Quentin Ladetto et al.

3D Continuous Positioning for Security Applications

Introduction

In a security, police or military context, the localisation of one's own personnel and assets, as well as the ones of the adversary, has become a crucial aspect and plays a key role in almost any intervention. Whether the goal is to direct people and systems towards a common objective or to avoid fratricide shooting, real-time geopositioning enables a quick overview of the situation on the field, facilitating a well-informed and effective decision process.

Defining areas of activities for troops or police patrols is not sufficient anymore in a world where information and people travel quickly and dynamically. For example, effectively engaging patrols in areas of responsibility of 25 km² (urban), 100 km² (semi-urban), or 400 km² (rural) to respond to a robbery lasting only a few seconds is much easier (especially to organize checkpoints), if the precise position of all patrols is known beforehand and in real-time. In the same vein, the optimal positioning of troops in urban areas can improve the occupation of terrain, while avoiding a series of annoyances such as friendly fire or traffic.

In the age of high-bandwidth communication technology it is becoming easier to find and to provide information on a wide range of topics. This also applies to anything related to the geolocation of people and objects. In addition to their location on a map, which is in itself very interesting, it is now possible to provide contextual information related to that position in real-time. Using state-of-the-art RFID readers, one can nowadays not only identify personnel or objects, but also determine their status and locations within a zone. This combination of localization and identification provides so-called "geo-identification" information.

It is well known that in forests, tunnels, urban canyons and other caves GPS (Global Positioning System) signal reception is degraded, which poses a significant problem for security operations increasingly taking place in these environments. While geolocation is broadly associated with GPS and its Russian, European and Chinese counterparts, it is clear that major technological advances were achieved in recent years offering the possibility of geo-positioning of individuals and objects in a continuous manner, outside as well as inside buildings, in all terrain and regardless of received satellite signal strength. Modern algorithms combine measurements from cameras and inertial sensors to enable not only precise localisation in almost any environment, but also a three dimensional reconstruction of the surrounding environment. Current active RFID transponders or tags can then provide additional information based on sensors such as motion, temperature, gas, smoke, etc. thus increasing the security level of the system.

All information is useless unless it is presented in an intuitive way. Only this enables an effective decision process under pressure. Thus once the position of people and objects of interest is known, this information along with other parameters is displayed on a two or three dimensional map in a 'Virtual Globe'. As important as reliable and accurate geolocation is a detailed representation of the environment. We show how dense, accurate, two or three dimensional maps of the environment can be acquired apriori using existing maps or LiDAR (Light Detection and Ranging) and CMOS (Complementary Metal–Oxide–Semiconductor) camera sensors. We also outline a method for creating such maps in real-time, relying only on low-cost, easily portable sensors while requiring no infrastructure nor any apriori knowledge. Furthermore, we provide a detailed comparison of the accuracy and density of the obtained maps.

In that given context, Armasuisse Science and Technologies has launched the development of a demonstrator combining all the necessary technologies to provide, access and transmit the following information:

- The exact and continuous 3D location and orientation of people in outdoor as well as GPS-denied environments
- Representation of the information in a three-dimensional model created on-line
- Real-time transmission and visualization of all information

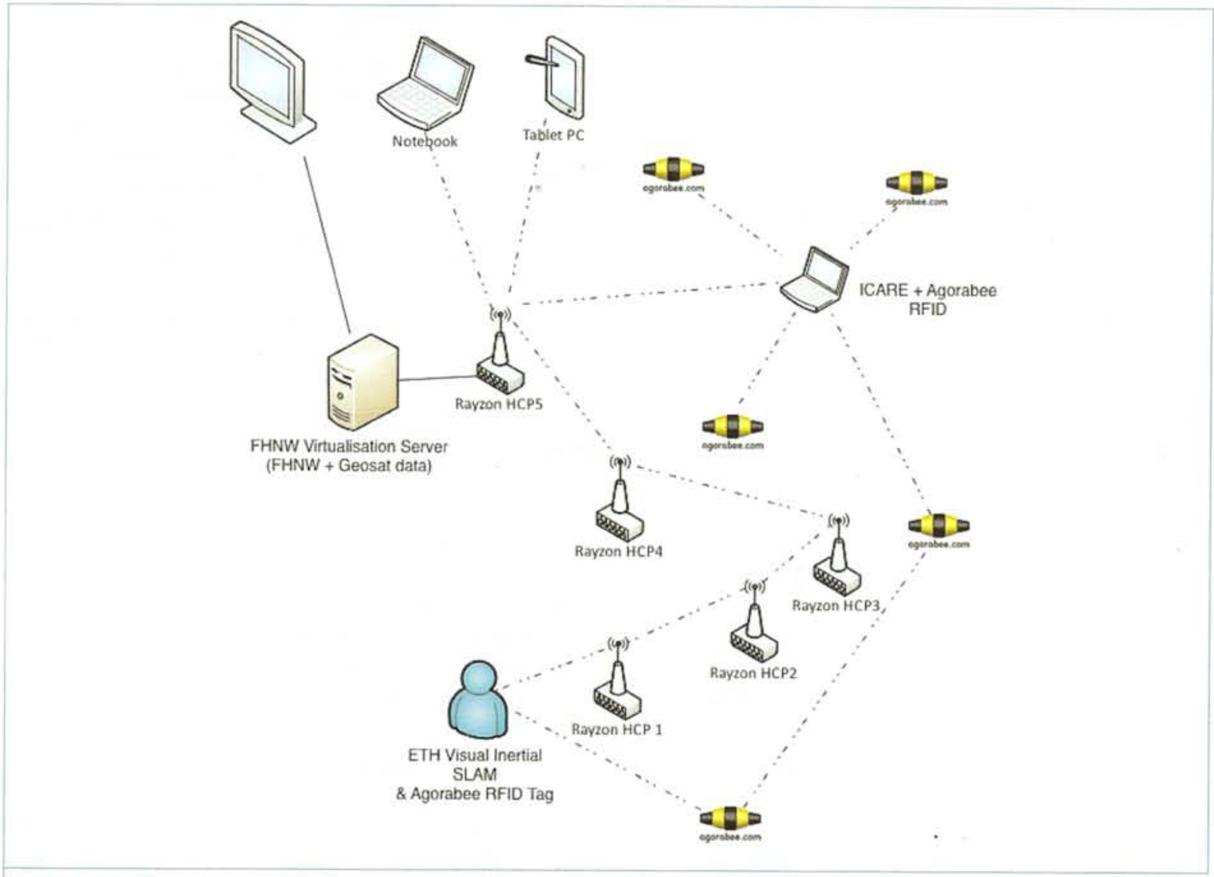


Figure 1 : Overview of the system architecture integrating the Visual-Inertial SLAM sensor reporting data over the Rayzon multi-hop network to the visualisation server. The position and friend-foe identification of the person is also identified via active RFID tags to the same server. The visualisation is then available in real-time on a fixed station, a notebook or a tablet PC.

- Identification of people through a series of readers / RFID tags enabling calibration of continuous geo-referencing systems
- Preview of the upcoming compartments of terrain, buildings or rooms
- Continuous observation of the engagement area in real-time
- Real-time guidance within complex environments

A dedicated demonstrator that fulfills these requirements was developed. Figure 1 gives an overview of all the subsystems involved.

1. The 3D environment is reconstructed using laser and stereovision.
2. The data is visualized in a 3D virtual globe enabling real-time display of the position and orientation of personnel as well as real-time mapping of the environment.
3. Localisation of personnel is based on a visual-inertial SLAM (Simultaneous Localization and Mapping) system, optionally supported by RFID technology for friend-foe detection and global position updates.
4. Real-time data transmission is secured by a multi-hop network which relays all the information to a centralised server.
5. An on-line video stream provided by the head mounted cameras is transmitted in parallel and enables viewing the scene in the Virtual Globe as seen by the person being tracked

This article presents the different components of the system, the results obtained as well as the future work and improvements to be done.

3D Terrestrial Laser scanning

This technology was used to create the 3D model of the environment prior to the live demonstration. The registered point cloud was integrated into the Virtual

Globe for 3D-representation of the building (indoor and outdoor). As this method provides very accurate and comprehensive 3D data, it was considered as the reference for the quality analysis of the two other 3D mapping methods used in the framework of the project, namely the Visual-Inertial SLAM and the Stereovision Mobile Mapping.

Hardware

The instrument used for the acquisition of the 3D point cloud of the building is a Riegl-VZ400 (Figure 2) with integrated compass, inclination sensor, GPS and high-precision mounting pads with digital camera (Nikon D700). The laser scanner was operated in stand-alone mode using the in-built user interface of the front panel. It has multi-target echoes capabilities, a line scanning mechanism based upon a fast rotating multi-facet polygonal mirror providing fully linear, unidirectional and parallel scan lines. The maximum measurement rate is 122'000 points/seconds with a range of 0.5m to 600m, and a precision of ±3mm (repeatability). Simultaneously to the laser acquisition, the digital camera shot pictures of the scanned environment with a resolution of 12.1 MPixel.

Acquisition and processing

The acquisition phase for the 3D mapping of the Armasuisse building in Thun was performed within six hours, by 30 laser stations operated in static mode, each of them lasting around three minutes (including image acquisition). The raw data based on the acquisition consists of 30 different point clouds that need to be processed, filtered and assembled together in the official Swiss coordinate system. The resolution of the scans was set to 0.06 degree, which results in one point each millimeter at a distance of 10 meters.

In order to assemble and register the 30 scans together, we used inter-scan targets (six per scan) visible at least on two successive scans. Few targets were accurately known in the Swiss coordinates system, allowing a global referencing of the assembled point cloud in the Swiss grid system.

The complete outdoor environment of the building was scanned (Figure 3), including the direct surroundings, each façade and the roof of the building itself (15 stations). The indoor environment was partly scanned (Figure 4). We chose to acquire the ground floor and the fourth floor (15 stations).

The processing phase consisted in importing the different point clouds in the RiScan Pro software of Riegl, applying few filters to eliminate wrong or unnecessary points, assembling and referencing the point cloud prior to import it into the Virtual Globe. The register-

Figure 2 : Riegl VZ-400 laser scanner with Nikon D700 digital camera

Figure 3 : View of the processed outdoor point cloud, Armasuisse building, Thun.

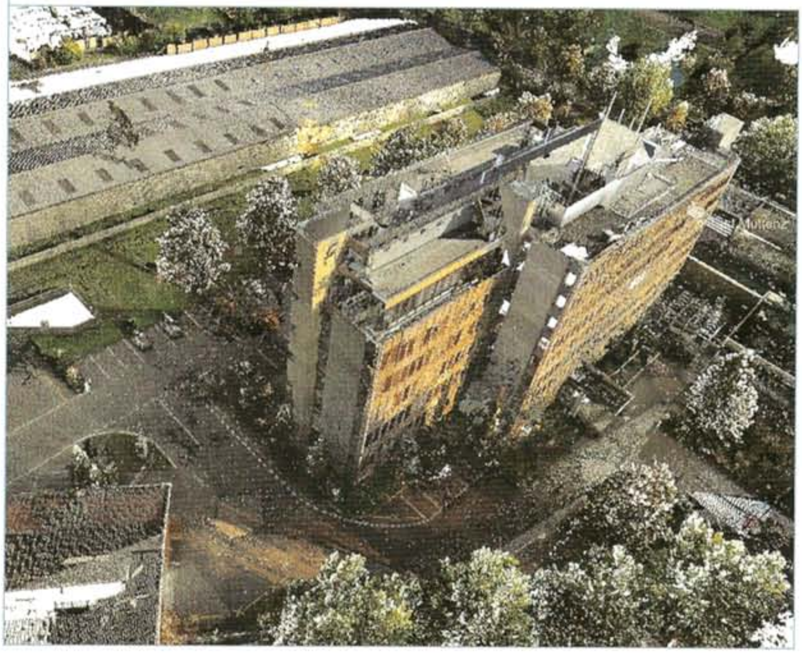


Figure 4 : View of the processed indoor point cloud, 4th floor - Armasuisse building, Thun.

ing procedure within RiScan Pro gave an estimated global accuracy of 8 mm for the final point cloud.

This set of data has been considered as the reference for the quality analysis of the two other 3D mapping methods. The result of this comparison is given in the next chapter.

Comparison of the 3D mapping methods

In the scope of the demonstrator, three different 3D mapping methods were employed and their results imported in the Virtual Globe:

- ☐ 3D Laser Scanning (LS)
- ☐ Stereovision Mobile Mapping (SMM)
- ☐ Visual-Inertial SLAM (VI-SLAM)

In order to get a comparison and analysis of these methods, Geosat imported the point clouds of the other two mapping methods listed above, and described further in this article, into the processing software of Riegl. In that way these methods have been compared to the reference data set made by laser scanning.

Laserscan vs Stereovision Mobile Mapping

The stereovision mobile mapping (SMM) offers the advantage of a mobile acquisition. But it has the disadvantage of being mounted on the roof of a car; data can then only be acquired if the environment of interest is accessible by road. The accuracy of the SMM method is very good in horizontal acquisition procedure, at short distances (< 20m) and with accurate trajectory data (5-10 cm). But for the surveying of the Armasuisse building the camera must be mounted obliquely in order to get the top of the façades. Using oblique acquisition procedures, data becomes noisier and less accurate (50-80 cm). Hereafter a chart comparing the main features of the LS and SMM methods:

3D Laser Scanning (LS)

- ☐ Static system, on a tripod
- ☐ Indoor and outdoor
- ☐ Accuracy ± 0.8cm
- ☐ Very sharp and accurate data
- ☐ Post-processing required

Stereovision Mobile Mapping (SMM)

- ☐ Mobile system, roof mounted on a car
- ☐ Outdoor only
- ☐ Max. accuracy in good acquisition conditions (horizontal shooting, accurate trajectory) ± 10cm
- ☐ Noisy data at high elevation (50 - 80 cm)
- ☐ Post-processing requested

Laserscan vs Visual-Inertial SLAM

Visual Inertial SLAM offers the main advantage of a mobile acquisition system without any post-processing. But the accuracy of the data is directly linked to the accuracy of the trajectory. In indoor environments the accuracy is around 0.5-5 meters, depending on the performance of the inertial platform, decreasing in time. One other disadvantage is the resolution of data, which is non-homogenous and varying with the movement of the sensor. Hereafter a chart comparing the main features of the LS and SMM methods:

3D Laser Scanning (LS)

- ☐ Static system, on tripod
- ☐ Heavy ≈ 12kg
- ☐ Post-processing required
- ☐ Accuracy ± 0.8 cm
- ☐ Regular point density 1pt/cm
- ☐ Very sharp and accurate data

Visual inertial SLAM (VI-SLAM)

- ☐ Mobile system, head-mounted on helmet
- ☐ Light, hand-held
- ☐ Real time acquisition
- ☐ Accuracy ± 2 m
- ☐ Irregular point density
- ☐ Noisy data

The OpenWebGlobe Virtual Globe Technology

Motivation

In recent years, Virtual Globes became an important tool for interactively visualizing and investigating three-dimensional geospatial data covering large areas. Virtual globes are capable to stream the huge necessary amounts of imagery, elevation data and other geospatial contents over the Internet. Emerging Internet technologies, such as HTML5 and WebGL (Khronos Group), offer new possibilities to develop 3d virtual globes running in web-browsers without a need for browser extensions (plugins).

The OpenWebGlobe Project and Technology

The OpenWebGlobe SDK (Software Development Kit) is an open source virtual globe environment using WebGL (www.openwebglobe.org). Unlike other (web-based) 3d geovisualisation technologies and toolkits, the OpenWebGlobe SDK not only supports the content authoring and web visualization aspects, but also the geospatial data processing functionality for generating multi-terabyte terrain, image, map and 3d point cloud data sets in high-performance and cloud-based parallel computing environments.

Virtual globes typically comprise geospatial base data in the order of TeraBytes to PetaBytes. In order to keep the downloading time as short as possible, only currently visible data will be downloaded to the client. So there is a need to preprocess bulk image, elevation and point cloud data by splitting it up into small, streamable data fragments at different resolutions (Christen M. and al, 2011). There is a preprocessing part within the OpenWebGlobe SDK which provides this functionality for imagery and elevation data. The processing SDK is supporting the highly parallel processing of multi-terabyte data sets on high-performance computing (HPC) architectures and on cloud-based computing platforms.

On the other side there is the viewer part of the OpenWebGlobe SDK. The viewer part is written in JavaScript and offers more than 100 different functions for the integration of geospatial content into a virtual globe application like different image layers, elevation layers, POIs (Point of Interests), textured 3d models, 3d point cloud data, and others. The viewer part is also intended to build rich 3d Web GIS applications using JavaScript and HTML5.

OpenWebGlobe in the Continuous Location Demonstrator: Dynamic Point Cloud Streaming

The OpenWebGlobe SDK was used as a base technology for the Continuous Location Demonstrator application. The application offers possibilities to integrate and compare the different 3d point clouds captured from laser scanning or stereo vision mobile mapping systems. Bulk 3d point cloud data was preprocessed prior to its integration into OpenWebGlobe. Data from the ETH stereo vision SLAM device such as point cloud fragments, position and orientation data is continuously loaded and displayed by the application.

Figure 5 shows a Screenshot of the application with a point cloud of the building exterior in an overview. Different predefined or dynamic view positions allow the user to investigate the building in more detail or to dynamically follow and observe moving objects or targets. Viewing modes like the 'follow mode' in which the user sees the virtual globe in the view of a individual or the 'third person mode' which allows observing a scene from a predefined distance and automatically follows any moving target are other useful implemented features. In Figure 6 a semi-transparent view frustum shows the current location and viewing direction of a individual and the point cloud fragments captured along his path.

As a part of the Continuous Location Demonstrator project, we implemented a scalable point cloud support into the OpenWebGlobe SDK. For this purpose the viewer part was extended by a set of functions to load, display, show and hide point cloud data. On the preprocessing side additional algorithms for coordinate transformations, discretization and the thinning-out of large point clouds had to be implemented.

In the Continuous Location Demonstrator Project OpenWebGlobe was installed on a dedicated visualization web server. Client devices like normal laptops, tablets or mobile phones can easily connect to the web server using an onboard web browser in which the interactive 3d application page is loaded and executed.

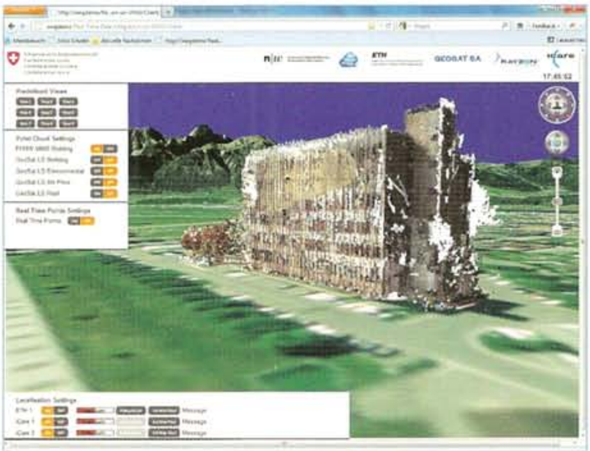


Figure 5 : Continuous Location Demonstrator application with point cloud of building acquired by stereo vision mobile mapping system

Figure 6 : Individual view frustum and dynamically captured point cloud fragments

An additional effort was made to implement the possibility to display real time data. There were two kinds of real-time data which were loaded by the application. First, the position and orientation information of an individual's head mount device, with an update rate of around 10 Hz. For this communication we used simple UDP messages which were sent from the individual's device to our visualization server. Second, the captured point cloud fragments (around 1000 points per fragment) from a individual's device, were to be loaded onto the viewer with an update rate of 3 seconds. For this purpose, the point cloud fragments were transferred as ASCII-files. These files were automatically copied to the visualization server and continuously loaded into the application.

Stereovision Mobile Mapping

Motivation

One of the aims of the Continuous Location Demonstrator Project was to establish and investigate the structure of an unknown building as quickly as possible. While explicit geometric 3d modelling based on image or laser scanning data still requires a considerable amount of human intervention, point cloud based 3d models provide a very powerful solution including rapid acquisition, rich details, free navigability and highly automated production (Nebiker S. and al, 2010). As shown in the earlier section on laser scanning, 3D point clouds have mainly been acquired using static or kinematic laser scanning systems. However, recent progress in sensor technology and image matching algorithms have made stereovision based dense point cloud generation an interesting alternative. For this reason, we used a stereovision mobile mapping solution for dynamically capturing the imagery of the building and for the subsequent extraction of point-cloud data.

Stereovision Acquisition and Processing Technology

The mobile mapping system consists of a Navigation System POS LV 210 from Applanix and two or more stereo camera systems, which are mounted on a vehicle. The system captures accurately georeferenced stereo image pairs, which can be used to survey any object-seen in the images. The cameras allow capturing up to 32 images per second in different resolutions up to 11 megapixels. They are triggered synchronously and referenced to GPS time. These georeferenced stereo image pairs are typically used to generate stereo imagery and 3d videos allowing interactive stereoscopic 3d measurements. It is also possible to automatically extract fully coloured dense point clouds by using dense image matching algorithms. The absolute 3d point measuring accuracies is around 5 cm horizontally and about 3 cm vertically. The relative 3d accuracy is around 1 cm for points measured within a stereo frame. The geometric accuracy of derived points clouds are approximately 5-10 cm for ranges up to 20 m.

Figure 7 shows the stereo vision mobile mapping system with cameras mounted for road investigation.

Rapid Acquisition of 3D Building Data

The stereovision mobile mapping system is typically used for capturing the 3d roadside environment. The investigations in this project were done in order to

Figure 7 : Stereo Vision Mobile Mapping System

demonstrate the feasibility and quality of capturing the geometry of multi-storey buildings with a stereo-vision based mobile mapping system. Because of the building height it was necessary to mount the two stereo camera systems onto the same side of the car and oblique, see Figure 8.

Images of the entire building were captured by driving around the building twice. The image capturing part was finished within minutes. After the capturing we generated point cloud data out of the images by using a semi-global matching algorithm, see Figure 5. This point cloud data was then integrated into the Continuous Location Demonstrator Application.

Stereo vision Conclusion and Outlook

Using stereo vision mobile mapping for a building acquisition offers a range of advantages like rapid acquisition at regular driving speeds (similar to mobile laser scanning systems), fully automatic processing of the stereo imagery and automatic extraction of dense, coloured point clouds. While points clouds could be considered a by-product, the accurately georeferenced stereo imagery itself bears an enormous potential. The high-quality metric imagery is much easier to interpret than point clouds and supports an abundance of automatic object recognition, extraction and mapping tasks, see for example (Cavegn S., 2011), in the future.

Figure 8 : Cameras mounted for building capturing

Visual-Inertial Simultaneous Localization and Mapping

The method presented here provides a solution that enables *real-time* localization of personnel or equipment in in- and outdoor environments. In contrast to most existing approaches it does not rely on any infrastructure such as GPS or RFID equipment. No apriori information about the scene such as floor plans are required. In addition to accurate position information this method allows for on-line reconstruction of a sparse map of the environment without the need for any manual post processing. The hardware consists of a set of low-cost, lightweight and low-power sensors. This makes the system portable and allows operation for extended periods of time.

A tight combination of visual and inertial sensors improves localization accuracy of traditional vision-only methods and allows for operation in poorly textured environments. Additionally, errors in the roll and pitch axis can be bounded through gravity. The following sections will present the prototype sensor suite and give an overview on the various processing stages involved. Results are discussed with a focus on weaknesses and future improvements.

Hardware

The custom sensor suite was designed aiming at low power consumption and light weight in order to be head-mountable. The hardware prototype is depicted in Figure 9 and consists of up to four machine vision grade CMOS camera sensors, a commercial-grade strapdown inertial measurement unit (IMU), two high-power LEDs and an embedded computer. The cameras record high dynamic range images at a rate of 10 Hz. The IMU provides measurements of linear acceleration and angular velocities at 1 kHz, in synchronization with the cameras. The high power LEDs are synchronized with the camera shutter and allow for operation in complete darkness.

Processing of Visual Information

The initial step in motion estimation is the detection of a set of distinct interest points in the images, as illustrated in Figure 10. To enable real-time operation, given very limited computational resources, a FAST keypoint detector is employed (Rosten & Drummond, May 2006). An adaptive thresholding scheme guarantees the detection of a set of well-distributed features in the image plane. Each keypoint is then described by a lightweight binary descriptor, see (M. Calonder, 2010). Features from successive frames are matched and gross outliers are rejected

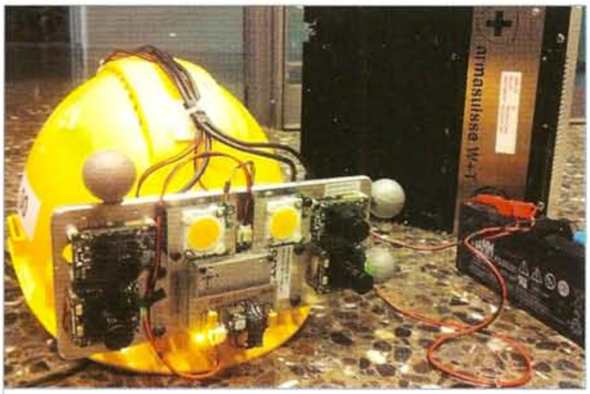


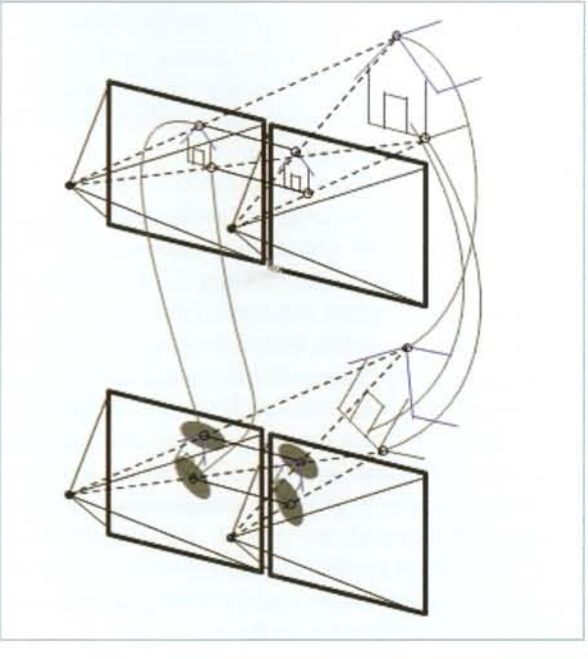
Figure 9 : Head-mounted sensor rig consisting of four machine vision CMOS camera sensors, a commercial grade inertial measurement unit and two high power LEDs.

based on inertial measurements. A P3P (L. Kneip, 2011) and RANSAC (Bolles, 1981) stage provide an initial motion hypothesis that is further optimized in an iterative, non-linear refinement step (Konolige, 2010). The pose estimation stage lasts no longer than 70ms on a single core 1.6GHz Intel ATOM embedded computer.

Vision-IMU Fusion

Modern Visual SLAM (Simultaneous Localization and Mapping) algorithms exhibit a very low drift but often lack the robustness required for localization applications in real scenarios. If visual tracking fails even for a short period of time e.g. due to the lack of texture or occlusion, position and attitude are lost and cannot be recovered unless a loop is detected. However, if visual motion estimation is tightly coupled with inertial readings, short periods of tracking failures can be tolerated, and error in certain axes can even be bounded globally, see (R. Voigt, 2011).

Figure 10 : Distinctivevisual features are tracked across successive frames yielding a chain of transformations.



We propose a non-linear state estimator that solves recursively for position, attitude, gyroscope and accelerometer biases, and velocity. The recursive nature of the filter allows for minimal resource usage. Further details can be found in (R. Voigt, 2011).

Results and Future Works

Figure 11 illustrates the sparse map that was generated in real-time while entering and walking through an Armasuisse building in Thun, Switzerland. The height is estimated with an error of < 0.5m. The more detailed view on the left shows a dense reconstruction also generated in real-time but with an update rate only slightly above one Hertz. While position, orientation and sparse map were continuously transmitted to the base station, the dense map was only stored locally to enable the simultaneous transmission of a live video stream.

In contrast to absolute positioning systems such as GPS, Visual-Inertial SLAM provides only relative measurements of position. Unless a position fix through e.g. GPS is available at least once, absolute position cannot be determined. A further drawback of the proposed method is its inherent drift. Once position has been estimated erroneously, this error cannot be corrected unless a previously seen position is revisited.

To further increase robustness and accuracy, higher resolution cameras with a larger field of view need to be integrated. Additional sensors such as altimeter and magnetometer may be tightly integrated as well. However, integration of low-quality MEMS magnetometers for indoor applications is non-trivial and requires careful design.

Visual mapping techniques are ideally suited for collaborative mapping tasks enabling multiple users to explore unknown environments simultaneously and construct one single, consistent map of the environment. If further optimized in terms of performance, power consumption and map consistency, such a Visual-Inertial SLAM system could prove invaluable for numerous applications including augmented reality.

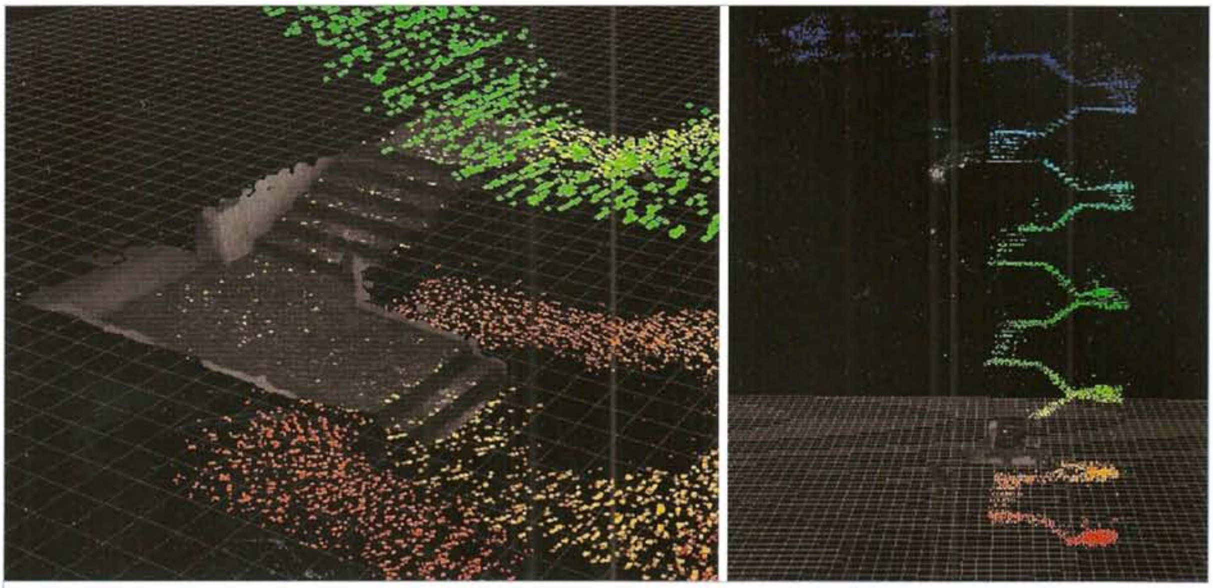


Figure 11 : Sparse map of an Armasuisse building in Thun computed in real-time while entering and walking through the building with the head-mounted sensor rig. Height is estimated with an accuracy of < 0.5m while climbing over five floors.

Localization and Identification of Friend or Foe using Active RFID

Motivation

The mechanism of traceability in a building are important in terms of safety and protection of persons as well as of assets (geofencing). It is however necessary to detect individuals at the earliest in order to provide an adequate response and play on (avoid or create) the element of surprise. To do this, the system should be able to identify and geolocate individuals outside of the building on an admissible perimeter. Note also that the system should be easy to use and deploy, especially in military engagement phase, and present accurate and reliable information.

Ad hoc Platform for Indoor and Outdoor localization and identification

Ad hoc network technology allows the easy construction of active RFID readers networks. Their configuration of the network nodes is orchestrated by a master reader and is fully automatic and secure. Additional readers can enter the network and extend the overall radio-frequency coverage. A simple structure of such a network is illustrated in Figure 12. The structure used in the project involved several such nodes in order to attain a sufficient density, thus reaching an acceptable localization accuracy.

Active RFID tags embedded or not within various sensors transmit their information in the ad hoc network. The messages are either received directly by the master reader or indirectly through registered network nodes (those nodes are called relays, repeaters or Xtenders). When a node gets a message it will add its unique identifier to it and relay it until the master reader. This feature allows real-time localization by zone. The shape of the tags used in the project are depicted in the Figure 13.

The tags are powered by coin cell batteries and have an autonomy of several years, making them attractive for wearable sensors in general and for the case of Friend or Foe identification in particular. The nodes or relays are also low power, and can hence be powered by portable batteries.

In the case of localization and identification of Friend or Foe, each soldierswill at least be wearing a tag identifying him/her. Additional tags with sensors can be worn by the soldier to gather information about his/her state such as the body temperature, if one is moving or not, etc. Nodes with limited coverage allow the localization of the soldiers inside the building in a confined space. The network was constructed with these nodes integrating a Relay, which role is to provide the backbone over which the gathered information is transmitted to the master. The nodes can be mobile (powered by battery), and therefore be worn by soldiers (i.e. the soldier node) offering absolute and relative positioning between them.

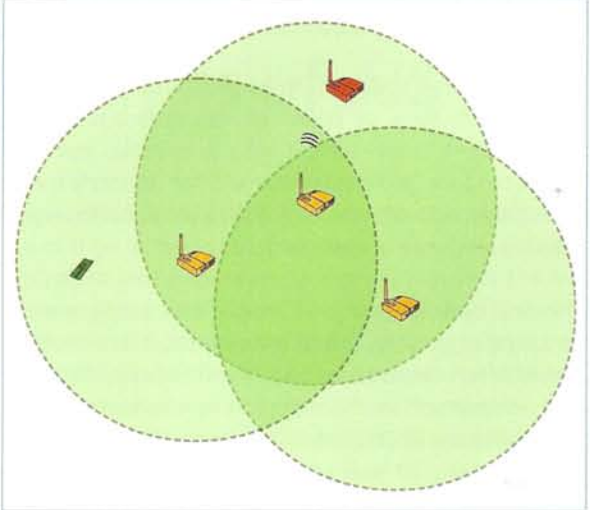


Figure 12 : In this figure, the red reader is the master which orchestrates the establishment of the adhoc network. The orange units are the relays. The information sent from the detectors is received by at least 2 relays which retransmit them to the master reader.

Multi-hop networking in an unreliable environment

When it comes to the integration of visual sensors with 3D rendering & visualization engines, there can easily be large amounts of data to be exchanged within short time periods. Depending on the preprocessing done at the sensors these data consist of sparse or dense point clouds with 3D coordinates and additional coloring information used to optimize visualization. For example, laser based sensors can nowadays generate millions of such points within a few seconds, whereas the sensing speed is still growing fast with each new generation of sensors.

Figure 13 : RFID active tags used for simultaneous friend-foe identification and localization.

Within the aforementioned vision of continuous 3D localization, each individual is carrying a sensor permanently scanning its environment. The acquired point clouds are then immediately spread to the other individuals. This enables each individual to additionally get the visualization of the environment scanned by his/her teammates. One can think of a distributed but collaborative scanning of the environment delivering a complete visualization of the environment within short time.

The distribution of the data acquired by the different sensors is key for successful realization of such a collaborative scanning and visualization of the environment. To enable broadband communication among all sensors and 3D rendering & visualization engines it perfectly makes sense to consider adhoc and meshed networking mechanisms, where the individuals are also carrying the communication equipment, and hence forming the network themselves. Although, depending on the number of individuals, the size of the field of application, and hence the expected distance between the units it might be required to additionally place relay nodes to achieve the required connectivity level.

Within our demonstrator we integrated the *Heterogeneous Communication Platform (HCP)* provided by Rayzon Technologies Ltd. These nodes combine router, server, and terminal functionality in a portable form factor. With the help of multiple communication interfaces and a simple LED-based indication of the connectivity level it enables easy deployment of an adhoc/meshed network providing the high bandwidth communication required to rapidly exchange 3D localization data among participants.

In a first step we incorporated a single individual carrying a helmet with the sensors. The generated data (i.e. files with point clouds) is then transferred to the first HCP node via a IEEE 802.11b/g WLAN connection using a shared network drive. Within the HCP the data is further processed by the *MORPHEUS* software, which is basically responsible for abstracting complex communication mechanisms towards applications. The data is split into small junks and spread to all other HCP nodes using well known Peer-to-Peer mechanisms. Especially in an unreliable communication environment based on adhoc/meshed networking mechanisms including multi-hop connections, the application of Peer-to-Peer mechanisms significantly improves the overall performance and robustness due to the small transfer units, the used forward error correction, and multi-path/multi-source data transfers.

The MORPHEUS software allows the seamless usage of any sort of communication technology to be used to interconnect the HCPs. Within the demonstrator multiple IEEE 802.11b/n (3x3 MIMO) WLAN interfaces were used to enhance robustness and increase available bandwidth between the HCPs. The HCP nodes were placed in a row one after the other resulting in a dynamically deployed multi-hop network, see Figure 14, comparable to the scenario where an individual is entering a building and continuously sending information about the scanned environment and its own position to the rest of the team waiting at the entrance of the building.

The *Virtual Globe Server* provided by the FHNW incorporates the 3D rendering and visualization of the received data generated by the sensor unit.

Additionally to the Peer-to-Peer based data distribution among the HCP nodes, a live video stream captured by a camera mounted onto the sensor unit helmet was routed through the adhoc/meshed network to a dedicated Video Server, enabling a direct comparison of the dynamically generated virtual environment with the real environment. The position and orientation of the sensor unit was transmitted using a UDP based protocol and routed to the Virtual Globe Server enabling the visualization of the very same view point (Figure 15).

Conclusion and outlook

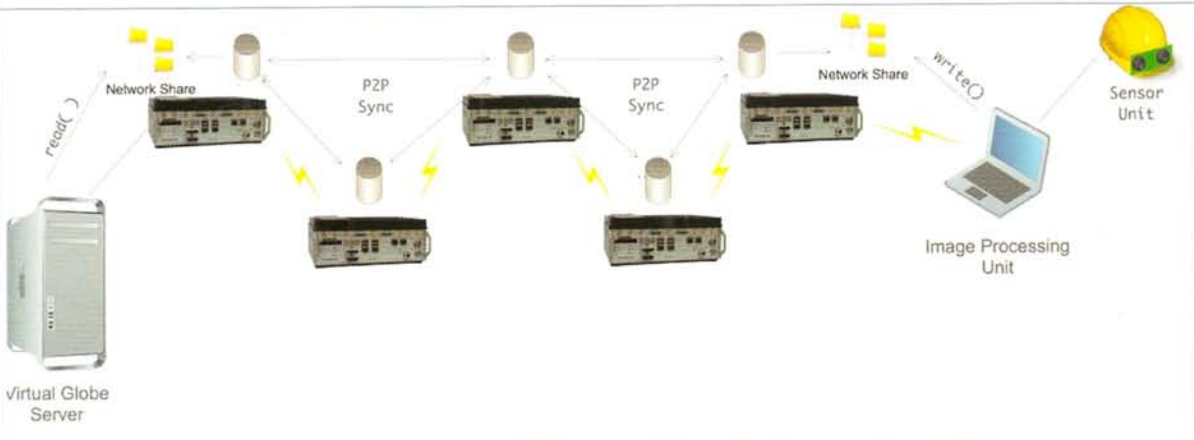
The Continuous Location Demonstrator is at this stage only a prototype or proof of concept, but further development could turn it into a reliable, essential tool for indoor geolocalization and for remotely monitoring dynamic objects and actors in complex – and possibly unknown – environments.

The real-time mapping of indoor and outdoor environment is a very challenging issue. Each method tested offers advantages and disadvantages. Accurate and comprehensive 3D mapping methods can nowadays difficially provide real-time maps. New handheld and kinematic 3D laser scanners will help reducing the time of acquisition and processing. Possible existing 2D floor plans, in combination with the 3D point clouds, will also help getting a quicker and more reliable 3D map of the environment.

The display of the building in different views, the real-time tracking of individuals simultaneously to their identification, and the representation of the collected point cloud fragments gives the user a good understanding of what is happening in reality. It is possible to recognize and dynamically map unknown building structures, floor levels and room arrangements. The use of the OpenWebGlobe SDK proved to be a good choice because of the flexibility of the SDK. Additionally, it allows to easily enhance and adapt the required functionalities and offers a rich set of already implemented features like fly-to animations and camera handling possibilities.

With respect to positioning, the different technologies used have shown a good level of complementarity and

Figure 14 : Setup of the communication system used within the demonstrator.



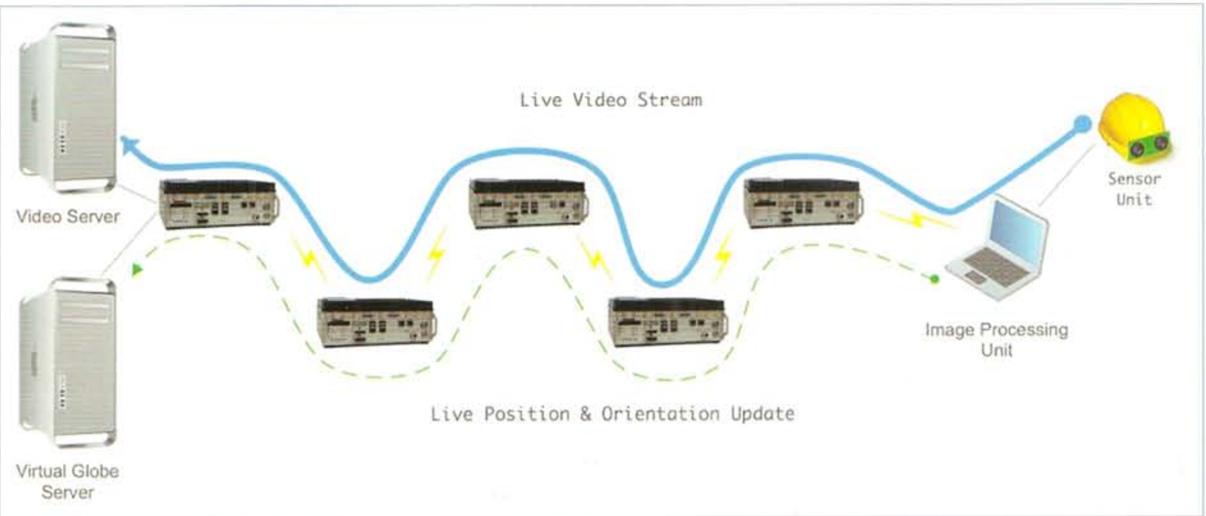


Figure 15 : Peer-to-Peer live video stream and live position & orientation data routed from the sensor unit to the servers.

their integration, together with some map matching capabilities, will for sure enhance the precision as well as the reliability of such system.

Acknowledgement

This research was financed by Armasuisse Science + Technology, the hub for research and technology of the Swiss department of defence.

References

Bolles, M. A. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, (pp. 381-395).

Cavegn, S., 2011. Automatische Verkehrszeichenkartierung aus mobil erfassten Stereobildsequenzen. Master-Thesis. Institute of Geomatics Engineering at the University of Applied Sciences Northwestern Switzerland, FHNW.

Christen M, Nebiker S. Large Scale Constraint Delaunay Triangulation for Virtual Globe Rendering. In: Kolbe TH, König G, Nagel C, eds. Advances in 3D Geo-Information Sciences. Berlin, Heidelberg: Springer Berlin Heidelberg; 2011:57-72. Available at: http://www.springerlink.com/content/qm8x71x25205k287/. Accessed March 31, 2011.

Khronos Group, WebGL Reference http://www.khronos.org/webgl

Konolige, K. (2010). Sparse sparse bundle adjustment. BMVC (pp. 102.1-102.11). British Machine Vision Conference.

L. Kneip, D. S. (2011). A novel parameterization of the perspective-three-point problem for a direct computation of absolute camera position and orientation. CVPR. International Conference on Computer Vision and Pattern Recognition.

M. Calonder, V. L. (2010). Brief: binary robust independent elementary features. European conference on Computer vision, (pp. 778-792).

Nebiker S, Bleisch S, Christen M. Rich point clouds in virtual globes – A new paradigm in city modeling? Computers, Environment and Urban Systems. 2010;34(6):508-517. Available at: http://dx.doi.org/10.1016/j.compenvurbsys.2010.05.002. Accessed November 29, 2010.

R. Voigt, J. N. (2011). Robust Embedded Egomotion Estimation. IROS. International Conference on Intelligent Robots and Systems.

Rosten, E., & Drummond, T. (May 2006). Machine learning for highspeed corner detection. European Conference on Computer (pp. 430-443). European Conference on Computer Vision.

Authors

Dr. Quentin Ladetto, http://www.linkedin.com/in/ladetto
Janosch Nikolic, http://www.asl.ethz.ch/people/nikolicj
Martin Rufli, http://www.asl.ethz.ch/people/ruflim
Dr. Cédric Pradalier, http://www.asl.ethz.ch/people/cedricp
Benjamin Loesch, http://www.fhnw.ch/personen/benjamin-loesch
Martin Christen, http://www.fhnw.ch/personen/martin-christen
Prof.Dr. Stephan Nebiker, http://www.fhnw.ch/personen/stephan-nebiker
Franck Schmidt, http://www.geosat.ch/?p=geosat&id=2#2
Patrick Lathion, http://www.geosat.ch/?p=geosat&id=2#2
Dr. Marc Danzeisen, http://ch.linkedin.com/pub/marc-danzeisen/0/346/aa1
Gilles Cherix, http://ch.linkedin.com/in/cherix
Christophe Abbet
David Werlen, http://ch.linkedin.com/in/davidwerlen
Dr. Jari Curty, http://ch.linkedin.com/in/jpcurty
Dr. Luis Harik, http://ch.linkedin.com/pub/dr-louis-harik/1/388/9b4
Dr. David Humair, http://www.linkedin.com/pub/david-humair/0/558/911