steps produce an unscaled point cloud in the camera's reference frame, that is transformed to a scaled inertial reference frame using the telemetry data. The final step is to remove outliers from the point cloud,

using their location and reprojection error.

Algorithm 2 Dense reconstruction algorithm

Inputs: images, IP, refined camera poses, GPS, altitude and camera orientation (given by the IMU).
Outputs: dense point cloud in an inertial North-East-Down reference frame.

1: Initialize the KLT point tracker with Min Eigen features extracted from the first image.
2: **for all** remaining images **do**
3: Find the previous tracked features on the next image, using the KLT.
4: **end for**
5: Estimate the 3D location of all the points tracked, using multi-view triangulation.
6: Exclude points with reprojection error higher than 2 pixels.
7: Refine the camera poses and feature 3D locations using Bundle adjustment.
8: Rotate the point cloud using the pitch and heading of the first camera pose.
9: Translate the point cloud using the altitude of the first camera pose.
10: Scale the point cloud using the ratio between the refined camera poses and the GPS.
11: Remove outliers from the point cloud using their distance and elevation.

Point cloud registration improves the point cloud's precision by matching it to a real DEM of the surrounding area using ICP, *i.e.* translating and rotating the point cloud in a way that minimizes the distance between the real map and the reconstruction. This step's algorithm can be seen in Algorithm 3. The DEM is interpolated to match the point cloud's resolution, which is downsampled to a 2.5m resolution. This specific downsampling and interpolation step for the point cloud and the DEM was chosen, because it yielded a better horizontal accuracy than using the standard 25m resolution DEM . Finally, the ICP's RMSE is used to evaluate the algorithm success up to this point: if the RMSE is too high, it means that the 3D reconstruction is of poor quality, thus the algorithm discard the current point cloud and goes back to the first step, motion estimation.

Target georeferencing computes the targets' geolocation using the refined dense reconstruction from the last step. This the simplest and quickest of the four main blocks, as it simply averages the 3D position of the 10 closest features to the target in order to estimate its location. This final step's algorithm can be seen in the last four lines of Algorithm 3.

Algorithm 3 Registration and target georeferencing algorithm

Inputs: DEM, dense point cloud of the terrain, GPS coordinates, target pixels.
Outputs: target coordinates , rectified point cloud.

1: Select a relevant subsection of the EU DEM v1.1 using the GPS coordinates.
2: Interpolate the selected DEM to a 2.5 metre resolution.
3: Downsample the dense point cloud to a 2.5 metre resolution.
4: Perform ICP between the downsampled point cloud and the interpolated DEM. Use a point to plane minimization metric.
5: Apply the rigid transformation provided by the ICP to the full dense point cloud.
6: **for all** target pixels **do**
7: Find the 10 closest features to each target pixel and their corresponding 3D locations in the rectified point cloud.
8: Average the location of the 10 closest features, using Inverse distance weighting, where each location estimate is weighted according to its pixel distance from the target.
9: **end for**

4. Results

The dataset used to validate the algorithm is a series of frames taken from a video captured by a fixed wing UAV loitering above a forest fire near Pombal, Portugal at 39.832856N -8.519885E at the 16th of August 2019.

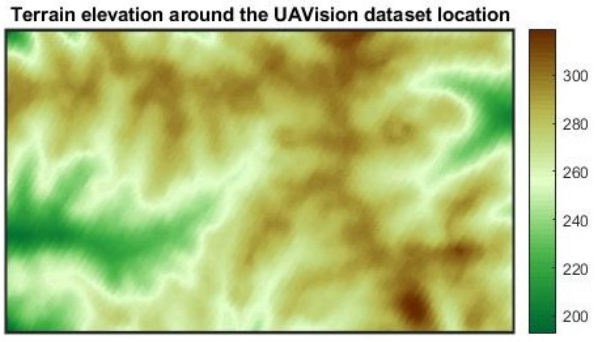


Figure 2: DEM around the fire

Figure 3 shows two images from the dataset. This dataset was sampled at a rate of 2Hz from 15s of continuous footage. The footage is unstable, with some sharp camera movements and video cutoffs that degrade the reconstructions quality, hence only

a short part of the video was usable.

The two frames in figure 3 show a typical Portuguese forest fire: a small village surrounded by a dense and vast forest and a large column of smoke above it. This column of smoke obscures part of the terrain under and behind it, hindering the ICP registration. At the same time, the forest surrounding the town is dense and has few distinct features, further complicating the reconstruction process. The last hindrance can be seen in Figure 2: the terrain elevation range is only 100m and there are no large hills to assist the ICP registration. However, the terrain is still complex and distinct enough to perform Sfm+ICP georeferencing.

(a) Frame 0

(b) Frame 30

Figure 3: First and last frame used

4.1. Reconstruction results

Figure 4 shows the reconstructed point cloud and the true camera trajectory. The aircraft performs a coordinated turn heading northwest with a slight bank angle, naturally changing its distance to the fire in all three spatial coordinates. The camera changes its heading and pitch throughout the video in order to track the fire's progress, and inadvertedly changes its rotation due to flight induced disturbances.

The reconstruction can also be seen in detail in figure 5. Comparing this reconstruction with figure 3, one can see that all of the village is reconstructed with a high point density. The forests surrounding it are also well represented, although with a smaller point density. The smoke seen in figure 3 obscures the terrain northeast of the town, and most of the smoke is not considered as a feature, with only a small amount of gray points hovering north of the town at 400m altitude. These smoke points should be considered outliers, as they serve no purpose for the georeferencing algorithm, but the fact that a small amount of them managed to bypass the outlier removal block is not serious, as the ICP can easily classify these points as outliers, since the rest of the reconstruction is close enough to the real DEM.

Table 1 provides statistics regarding Sfm's performance with this dataset: 48ha were reconstructed, with an average of 96 points per hectare. While this density is not as high as one might have predicted, it is more than enough to accurately georefer all the buildings in the village and several targets outside of it, as well as the firefront itself. Only 48ha were reconstructed, however, this is mostly due to the smoke obscuring the terrain behind it, so the camera is unable to capture more.

Table 1: Reconstruction statistics

Reconstruction area [ha]	Average points per reconstruction	Points per hectare
48	4616	96

4.2. Georeferencing results

Figure 6 shows the average XY and Z errors for several sequential algorithm runs, and table 2 presents basic statistics regarding those results. All of these runs were made sequentially and with the same inputs and tuning parameters, however, they do present different outputs, since Sfm is not a deterministic process, having some degree of stochasticity. They show an average XY georeferencing error of just 49.1m, which is on par with state of the art medium altitude georeferencing algorithms. The XY results show a good error dispersion, with a relatively small standard deviation, meaning that the algorithm is, not only accurate, but also consistent. There is one clear outlier with 81.2m XY error, however most results are roughly in the 30m-60m range.

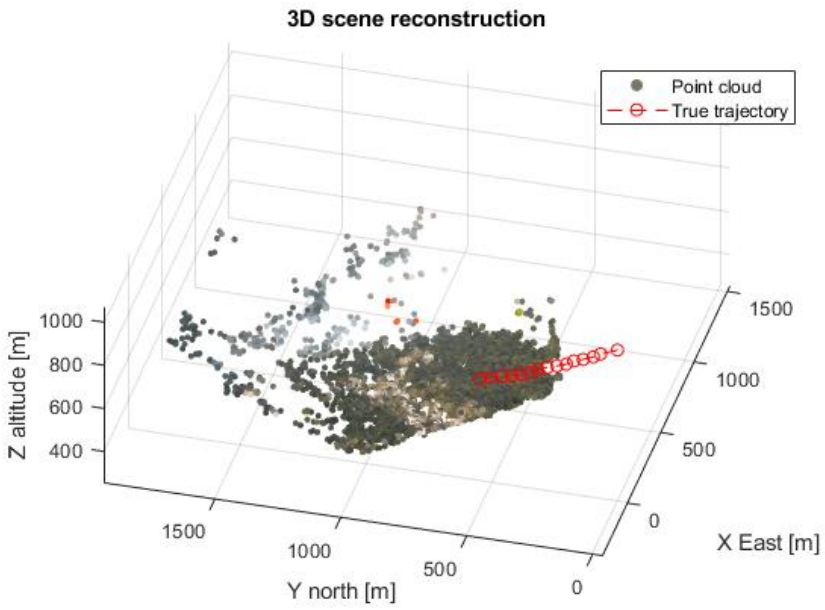


Figure 4: 3D reconstruction and aircraft trajectory

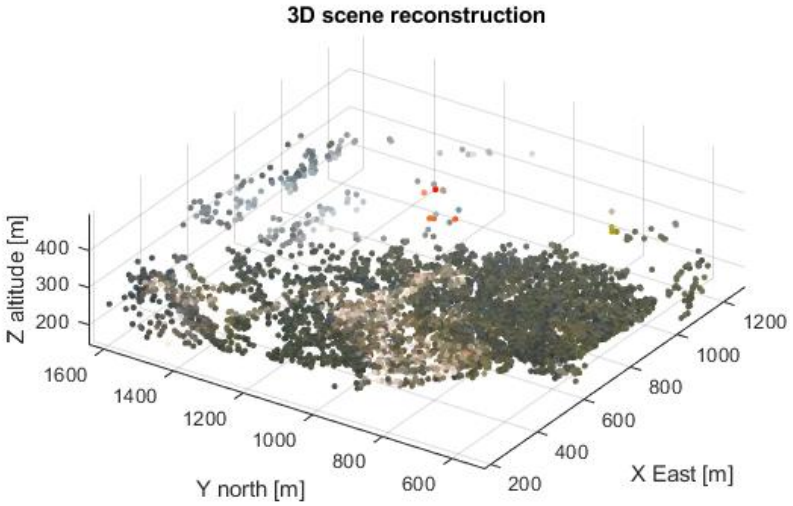


Figure 5: 3D reconstruction with equal scale

Table 2: Georeferencing accuracy statistics after ICP

	XY [m]	Z [m]
Max	81.2	6.9
Mean	49.1	4.7
Median	49.4	4.7
Standard deviation	16.7	1.3
Range	52.7	3.7

The average Z error shown in figure 6 and Table 2 is just 4.7m, and its standard deviation is only 1.3m. These statistics demonstrate that Sfm+ICP can also consistently and accurately determine a target's elevation, even if the camera is flying at 900m altitude. While this Z accuracy can be a useful property in some instances, a target's elevation can also be estimated by sampling a DEM at the desired coordinates. Notwithstanding, the low Z error also means that the reconstruction is consistent with the terrain, therefore the dense reconstruction

is useful for more than just georeferencing, for instance, it can be used as a new dynamic map of the disaster area.

Figure 6 reveals a correlation between the XY and Z errors: reconstructions with high XY also tend to have a high Z error, which is unsurprising due to the way that the ICP corrects both the XY Z errors. However, it shows that, since the Z error can be in part estimated using a DEM, this correlation can be used to assess if a reconstruction is horizontally accurate or not.

Figure 7 shows the XY and Z error distribution for the four targets georefered. The results show that the four targets have a low average XY error, and that some even have a similar accuracy as the DEM itself. There is one clear outlier reconstruction that had a very poor accuracy, however, apart from it, the algorithm performed consistently and accurately across all targets.

The Z error also shows a low error somewhat consistently across the four targets, however, here the outlier's impact is more clear: the mean Z error of each target (represented by red lines) is much lower than the median, for every single target. While the difference between the mean and the median for each target is no more than 2m, it shows that the vertical error is much more sensible to bad registrations than initially thought. This is likely due to this dataset's high pitch error (around 5 degrees), which makes the registration step all the more difficult. Despite this, the vertical results are still on par with the state of the art, and could be further improved if more stringent reprojection error and ICP RMSE constraints were applied, at the cost of a longer algorithm run time and sparser reconstruction.

5. Conclusions

This work presented a new and robust Sfm+ICP georeferencing algorithm designed for medium altitude forest fire monitoring. Its secondary achievement is the algorithm's ability to perform a reconstruction of the area around the fire. Both these features are useful for the FIREFRONT project, as they further the project's ability to perform real time georeferencing in adverse conditions, while also opening the door to other Sfm based research in the field of forest fire monitoring.

The algorithm is designed to work with any medium altitude aircraft with a camera, IMU and GPS, and it is able to match state of the art algorithms accuracy-wise, even when these sensors have a high level of noise. This robustness to IMU and GPS errors is especially relevant since optic-ray-based georeferencing algorithms are unable to fix these high sensor noises by themselves, therefore having a Sfm-based algorithm, such as the one presented in this work, provides the FIREFRONT project with a valuable georeferencing alternative.

The next step in Sfm+ICP research is to test it in real time with a predefined flight pattern and an accurate way to assess the algorithm's georeferencing accuracy. This endeavour would also be invaluable to other georeferencing algorithms, as all of them would be uniformly and accurately evaluated using a dataset especially made for just that.

The algorithm can still be subject to some alterations to increase its accuracy and efficiency, namely: a more complex feature tracking solution, having multiple ICP iterations with different parameters, or integrating fire segmentation data with the target georeferencing step.

Lastly, Sfm+ICP could be used as a measurement model in a SLAM implementation, or as an initialization step in a direct georeferencing algorithm (or vice-versa).

Acknowledgements

The author would like to thank UAVision for providing the forest fire footage used in this work, as well as Alexandre Bernardino and Ricardo Ribeiro for assisting in the design of this algorithm.

This work was made in the context of the FCT project FIREFRONT (PCIF/SSI/0096/2017).

References

[1] EU-DEMv1.1-flyer. https://land.copernicus.eu/user-corner/publications/eu-dem-flyer/at_download/file. Accessed: 2021-05-14.

[2] Projections of Fire Weather Index (PESETA III) — European Environment Agency. https://www.eea.europa.eu/data-and-maps/data/external/peseta-iii-projections-of-fire. Accessed: 2021-05-14.

[3] Forest fires — European Environment Agency. https://www.eea.europa.eu/data-and-maps/data/external/peseta-iii-projections-of-fire, 2019. Accessed: 2021-05-14.

[4] M. Alonzo, H. E. Andersen, D. C. Morton, and B. D. Cook. Quantifying boreal forest structure and composition using UAV structure from motion. *Forests*, 9(3):1–15, 2018.

[5] M. Beighley and A. C. Hyde. Portugal Wildfire Management in a New Era: Assessing Fire Risks, Resources and Reforms. *Independent report*, 9(1):52, 2018.

[6] P. J. Besl and N. D. McKay. A Method for Registration of 3-D Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

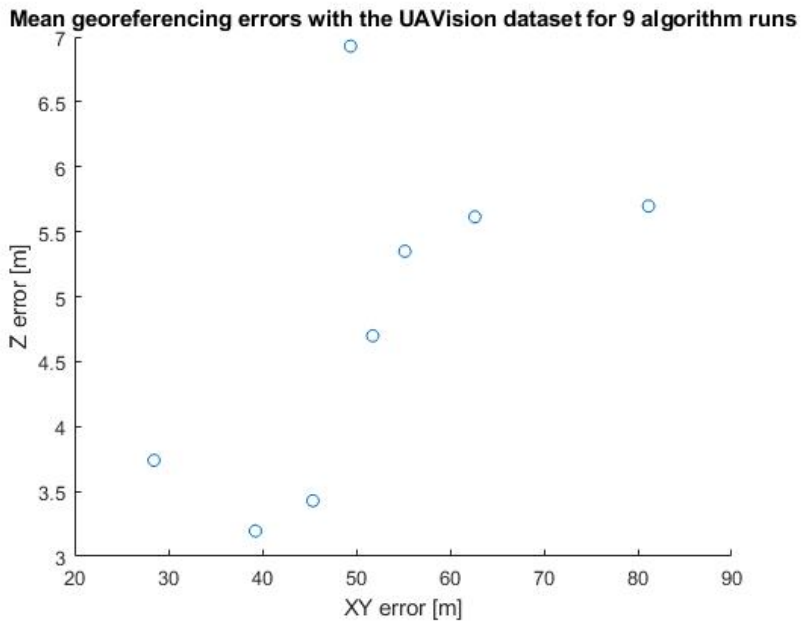
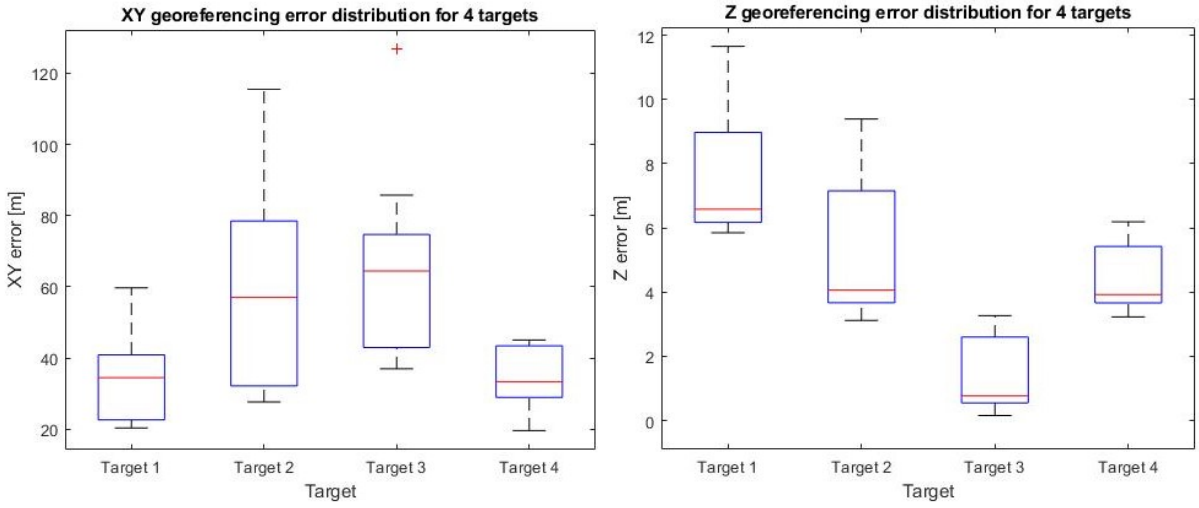


Figure 6: Georeferencing accuracy after ICP for the UAVision dataset



(a) XY error by target

(b) Z error by target

Figure 7: XY and Z error for each target

[7] G. Conte, M. Hempel, P. Rudol, D. Lundström, S. Duranti, M. Wzorek, and P. Doherty. High accuracy ground target geo-location using autonomous micro aerial vehicle platforms. *AIAA Guidance, Navigation and Control Conference and Exhibit*, pages 1–14, 2008.

[8] M. Hamidi and F. Samadzadegan. Precise 3D geo-location of UAV images using georeferenced data. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 40(1W5):269–275, 2015.

[9] R. Hartley and A. Zisserman. Multiple View Geometry in Computer vision. *Computer-Aided Design*, 16(2):672, 2003.

[10] J. Iglhaut, C. Cabo, S. Puliti, L. Piermattei, J. O'Connor, and J. Rosette. Structure from Motion Photogrammetry in Forestry: a Review. *Current Forestry Reports*, 5(3):155–168, 2019.

[11] J. L.Carrivick. *Structure from motion in the geosciences*. 1377.

[12] M. Lhuillier. Incremental Fusion of Structure-from-Motion and GPS using Constrained Bun-

dle Adjustment. 34(12):1–20, 2017.

[13] F. Lindsten, J. Callmer, H. Ohlsson, D. Törnqvist, T. B. Schön, and F. Gustafsson. Geo-referencing for UAV navigation using environmental classification. *Proceedings - IEEE International Conference on Robotics and Automation*, pages 1420–1425, 2010.

[14] A. Matese, S. F. Di Gennaro, and A. Berton. Assessment of a canopy height model (CHM) in a vineyard using UAV-based multispectral imaging. *International Journal of Remote Sensing*, 38(8-10):2150–2160, 2017.

[15] O. Özyeşil, V. Voroninski, R. Basri, and A. Singer. A survey of structure from motion. *Acta Numerica*, 26:305–364, 2017.

[16] B. Santana, A. Bernardino, and R. Ribeiro. Direct Georeferencing of Forest Fire Aerial Images using Iterative Ray-Tracing and a Bearings-Range Extended Kalman Filter. Master's thesis, Tecnico Lisboa, 2020.

[17] D. g. Sim and R. h. Park. Localization Based on DEM Matching Using Multiple Aerial Image Pairs. *EEE Transactions on Image Processing*, 11(1):52–55, 2002.

[18] H. Xiang and L. Tian. Method for automatic georeferencing aerial remote sensing (RS) images from an unmanned aerial vehicle (UAV) platform. *Biosystems Engineering*, 108(2):104–113, 2011.

[19] C. Xu, D. Huang, and J. Liu. Target location of unmanned aerial vehicles based on the electro-optical stabilization and tracking platform. *Measurement: Journal of the International Measurement Confederation*, 147:106848, 2019.

[20] C. Yuan, Y. Zhang, and Z. Liu. A survey on technologies for automatic forest fire monitoring, detection, and fighting using unmanned aerial vehicles and remote sensing techniques. *Canadian Journal of Forest Research*, 45(7):783–792, 2015.

[21] X. Zhuo, T. Koch, F. Kurz, F. Fraundorfer, and P. Reinartz. Automatic UAV image geo-registration by matching UAV images to georeferenced image data. *Remote Sensing*, 9(4):1–25, 2017.