J Intell Robot Syst (2011) 64:57–76
DOI 10.1007/s10846-011-9546-8

Safety, Security, and Rescue Missions with an Unmanned Aerial Vehicle (UAV)

Aerial Mosaicking and Autonomous Flight at the 2009 European Land Robots Trials (ELROB) and the 2010 Response Robot Evaluation Exercises (RREE)

Andreas Birk · Burkhard Wiggerich · Heiko Bülow · Max Pfingsthorn · Sören Schwertfeger

Received: 2 April 2010 / Accepted: 18 January 2011 / Published online: 29 January 2011

Abstract Several missions with an Unmanned Aerial Vehicle (UAV) in different realistic safety, security, and rescue field tests are presented. First, results from two safety and security missions at the 2009 European Land Robot Trials (ELROB) are presented. A UAV in form of an Airrobot AR100-B is used in a reconnaissance and in a camp security scenario. The UAV is capable of autonomous waypoint navigation using onboard GPS processing. A digital video stream from the vehicle is used to create photo maps—also known as mosaicking—in real time at the operator station. This mapping is done using an enhanced version of Fourier Mellin based registration, which turns out to be very fast and robust. Furthermore, results from a rescue oriented scenario at the 2010 Response Robot Evaluation Exercises (RREE) at Disaster City, Texas are presented. The registration for the aerial mosaicking is supplemented by an uncertainty metric and embedded into Simultaneous Localization and Mapping (SLAM), which further enhances the photo maps as main mission deliveries.

Keywords Unmanned Aerial Vehicle (UAV) · Autonomy · Registration · Mosaicking · Visual SLAM

A. Birk (✉) · H. Bülow · M. Pfingsthorn · S. Schwertfeger
Jacobs University Bremen, Campus Ring 1, 28759 Bremen, Germany
e-mail: a.birk@jacobs-university.de
http://robotics.jacobs-university.de

B. Wiggerich
AirRobot GmbH & Co. KG, Werler Strasse 4–8, 59755 Arnsberg, Germany
e-mail: info@airrobot.de
http://www.airrobot.de

1 Introduction

Aerial vehicles are obviously well suited to give a bird's eye view over an incident, surveillance, or reconnaissance site. They offer hence interesting options as robotics tools for assisting in safety, security, and rescue missions. We present here results from using an UAV in two field tests where different realistic response scenarios were investigated, namely the 2009 European Land Robot Trials (ELROB-2009) and the 2010 Response Robot Evaluation Exercises (RREE-2010).

ELROB-2009 was the fourth event in an annual series of evaluation events for Safety, Security, and Rescue Robotics (SSRR) [10]. ELROB-2009 took place from 14.-18. June 2009 in Oulu, Finland. The results from ELROB presented in this article are based on a joined participation of the German company Airrobot with the Jacobs Robotics Group. The team used an Airrobot AR100-B Unmanned Aerial Vehicle (UAV) in two SSRR missions, namely reconnaissance and camp security, i.e., two safety and security related missions. The Airrobot AR100-B is a quadrocopter (Fig. 1), which is capable of onboard processing of GPS signals and of stored way points, hence allowing fully autonomous operations. It is also equipped with a high quality digital video link to the operator station.

A third application scenario was investigated at the 2010 Response Robot Evaluation Exercise (RREE-2010) at Disaster City in College Station, Texas [24]. RREE is an annual event organized by the Intelligent Systems Division (ISD) of the National Institute of Standards and Technology (NIST). At this event, the task for the aerial vehicle was to gather an overview of a rubble pile, i.e., the mission is more related to rescue than safety and security. A slightly different Airrobot model was used at RREE-2010, namely an older AR100. Both UAV weigh less than one kilogram and have about one meter diameter, i.e., they are easily transportable and deployable.

Data from an onboard video link was used in all missions for photo mapping—also known as mosaicking—by registering the video frames with each other and hence providing a larger overview. These photo maps serve two main purposes. First, they help the operator in the mission: they stabilize the video stream, which is quite shaky, and they provide a much larger overview than a video. Second, they are an important mission deliverable. Even though there are many Geo Information Systems (GIS)

Fig. 1 The Airrobot AR100-B taking off for a mission at ELROB 2009 (*left*) and hovering in the air (*right*)

available including widespread popular ones like Google Maps [12] and Microsoft Bing Maps [18], they have severe limitations. First of all, there are still surprisingly many regions that are not or only poorly covered by certain systems. Google Maps has for example only an extremely low resolution coverage of Oulu and especially of the competition site where ELROB-2009 took place (Fig. 2). Also, GIS may provide inaccurate information (Fig. 3) and it is even in the best case less precise than what can be achieved with a low flying UAV. But the most important motivation for real be achieved with a low flying UAV. But the most important motivation for real time photo mapping is that it provides an older data stored in a GIS.

In addition to the use of photo-maps for aiding the operator and for using them as mission deliverable, there is a third interesting option, namely the use of the photo-maps in the context of cooperating UAV and ground robots. As pointed out in [15], UAV can easily cover larger areas searching for targets, whereas unmanned ground vehicles (UGV) are in comparison slow and difficult to maneuver as they have to avoid obstacles like fences, walls, and so on. But UGV offer more payload, typically have more accurate sensors, and they can be equipped with direct intervention possibilities, e.g., in form of a manipulator. Cooperating UAV and UGV are therefor an option to combine the best of both worlds, i.e., to use UAV to generate large overview maps in a fast manner, to locate targets, and to guide UGV to them that can then gather more information and intervene with the targets.

Naively, the generation of a photo map is trivial as long as the exact pose of the vehicle is know. But GPS localization is very imprecise, the 3D orientation of the UAV is hard to determine, and so on. An alternative is to use registration of subsequent frames from a camera on the vehicle. Related work in the areas of visual odometry and visual Simultaneous Localization and Mapping (SLAM) predominantly uses feature based approaches building for example upon the scale

Fig. 2 The ELROB 2009 competition site is only poorly covered in Google maps; the *above image* shows the best available resolution of the area

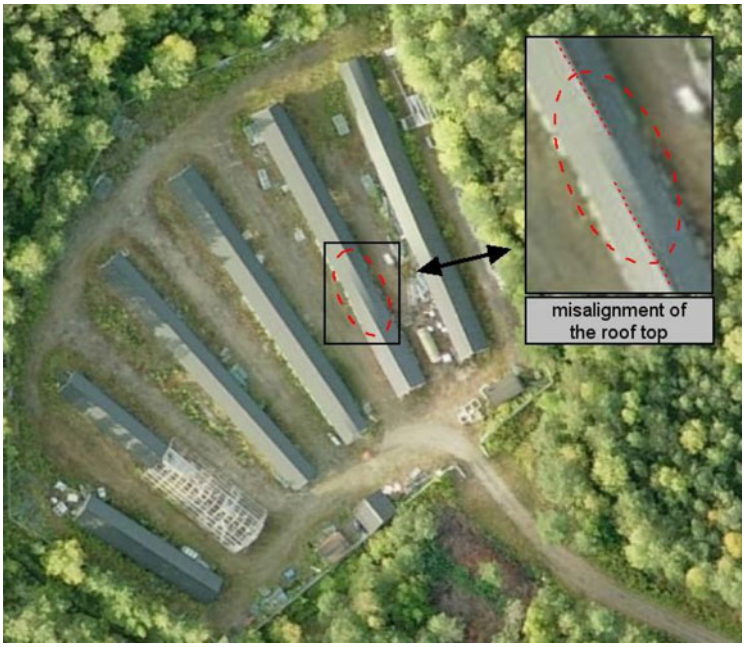


Fig. 3 Even professional map services are far from perfect. The *above image* shows an obvious bug in the coverage of the ELROB reconnaissance site by Microsoft Bing Maps

invariant feature transform (SIFT) [16, 17] or the Kanade-Lucas-Tomasi Feature Tracker (KLT) [22]. This holds for the area of Visual SLAM (VSLAM) in general [1, 2, 5, 6, 8, 9, 11, 13] and in particular for work dealing with UAV and photo mapping [1, 23].

In our approach, not only local descriptors but the whole information within the images is used, which seems to be very beneficial in the context of photo mapping as described in more detail in [3]. An variant of the Fourier Mellin Invariant (FMI) transform for image representation and processing [7, 21] is used with two modifications in comparison to previous work. First, a logarithmic representation of the spectral magnitude of the FMI descriptor is used. Second, a filter on the frequency where the shift is supposed to appear is applied. An analysis of the benefits of this improved FMI (iFMI) is provided in [3]. The iFMI provides 2D translations, rotations, and scaling, i.e., registrations when the UAV is moving horizontally, rotating its yaw, and changing its altitude. In addition, it requires in theory a strictly down looking camera and camera and a flat world hypothesis is assumed. But as shown in the experiments, these constraints do not have to be strictly achieve meaningful results.

We furthermore developed an uncertainty metric for iFMI registrations [20], which is up to our knowledge the first time that this was done for a spectral registration method. This allows to embedded the iFMI registration in SLAM, which further improves the mosaicking results as also shown in this article.

The rest of this article is structured as follows. Section 2 gives an overview of the iFMI and the related uncertainty metric used for image registration and SLAM.

Section 3 describes results from a reconnaissance and a surveillance trial at ELROB-2009. During these safety and security missions, the UAV operated fully autonomously for extended periods in time and generated several maps with iFMI registration. In Section 4, results from RREE-2010 are presented where the UAV engaged in a more rescue oriented response task, namely the inspection of a rubble pile. The augmentation of iFMI with an uncertainty estimation is used in this mission to embed the registration within pose-graph SLAM. It is shown that the resulting maps are improved. Section 5 concludes the article.

2 The Underlying Methods of the Photo-Mapping

2.1 The Improved Fourier Mellin Invariant Registration

The basis for the Improved Fourier Mellin Invariant (iFMI) registration [3, 4] used in this work is a Phase-Only Matched Filter (POMF). In this correlation approach, the signal shift is calculated based on the phase difference of the image spectra. If the POMF is applied to two periodically shifted signals, the result is an ideal Dirac pulse indicating the signal shift. This resulting Dirac pulse deteriorates with changing signal content of both signals. As long as the POMF yields a clear detectable maximum this method can be used for matching two signals.

The POMF can be used to detect both translation in the image content as well as rotation and scaling. Translation and rotation/scaling is decoupled because the 2D spectrum rotates and scales in the same way as the image in the time domain. However, the spectrum magnitude is invariant in translation. The spectrum magnitudes are resampled in polar coordinates using a logarithmic radial axis. A translation in this log-polar space signifies a rotation along the angular axis and a scaling along the radial axis. The logarithmic resampling is a way to compute the Mellin transform with a Fourier transform, hence the name [7, 21]. It is thus possible to use the 2D spectrum magnitudes of the images to register rotation and scaling parameters, and the image content itself to register translation.

This is used in the Fourier Mellin based mapping in a straightforward way. Algorithm 1 shows the calculation of the Phase Only Matched Filter, and Algorithm 2 details the iFMI registration method. The image registration is done incrementally, keeping track of the global transformation parameters from a global reference frame to the current image.

Algorithm 1: The Phase Only Matched Filter, later used as procedure $\Delta = \mathrm{POMF}(A, B)$

Data: Matrices A and B
Result: A matrix Δ containing a discretized Dirac pulse
$F_A = \mathrm{FFT}_{2D}(A)$;
$F_B = \mathrm{FFT}_{2D}(B)$;
$D = \mathrm{PHASE}(F_A) - \mathrm{PHASE}(F_B)$;
$\Delta = \mathrm{IFFT}_{2D}(\exp(iD))$;
return Δ;

Algorithm 2: The iFMI Registration Algorithm

Data: Grayscale Images A, B
Result: Transformation Parameters T: Translation x, y, Rotation θ, Scaling γ
`/* Calculate spectra */`
$S_A = \mathrm{FFT}_{2D}(A)$;
$S_B = \mathrm{FFT}_{2D}(B)$;
`/* resample magnitudes into polar coordinates with logarithmic radius */`
$P_A = \mathrm{LOGPOLAR}(|S_A|)$;
$P_B = \mathrm{LOGPOLAR}(|S_B|)$;
`/* apply POMF */`
$\Delta_{rot/scale} = \mathrm{POMF}(P_A, P_B)$;
`/* find dirac peak within scaling/rotation parameter space */`
$\{\theta^*, \gamma^*\} = \arg\max_{\theta,\gamma} \Delta_{rot/scale}(\theta, \gamma)$;
`/* rotate and scale second image into reference frame */`
$B' = \mathrm{ROTATE\_SCALE}(B, \theta^*, \gamma^*)$;
`/* apply POMF once more */`
$\Delta_{trans} = \mathrm{POMF}(A, B')$;
`/* find dirak peak within translation parameter space */`
$\{x^*, y^*\} = \arg\max_{x,y} \Delta_{trans}(x, y)$;
return $T(x^*, y^*, \theta^*, \gamma^*)$;

A first reference image I_0 is acquired or provided to define the reference frame. For each subsequent image I_n the image registration is performed with the preceding image I_{n-1} resulting in a transformation estimate T^n_{n-1}. Two subsequent transformations T^m_n and T^k_m can be combined as follows:

$$T^k_n = T^m_n \oplus T^k_m$$

with specific parameters:

$$\begin{aligned} x^k_n &= x^m_n + \gamma^m_n \left(\cos(\theta^m_n) x^k_m - \sin(\theta^m_n) y^k_m\right) \\ y^k_n &= y^m_n + \gamma^m_n \left(\sin(\theta^m_n) x^k_m + \cos(\theta^m_n) y^k_m\right) \\ \theta^k_n &= \theta^m_n + \theta^k_m \\ \gamma^k_n &= \gamma^m_n \gamma^k_m \end{aligned}$$

At each iteration, the new transformation estimate T^n_{n-1} is combined with the previous global estimate T^{n-1}_0 to form the new global estimate T^n_0 for the image I_n. This new global transformation is used to project image I_n into the global image map.

Several strategies are used to improve the registration. First of all, the search area within the FMI descriptor is restricted to reasonable parameters between two image frames to make the registration more stable. As a second measure the FMI descriptor is, in addition to a standard spectral window, processed by a logarithmic function in order to suppress low frequencies which often lead to erroneous registration results. Furthermore, the domain after the inverse Fourier transform is processed by a FIR Fourier transform is processed by a FIR interpolation filter to suppress interference from aliasing image content and which can be distributed around several frequency cells. A more detailed discussion of the approach can be found in [3].

2.2 Uncertainty Estimations for iFMI as Basis for SLAM

Please note that the iFMI described above is a pure registration method, i.e., it determines a spatial transform between images. We have recently extended iFMI [20] with an uncertainty analysis, which is up to our knowledge done for the first time for a spectral registration method. A covariance matrix is extracted from the result of a Phase-Only Matched Filter, which is interpreted as a probability mass function.

The method is embedded in a pose graph implementation for Simultaneous Localization and Mapping (SLAM) [14, 19]. The covariances matrices computed in Algorithm 3 are combined into one block diagonal matrix for the parameters x, y, θ. Scaling (γ) is encoded into a height parameter z, as it is directly proportional to it in the SLAM setting. Within the implementation, the covariances are reordered and scaled accordingly.

Algorithm 3: Uncertainty Analysis of iFMI Registration result as computed by algorithm 2

Data: Parameter Estimates x, y, θ, γ, POMF Results $\Delta_{rot/scale}, \Delta_{trans}$, and Interpolation Neighborhood N

Result: Covariance Matrices $C_{rot/scale}$ and C_{trans}, for parameters θ, γ and x, y, respectively

`/* normalize POMF Result histograms */`

$$\Delta_{rot/scale} = \Delta_{rot/scale} \cdot \left(\sum_i \sum_j \Delta_{rot/scale}(i,j) \right)^{-1};$$

$$\Delta_{trans} = \Delta_{trans} \cdot \left(\sum_i \sum_j \Delta_{trans}(i,j) \right)^{-1};$$

`/* fit covariance to histograms around estimates */`

$$C_{rot/scale} = \sum_{i=\theta-N}^{\theta+N} \sum_{j=\gamma-N}^{\gamma+N} \Delta_{rot/scale}(i,j) \cdot (i-\theta, j-\gamma)^T (i-\theta, j-\gamma);$$

$$C_{trans} = \sum_{i=x-N}^{x+N} \sum_{j=y-N}^{y+N} \Delta_{trans}(i,j) \cdot (i-x, j-y)^T (i-x, j-y);$$

return $C_{rot/scale}, C_{trans}$;

Each registration result gives rise to an edge in a pose graph, while vertices represent poses at which images were acquired. This representation is especially useful as it allows for easy manipulation at a later time. More integrated representations, such as the image mosaics shown in the next section, are easily reconstructable from the graph. Loops in the pose graph, i.e. substantial overlaps between non-consecutive image frames, are detected automatically by comparing current pose estimates of two nearby vertices. Once all loops are found, the TORO pose graph global optimization library [14] is used to reduce the overall map error.

We demonstrate the benefits of iFMI-based SLAM in [20] through results in the underwater domain, i.e., mosaicking of sea beds. The results presented here from ELROB-2009 are based on registration only, which already gives very useful results as can be seen later on in the related Sections 3.1 and 3.2. However, the field tests from RREE-2010 show that the usage of SLAM further improves the map quality. The according results are presented in Section 4.

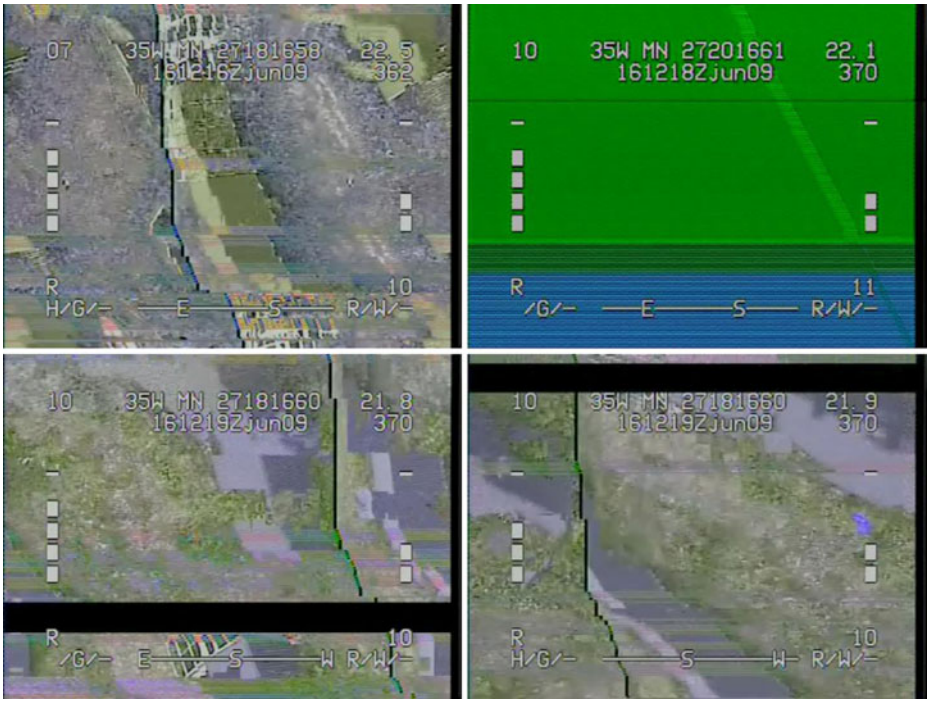


Fig. 4 The digital video link from the UAV had to cover up to 300 m through forest under humid weather conditions. There were occasional video errors despite using a high quality, military grade RF connection. The *above images* show some typical examples

3 Safety and Security Missions at ELROB-2009

As we will show in the following, the system generated useful photo-maps in real-time during the missions at ELROB-2009. The main limitation in the photo mapping were error frames in the video (Fig. 4), which could not be registered. When an error frame occurred, the generation of a new photo map was started. Table 1 shows the number of frames for the different maps that are presented later on. The photo mapping took place on a laptop with a Core-2 Duo 2.0 GHz CPU running Ubuntu 9.04 Linux, which received a copy of the video stream from the UAV.

As discussed in a bit more detail later on, the UAV performed in the field trials at ELROB-2009 significant amounts of time in autonomous mode (Table 2) despite challenging weather conditions. The operator never used line-of-sight control, i.e., he never looked directly at the vehicle, but he had to rely only on the

Table 1 Number of frames in the example maps

	Reconnaissance			Camp security	
Figures	6	7	8	11	12
#frames	31	52	87	337	784

Table 2 The autonomous flight times of the UAV

Scenario	Autonomy (mm:ss)	Overall (mm:ss)
Reconnaissance	11:00	22:00
Camp security	20:00	57:30

Please note that the overall mission times include the set up and the time for the provision of the mission deliverables like maps, photos of the intruders, etc. to the organizers

video/map-information. This is was explicitly required at ELROB and it is a realistic constraint for safety and security missions where the operator tends to be relatively far from the controlled vehicle itself.

3.1 Reconnaissance at ELROB-2009

The mission task for the reconnaissance mission was to detect hazardous spots in an Nuclear-Biological-Chemical (NBC) incident scenario. The hazardous spots were marked by Emergency Response Intervention Cards or short ERI-cards. ERI-cards are rectangular plates of about 40 cm × 30 cm, with black number codes for different hazardous materials on orange background.

The main site for the reconnaissance operations was about 360 m away from the operator's station. The site itself spans about 60 m by 60 m (Fig. 5). The UAV used GPS way points to autonomously reach the site and to explore it.

The UAV operated fully autonomously for 11 min in total out of 22 min mission time (Table 2). Please note that the mission time includes set up and the provision

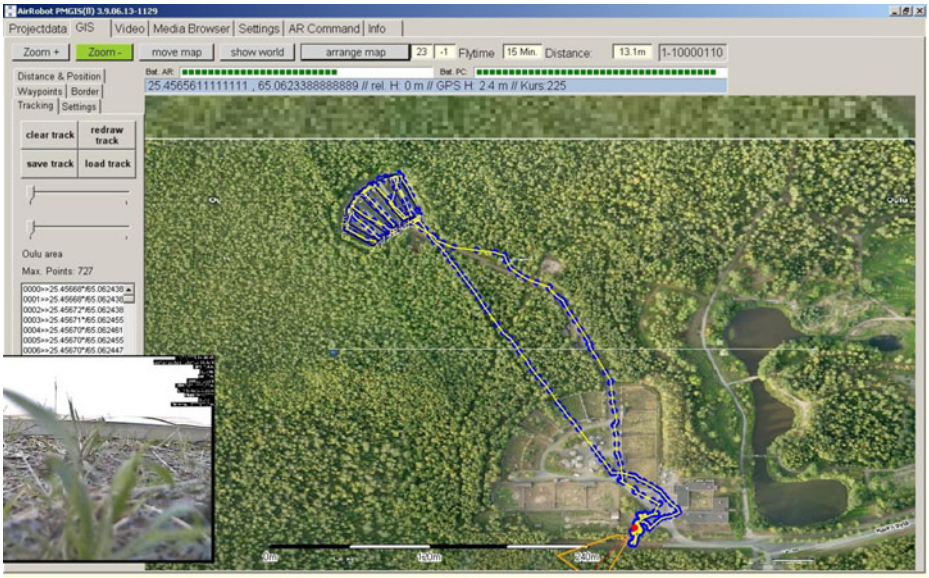


Fig. 5 The waypoints in the reconnaissance path

Fig. 6 An example of a photo map in the reconnaissance scenario (31 frames)

of mission deliverables to the organizers. There were only about 4 min of manual intervention during the actual flight time of about 15 min. The manual operations included start and landing as well as the upload of a second set of preplanned GPS way points to pursue a different search strategy. The vehicle was capable of

Fig. 7 A second example of a photo map in the reconnaissance scenario (52 frames)

autonomous operation under difficult wind conditions; wind gusts in the air exceeded 12 m/sec. The UAV did not detect any of the ERI-cards. But also five out of the six other teams—all of them using land robots—did not find any ERI-cards either. Only one team out of the seven teams managed to find 2 ERI-cards.

The mission scenario involved significant communication challenges for the video transmission. The maximum distance was about 400 m with significant amounts

of vegetation, namely forest in humid weather conditions, between the UAV and the operator's station. There were hence quite often error frames in the video data stream. As mentioned, we are currently working on an automatic detection of error frames to prevent a disruption of the map generation. Typical error-free video sequences consisted of about 75 frames. Figures 6 and 7 show two examples with 31, respectively 52 frames. It has to be noted that the approach also gives reasonable results when the camera is tilted though it should fail in theory under such conditions. Figure 8 shows an according example based on 87 frames. The maps were generated in real time at the operator's station with about 50 m/sec processing time for registration and visualization, i.e., 20 Hz update rate.

Fig. 8 An example of a photo map of the way of the UAV to the reconnaissance site (87 frames). Please note that there is a significant tilt of the camera and hence skew in the images. Though the algorithm is in theory not capable of registering under these conditions, the result is nevertheless quite usable of a rough orientation

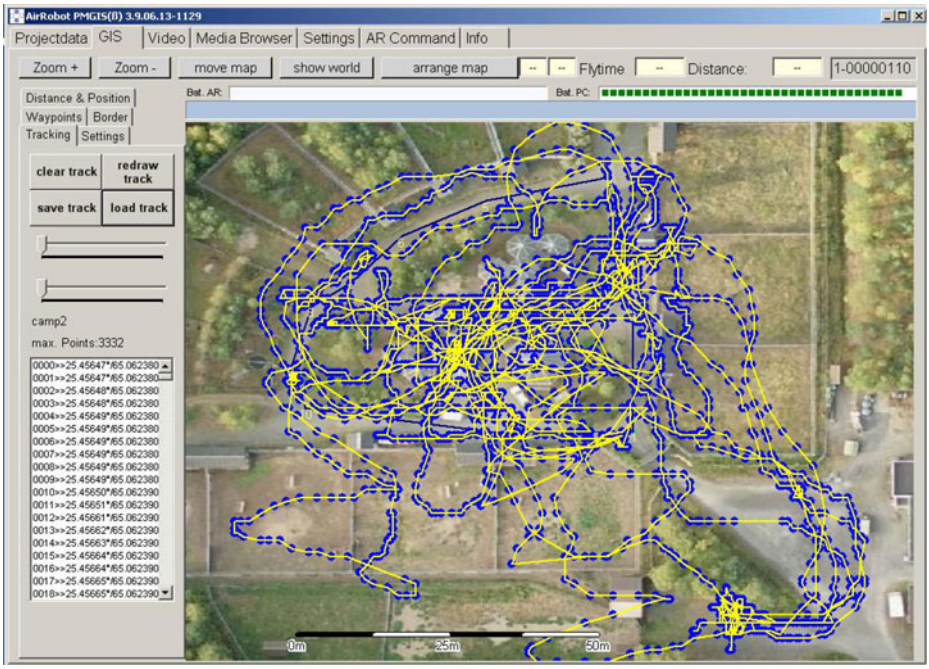


Fig. 9 The actual path of the UAV in the camp security scenario based on GPS track points

3.2 Camp Security at ELROB-2009

The mission task for the camp security was to detect intruders. The operator's station was located about 30 m from the camp site, which spans about 90 m by 90 m (Fig. 9). The UAV mainly autonomously hovered over the site and intruders were "chased" in manual mode by the operator when being detected. As mentioned before, the operator never used line-of-sight control, i.e., he always used only the data at the operator's station, which was inside a tent. The overall mission time was 60 min including set up and the provision of mission deliverables in the end.

In total, six intruders out of eight possible ones were detected and "caught" by the UAV. Two examples of detected intruders are shown in Fig. 10. Only a 2nd team out of the seven teams in total—all other six teams using land robots—was also able to catch six intruders. One other team caught 2, three teams caught 1, and one team caught 0 intruders. The two intruders that were missed entered the camp while the UAV was for a 5 min stop at the operator's station for swapping the batteries—as mentioned before, the maximum operation time of the UAV is about 25 to 30 min, i.e., batteries had to be changed once during this mission.

As shown in Table 2, the UAV operated fully autonomously for 20 min in total out of 57.5 min flight time, i.e., for about 34.8% of the time. The wind conditions were also quite challenging during this mission, the wind speed on ground was measured with up to 7.2 m/sec, wind gusts in the air exceeded 12 m/sec.

The mission scenario involved shorter communication ranges, less than 150 m maximum, and fewer vegetation, especially fewer fewer trees. But the "camp"—

Fig. 10 Two examples of intruders detected during the camp security mission

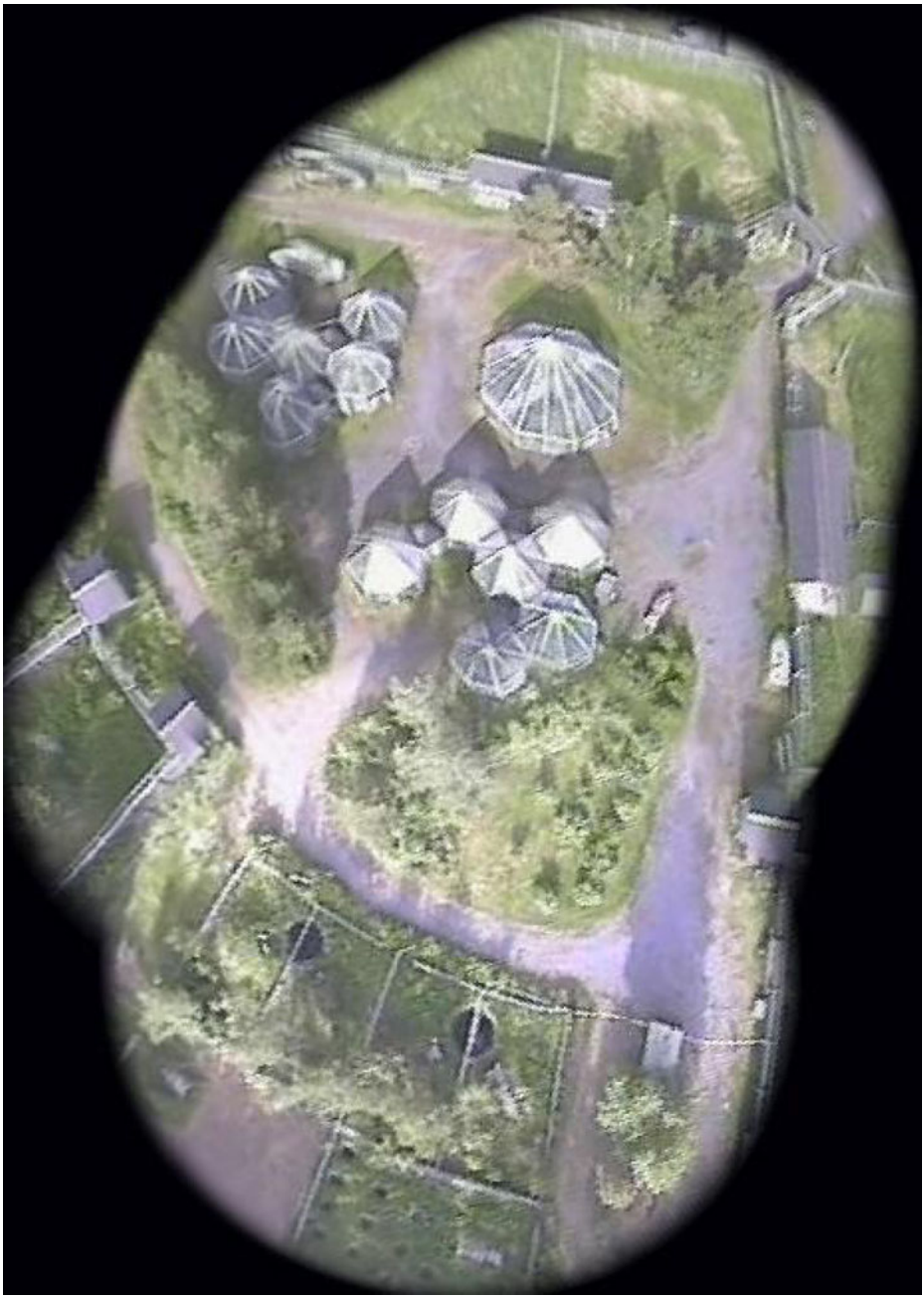


Fig. 11 An example of a photo map of the camp generated from 337 video frames in real time

being a part of the former Oulu Zoo—contains a significant amount of metal fences and cages with wire meshes. Nevertheless, the video link was much less error prone than for the reconnaissance mission. The video sequences between error frames were

hence significantly longer. On average, video sequences with 650 frames were turned into photo maps. Figures 11 and 12 show two examples with 337, respectively 782 frames. The maps were generated in real time at the operator's station with about 50 m/sec processing time for registration and visualization, i.e., 20 Hz update rate.

4 An Example for a Rescue Mission: Mapping a Rubble Pile at RREE-2010

The ELROB missions presented in the previous section required—quite typically for safety and security scenarios—a significant distance between the operator and the vehicle and hence quite some amounts of full autonomy or at least some skilled operation without line-of-sight control. The RREE mission presented in this section is in contrast quite typical for a rescue mission. The operator of the UAV can get

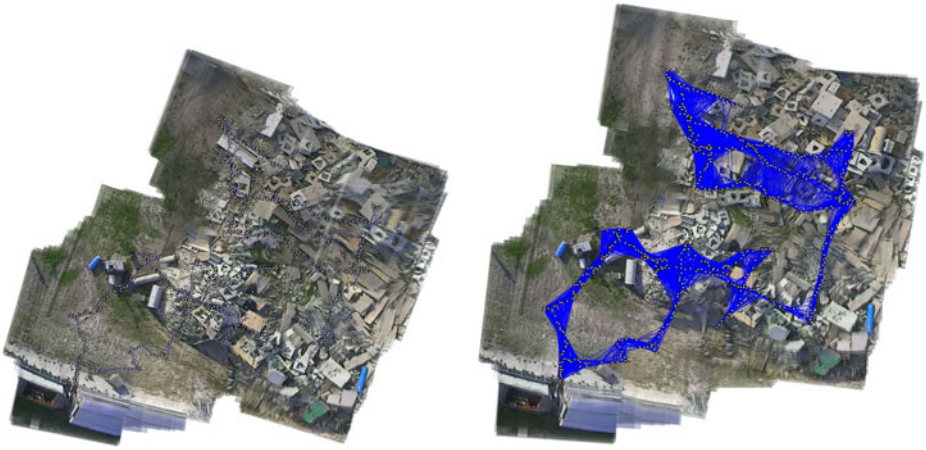
Fig. 12 A second example of a photo map of the camp generated from 784 video frames in real time

relatively close to the incident site and manual flight with eyes-on control, i.e., direct line-of-sight to the UAV, is possible.

The RREE scenario presented here is a very good example for a possible usage of UAV in rescue missions, namely the generation of an overview map of a rubble pile. Rubble piles are potentially very dangerous to physically access, both for responders as well as for the victims trapped inside of the pile. All operations have therefore to be carefully planned and a bird's eye overview generated from aerial video is a very good basis for this.

In the context of this article the most interesting aspect of the rubble pile map shown in Fig. 13 is that it has been generated—in contrast to the ELROB-2009 mosaics—with full SLAM and not only with sequential registrations. Concretely, 630 video frames were taken. Figure 14 shows on the left the path of the robot and the map generated on the left the path of the robot and the map generated by sequential registrations. On the right in the same figure, path and of edges that have been

Fig. 13 A map of a rubble pile at Disaster City generated at RREE-2010 from 630 video frames via iFMI-SLAM



(a) sequential only (b) including loop closures

Fig. 14 The path of the UAV over the rubble pile as pose-graph, once without (*left*) and once with loop closures edges for SLAM (*right*). The *little triangles* show the 630 places where a photo was taken.

generated by loop closures, i.e., the detection of non-sequential image pairs that can be registered with each other.

The loop closures lead to a significant increase in the number of edges, namely from the 629 sequential ones that represent the path of the robot to a graph with 12,364 edges. All edges are supplemented by uncertainty information, i.e., a covariance matrix that is extracted from the result of the phase-only matched filters

(a) sequential only (b) SLAM

Fig. 15 A detail of the rubble pile map, once generated with sequential registration only (*left*) and once embedded in posegraph SLAM (*right*)

in the iFMI registrations. The pose-graph is relaxed using the open source TORO optimizer [14]. As shown in Fig. 15, the SLAM improves the map quality.

The computation of the uncertainty information hardly influences the runtime of the registration at all, i.e., it can still be done in real-time with 20–30 Hz. The optimization of the pose-graph is quite fast, too. It takes 2.12 s in the example with the 12,364 edges. Please note that previously computed transformations on the images can be reused, hence speeding up the registrations. So while the UAV is flying, the iFMI already generates a reasonable photo map. Once the mission is finished or a higher quality overview is required, a loop closing and optimization process can be triggered and an improved map is generated after a few seconds via SLAM.

5 Conclusions

Several missions with an Unmanned Aerial Vehicle (UAV) in different realistic safety, security, and rescue exercises were presented. The missions can be considered to be prototypical examples for the usage of UAV in the different types of response operations. An important element in the presented UAV operations is the real-time generation of photo maps to aid the operator and to serve as mission deliverables. An improved version of Fourier Mellin based registration with two core modifications is used for this purpose. First, a logarithmic representation of the spectral magnitude of the FMI descriptor is used. Second, a filter on the frequency where the shift is supposed to appear is applied. Furthermore, the iFMI registration is augmented by an uncertainty metric. This allows to embed the registration into a pose-graph for Simultaneous Localization and Mapping (SLAM). The resulting photo maps provide high resolution, up to date information about the mission site and about targets found within them like intruders.

The experimental results come from two realistic field trials covering all three types of SSRR operations. First, two safety and security missions with a UAV at the 2009 European Land Robot Trials (ELROB) were presented. An Airrobot AR100-B was used in a reconnaissance and in a camp surveillance scenario. The UAV is capable of GPS based way point navigation, which was used for significant amounts of autonomous operations despite very challenging weather conditions with gusts of wind with top speeds of more than 12 m/sec. A digital video link was used for photo mapping in real time. Despite significant deviations from the ideal theoretical conditions, i.e., the assumption of a flat world and of a perfectly down-looking camera, useful maps are generated. Second, results from a rescue scenario at the 2010 Response Robot Evaluation Exercises (RREE) at Disaster City, Texas were presented. Concretely, a manually flown UAV was used to generate an overview map of a rubble pile. The iFMI registration for the aerial mosaicking is supplemented in this mission by an uncertainty metric for SLAM, which further enhances the photo map.

References

1. Angeli, A., Filliat, D., Doncieux, S., Meyer, J.A.: 2d simultaneous localization and mapping for micro aerial vehicles. In: Proceedings of the European Micro Aerial Vehicles (EMAV 2006) Conference (2006)

2. Broida, T.J., Chandrashekar, S., Chellappa, R.: Recursive estimation of 3d motion from a monocular image sequence. IEEE Trans. Aerosp. Electron. Syst. **26**, 639–656 (1990)
3. Buelow, H., Birk, A.: Fast and robust photomapping with an Unmanned Aerial Vehicle (UAV). In: International Conference on Intelligent Robots and Systems (IROS). IEEE (2009)
4. Buelow, H., Birk, A., Unnithan, V.: Online generation of an underwater photo map with improved Fourier Mellin based registration. In: IEEE OCEANS. IEEE (2009)
5. Castellanos, J., Martinez, J., Neira, J., Tardos, J.: Simultaneous map building and localization for mobile robots: a multisensor fusion approach. In: Martinez, J. (ed.) Proceedings IEEE International Conference on Robotics and Automation, vol. 2, pp. 1244–1249 (1998)
6. Castellanos, J., Neira, J., Tardos, J.: Multisensor fusion for simultaneous localization and map building. IEEE Trans. Robot. Autom. **17**(6), 908–914 (2001)
7. Chen, Q.S., Defrise, M., Deconinck, F.: Symmetric phase-only matched filtering of Fourier-Mellin transforms for image registration and recognition. IEEE Trans. Pattern Anal. Mach. Intell. **16**(12), 1156–1168 (1994)
8. Davison, A.J.: Real-time simultaneous localisation and mapping with a single camera. In: Proceedings of the Ninth IEEE International Conference on Computer Vision, pp. 1403–1410 (2003)
9. Davison, A.J., Reid, I.D., Molton, N.D., Stasse, O.: Monoslam: real-time single camera slam. IEEE Trans. Pattern Anal. Mach. Intell. **29**, 1052–1067 (2007)
10. ELROB: European Land-Robot Trial (ELROB). www.elrob2009.org (2009)
11. Eustice, R.M., Singh, H., Leonard, J.J.: Exactly sparse delayed-state filters for view-based SLAM. IEEE Trans. Robot. Autom. [see also IEEE Transactions on Robotics and Automation] **22**(6), 1100–1114 (2006)
12. Google: Google Maps. http://maps.google.com (2009)
13. Grandjean, P., Robert De Saint Vincent, A.: 3-D modeling of indoor scenes by fusion of noisy range and stereo data. In: Robert De Saint Vincent, A. (ed.) Proceedings of the IEEE International Conference on Robotics and Automation, vol. 2, pp. 681–687 (1989)
14. Grisetti, G., Stachniss, C., Grzonka, S., Burgard, W.: A tree parameterization for efficiently computing maximum likelihood maps using gradient descent. In: Robotics: Science and Systems (RSS). Atlanta, GA (2007)
15. Grocholsky, B., Keller, J., Kumar, V., Pappas, G.: Cooperative air and ground surveillance. IEEE Robot. Autom. Mag. **13**(3), 16–25 (2006)
16. Lowe, D.G.: Object recognition from local scale-invariant features. In: Proceedings of International Conference on Computer Vision, pp. 1150–1157 (1999)
17. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
18. Microsoft: Bing Maps. http://www.bing.com/maps/ (2009)
19. Olson, E., Leonard, J., Teller, S.: Fast iterative alignment of pose graphs with poor initial estimates. In: Leonard, J. (ed.) Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006, pp. 2262–2269 (2006)
20. Pfingsthorn, M., Schwertfeger, S., Buelow, H., Birk, A.: Maximum likelihood mapping with spectral image registration. In: IEEE International Conference on Robotics and Automation (ICRA). IEEE (2010)
21. Reddy, B., Chatterji, B.: An FFT-based technique for translation, rotation, and scale-invariant image registration. IEEE Trans. Image Process. **5**(8), 1266–1271 (1996)
22. Shi, J., Tomasi, C.: Good features to track. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR94) (1994)
23. Steder, B., Grisetti, G., Stachniss, C., Burgard, W.: Visual SLAM for flying vehicles. IEEE Trans. Robot. Autom. **24**(5), 1088–1093 (2008)
24. TEEX: Disaster City. http://www.teex.com/teex.cfm?pageid=USARprog&area=USAR&templateid=111 7 (2008)