STEREOPOLIS II: A MULTI-PURPOSE AND MULTI-SENSOR 3D MOBILE MAPPING SYSTEM FOR STREET VISUALISATION AND 3D METROLOGY

Nicolas Paparoditis[1], Jean-Pierre Papelard[1], Bertrand Cannelle[1], Alexandre Devaux[1], Bahman Soheilian[1], Nicolas David[1], Erwann Houzay[1]

1: IGN, Laboratoire MATIS; Université Paris-Est
Institut National de l'Information Géographique et Forestière, 73 avenue de Paris, 94165 Saint Mandé – FRANCE
firstname.lastname @ ign.fr

Résumé

Nous présentons dans cet article un système de numérisation mobile 3D hybride laser-image qui permet d'acquérir des infrastructures de données spatiales répondant aux besoins d'applications diverses allant de navigations multimédia immersives jusqu'à de la métrologie 3D à travers le web. Nous détaillons la conception du système, ses capteurs, son architecture et sa calibration, ainsi qu'un service web offrant la possibilité de saisir en 3D via un outil de type SaaS (Software as a Service), permettant à tout un chacun d'enrichir ses propres bases de données à hauteur de ses besoins. Nous abordons également l'anonymisation des données, à savoir la détection et le floutage de plaques d'immatriculation, qui est est une étape inévitable pour la diffusion de ces données sur Internet via des applications grand public.

Mots clés :
Numérisation 3D mobile, infrastructures de données spatiales, tête panoramique multi-caméra, base stéréoscopique, capteurs laser, calibration multi-capteurs, outils de saisie 3D de type SaaS, détection de piétons, détection de plaques d'immatriculation

Abstract

In this paper, we present a hybrid image/laser multi-sensor mobile mapping system allowing to capture a spatial data infrastructure which is compliant with several applications ranging from multimedia immersive visualisation to 3D metrology across the web. We detail the design of the system, its sensors, its architecture, and its calibration. We also present a web-based service on the acquired data that allows immersive stree navigations but also 3D plotting tools that can allow a client to enrich his own databases. We also address some data processing issues related to privacy, i.e. pedestrian and car plate detection and blurring, that are mandatory for the dissemination of data through web-based consumer applications.

Keywords :
Mobile mapping systems, spatial data infrastructure, multi-camera panoramic head, stereo rig, lidar sensors, multi-sensor calibration, SaaS 3D plotting tools, pedestrian detection, car plate detection

1. Introduction

Street-based Mobile Mapping Systems (MMS) have encountered a large success in the recent years with wide applications ranging from surveying to multimedia and city modelling (Goulette et al., 2006) as evidenced by the work of (Ellum and El-Sheimy, 2002) or even more recently in (Petrie, 2010). Street-based mobile mapping systems are complementary to geospatial imaging systems to capture building facades, road and pavements, and street furniture in urban canyons. Indeed, aerial nadir or oblique imagery can also be used for capturing urban canyons but only with centimetric resolution aerial surveys that are acquired at very low altitude with thus severe overcosts due to the multiplication of acquisition strips. Nevertheless, for many applications these aerial acquisitions would not be self-sufficient due to all vertical occlusions in the urban canyon such as trees, nor adapted to capture the geometry of vertical objects such as facades, and definitely not adapted for underground surveys, etc.

Many different street-based mobile mapping concepts have been developed with different sensors, and different geometrical designs which are usually adapted and tuned for a given application. Many of these systems are LiDAR-based and are designed specially for road-based inventory applications. Nevertheless some other systems use optical devices as primary sensors for surveying. The GeoInvent company (GeoInvent, 2012), used to design vehicles with modular convergent cameras and/or stereo rigs to offer manual photogrammetric plotting facilities for surveying of roads and streets network inventory to build road databases. TeleAtlas used the GeoInvent solution to build with massive human labour 3D city models for the GPS navigation systems. The Viametris company (Viametris, 2012) proposes automatic tools to extract 3D road objects (road signs, road marks, etc.) for road inventories from the

Figure 1 : The STEREOPOLIS II mobile mapping system.

data acquired by their MMS. Digital globe systems such as Google Maps (Anguelov et al., 2010) or Microsoft Bing Map StreetSide (StreetSide, 2012) are panorama-based in order to propose immersive visualisation possibilities on the web. Another system like the EarthMine (EarthMine, 2012), which can interconnect itself with a GIS software, is equipped with a vertical panoramic stereo rig that allows to build automatically 3D point clouds of the scene that can be used for image-based rendering to provide a higher degree of immersivity in street navigations and the system also provides image-based modelling tools. Some other systems mix laser and image acquisitions for 3D facade modelling and texturing to upgrade LOD2 city models to LOD3 (Bénitez et al., 2010). The Navtech true system car is also equipped with optical imagery (panoramic head) and lidars to collect complementary features necessary in GPS navigation systems or in Advanced Driver Assistance Systems.

2. A multi-purpose MMS for acquiring a 3D spatial data infrastructure ?

In (Van den Heuvel, 2009), the authors describe the project and the workflow for the collection of aerial and street-based image databases over Netherlands. Nevertheless, to our knowledge, no databases acquired by street-view systems are yet available and are open Spatial Data Infrastructures (open-SDI). No open street-view system proposes yet photogrammetric surveying/measuring services for either for public or professional use. Having an open open-SDI ecosystem lying on data with relative and absolute centimetric localisation would be extremely helpful to allow the development of new innovative applications and new markets at a national or international level. For instance, building semi-automatically or automatically 3D visual landmarks (Brenner, 2010) databases from this SDI could allow autonomous vehicles to navigate with low cost localisation systems as opposed to the Google autonomous vehicles which are currently using navigations systems as the ones used in mobile mapping systems (Markoff, 2010).

A central question is thus: Is there an ideal mobile system capable of acquiring infrastructure data in one go which is compliant with the needs of all applications? The point here is obviously to reduce the cost of the acquisition by adding new sensors allowing to address new applications to make the SDI ecosystem possible. If the data exists and is easily accessible, if its quality is sufficient and guaranteed, and if the economic model is smart, data would be used everywhere for every application. In this paper, we only focus on the technology that could allow to build that infrastructure.

3. A laser-based, an image-based or a hybrid-based system ?

As we stated previously, most of systems available on the market are laser-based. Nevertheless, the point clouds that they acquire are, on the one hand, difficult to visualise and manipulate due to their huge size. On the other hand, they are also difficult to interpret without additional information. Cameras can be optionality integrated on these systems but are not used for metrology but to ease the local interpretation of the point clouds and to texture the point clouds.

Conversely, the EarthMine system is image-based. It is composed of a vertical stereo rig with two very high resolution multi-camera panoramic heads. These panoramas allow a street-view like navigation but are also a plotting support for 3D image-based modelling tools. In this stereo acquisition geometry, each point of object space visible from the rig is seen in stereo by the two panorama heads. This panorama stereo allows to generate a depth map, which is fully superimposable with the geometry of one of the panoramas but nevertheless at a very high computation cost. The generated RGBD images (the three color channels plus depth as a channel) have interesting properties for image-based rendering in immersive navigation. As a consequence, every point of the dual point cloud has also a synchronous and genuine corresponding colour with less parallax problems that the ones encountered in the colouring of laser point clouds with images which are not acquired exactly at the same time. The point cloud has the same density as the images thus a much higher density and a spatially more homogeneous distribution than laser-based acquisition. The stereo baseline (close to 1 m) is rather short to ease image matching but at the cost of a loss in depth estimation accuracy. The accuracy is of several centimetres at ten meters instead of one centimetre at ten meters for laser-based measurements. This is thus insufficient for various surveying purposes for instance. Of course, point cloud based visualisation and 3D plotting tools can also be used with this system.

4. The design of STEREOPOLIS

4.1. General properties

The STEREOPOLIS system is composed of a navigation device hybridating measurements from two GPS, an inertial measurement unit, and a wheel odometer. They are fused through a tightly coupled compensation to provide a pose (absolute position and orientation of the platform reference system in a global reference system) whatever the GPS configuration and whatever the GPS outages and multiple paths. Such situations are very frequent, especially in the old town centers of European cities where urban canyons are narrow. This navigation device produces a direct georeferencing of the platform that allows to locate all data and imagery acquired by the MMS in a global reference datum.

4.2. Image sensors

STEREOPOLIS is equipped with a panoramic head composed of ten full High-Definition (HD) cameras mounted within a rigid body to ensure that the relative pose of the cameras remain unchanged along the acquisition. The cameras are perfectly synchronized, and mounted very closely to one another. This permits to avoid parallax, and have the same exposure times in order to build geometrically and radiometrically seamless panoramas. The cameras have been selected to have a high radiometric dynamic and a high signal to noise ratio (200-300), in order to cope with the strong variations in illumination between the shadowed and the lightened sides of the street. The panoramas which are acquired are very large (10,176*5,088 when the panoramas are generated) and are of very high resolution: the field of view of one pixel is 0.04°. Thus, the Surface Sampling Distance on a fronto-parallel surface at 10 m is around 0.5 cm.

The camera triggering is servoed on the displacement of the vehicle - thanks to the real time output of the navigation system - in order to acquire images at regular distance intervals in order to limit the volume of data which is registered. In our surveys, we usually define that spacing distance to 3 or 4 meters.

Figure 3 : The sensors of STEREOPOLIS. **Upper left**: the POS-LV220 navigation system. **Upper right**: one of the sixteen Pike full HD cameras. **Bottom left**: one of two RIEGL LMS-Q120i lidar. **Bottom right**: the HDL-64E of VELODYNE.

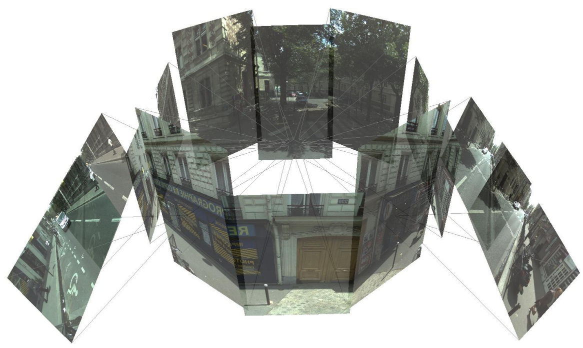

Figure 4 : The set of images acquired at one pose of STEREOPOLIS. One can see the images acquired by the panoramic head in the middle and also the images acquired by the stereo rigs on the left and on the right of the Figure. Each image is full HD.

Every point in the scene, which is across-track from the vehicle, is visible in multiple along-track panorama views and, can thus be measured in 3D. The system is also equipped with two complementary stereo rigs: one looking at the front and one looking at the back of the vehicle. They allow 3D measurements on objects which are along track (road marks, road signs, etc.). Each stereo rig plus the corresponding axial image of the panoramic head offers thus a triangular tri-stereo which is very favourable for a 3D automated reconstruction of

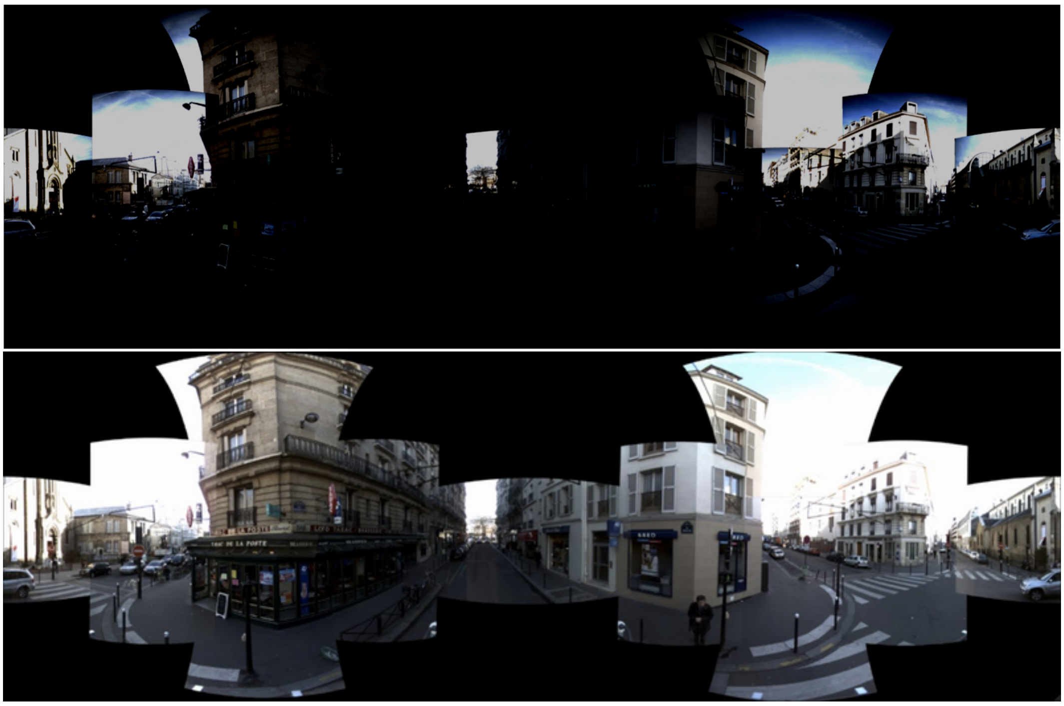

Figure 2 : The high dynamic of cameras on STEREOPOLIS. **Top**: the 12-bits panorama linearly converted to 8 bits. All details in shadows disappear. **Bottom**: the 12 bits panorama efficiently converted to 8 bits.

objects (no degenerated cases when matching contour lines).

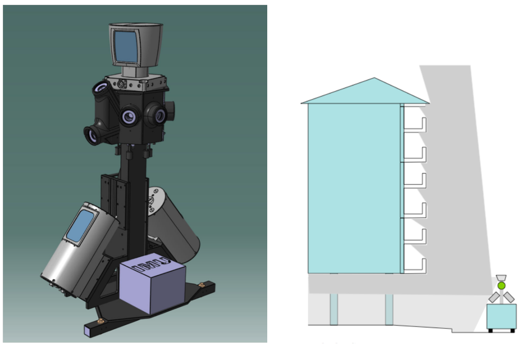

Figure 5 : **Left**:The rigid mount with the panoramic head, the IMU, the VELODYNE and the two RIEGL lidar devices. **Right**: Our acquisition geometry and coverage of the object space with the lidar sensors. The space covered by the RIEGL sensors and the VELODYNE are displayed in dark and light greys, respectively.

4.3. Lidar devices

In addition to the optical imagery, we also capture the geometry of the scene with both plane sweep and 3D lidars. Even though the spatial density of lidar-based point clouds is lower than image-based point clouds that can be generated by dense multi-view image matching (Penard et al., 2005), (Pierrot-Deseilligny and Paparoditis, 2006), the accuracy and above all the reliability of the 3D individual measurements provided by high-quality lidars are still superior. In addition, lidar-point clouds can be directly used at the end of the survey and do not need heavy computation loads.

In our pipelines, we use lidar to extract 3D features (Demantké et al., 2011) and 3D objects automatically (posts, trees, etc.) (Monnier et al., 2012) , conversely rather difficult to extract from images. Nevertheless, images are directly processed to accurately extract some specific 3D objects such as road marks (Soheilian et al., 2010) or road signs (Belaroussi et al., 2010).

On the one hand, two metrological LMS-Q120i plane sweep lidars with a centimeter depth accuracy with a field of 80° are placed on each side of the car to observe mainly the facades. Ideally, two other lasers looking downwards could allow a better capture of the bottom of the canyon. On the other hand, an HDL-64E lidar from VELODYNE is integrated to capture the bottom part of the canyon. Furthermore, with a scan rate up to 1.3 million points per second, it can be used to track and remove mobile objects from the point clouds for some applications. The high speed rotation (10-15 Hz) of the rake of 64 lidar fibers around the z-axis and, thus, the continuous variation of the lidars observation direction allows to potentially observe the same scene elements many times along the track of the vehicle This minimises the surface of occluded areas. Of course, the sampling properties of the surfaces differ from one laser to the other.
Both lidar points clouds are complementary. The 3D points of the VELODYNE are used to fill up shadow ar-

eas of the RIEGL lasers or to detect small linear objects like pavement limitation posts that could be missed. Indeed, the RIEGL laser only digitises up to 100 scan lines per second thus the spatial sampling in the along track direction is 10 cm which is close to the size of the object.

As shown on Figure 5, the IMU together with all the major sensors are mounted on a rigid structure in order to ensure that they are no relative deformations all along the survey.

5. The architecture of STEREOPOLIS

As mentioned previously, the triggering of the cameras is conditioned by the displacement of the vehicle in order to ensure a spatial regularity of image acquisition. The STEREOPOLIS managing application receives the real time poses of the navigation system and determines when the following node needs to be captured. In our surveys, it happens when the driven distance exceeds a few meters, or every few seconds when the vehicle is stopped, usually at a traffic light - in order to keep a reasonable continuity within the sequences. The application then sends a trigger that is physically duplicated with a splitter - in order to avoid delays - and sent to every camera to acquire an image. Once the master camera has acquired the image, it immediately sends an output trigger to the navigation system in order to date the event. In parallel, the image of each camera is sent on the network and written on the associated PC and on the 750 Gigabytes hard disk which is unique for every camera in order to be able to write up to six full HD 12 bits images per second. For each image acquired, thumbnails are sent to the master PC, and are displayed on the operator's screen to ensure the correct functioning of the cameras. The writing facility for each camera is also displayed in order to pre-identify possible problems on the network thus encouraging the vehicle to slow down if necessary. For both lidars which acquire data continuously, a Pulse Per Second entry connected to a GPS allows to date very accurately the different lidar measurements.

6. One survey example

The 12th district of Paris was surveyed in April 2008 - in the scope of the iTOWNS project (http://www.itowns.fr) of the French National Research Agency (ANR).
It was carried out in 12 hours, which represents 2 days acquisition. Although the street network is 120 linear kilometers, 180 km had to be driven due to the circulation constraints (one ways, etc.) with an average speed of 15 km/h.
This represents around 45,000 panoramas composed of 450,000 full HD images and around 800 million points for the LMS-Q120i lasers for a global storage of 3 to 4 TerraBytes without any compression.

Figure 10 : Footprint of the survey acquired on 12th district of Paris in April 2008 for the iTOWNS ANR project. One point in red corresponds to one image node. In blue, the IGN building database is displayed. **Left**: Overview of the area. **Right**: Focus on a roundabout.

7. Calibration and georeferencing of the MMS

As mentioned previously, our mobile mapping system is composed of many sensors. In order to merge relatively all the data acquired by the sensors and to interact with them into a global reference datum, we need to determine the relative pose of the lidars and of all cameras composing the panoramic head and the stereo rigs with respect to the vehicle (or navigation system). Our panoramic head can not be calibrated as a whole by itself due to the fact that the overlap between the images is very small (a few pixels). We will thus estimate independently the extrinsic parameters of each camera on the panoramic head.

The pose estimation of the RIEGL laser sensors relatively to the vehicle is relatively straightforward. Indeed, topometric targets can be mounted on specific slots on the laser frame. The 3D position of these slots in the laser reference frame are provided by the manufacturer. The position of the laser center and the rotation between the vehicle reference frame and the laser reference frame can thus be estimated by a topometric measurement of the slots positions in the vehicle reference frame.

Concerning the extrinsic calibration of the cameras, two methods providing similar results have been used to estimate the cameras extrinsic parameters (Cannelle et al., 2012). Figure 11 illustrates the two methods. The first one, named "Off-Line", is based on a network of topometrically surveyed targets which are observed and plotted manually by an operator in every image. The pose of the cameras is then estimated by bundle adjustment. This method requires preferably an outdoor network facility, which is not so easy to build and to maintain. The calibration facility can also be far away from the survey sites and can thus be done in conditions very different from that of the surveys (e.g. difference in temperature which can have an impact on the mechanical deformation of the metal body on which all the sensors are attached). Conversely, the "on-line" calibration directly uses the survey itself. The "on-line" calibration which is a self-calibration method uses the image content, i.e. the

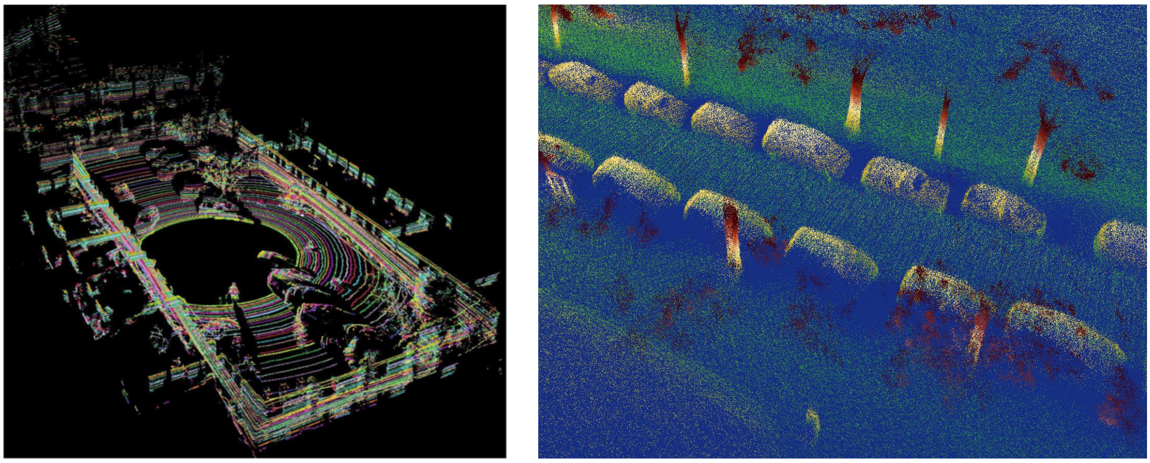

Figure 6 : Point clouds acquired by the VELODYNE HDL-64E lidar (subsampled by a factor 50). **Left**: Points acquired for a single rotation. The points corresponding to the same laser head on the 64 laser rake are coloured in the same color. **Right**: Points corresponding to one section with multiple rotations. The colors of the points are height-coded. One can see that the occluded areas are rather limited.

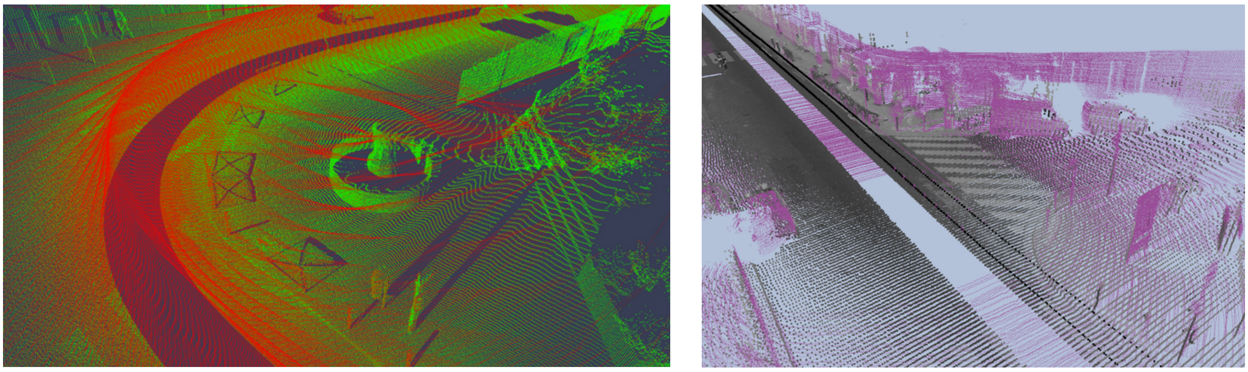

Figure 7 : Point clouds acquired by the two RIEGL devices. **Left**: Acquisition of the bottom of a urban canyon. The 3D points are displayed in green. The rays corresponding to the path of the laser beam from the aperture to the measured surface are shown in red. One can see the important shadows due to the acquisition geometry. **Right**: Superimposition of laser point clouds of the VELODYNE (black) together with points of the RIEGL (pink).

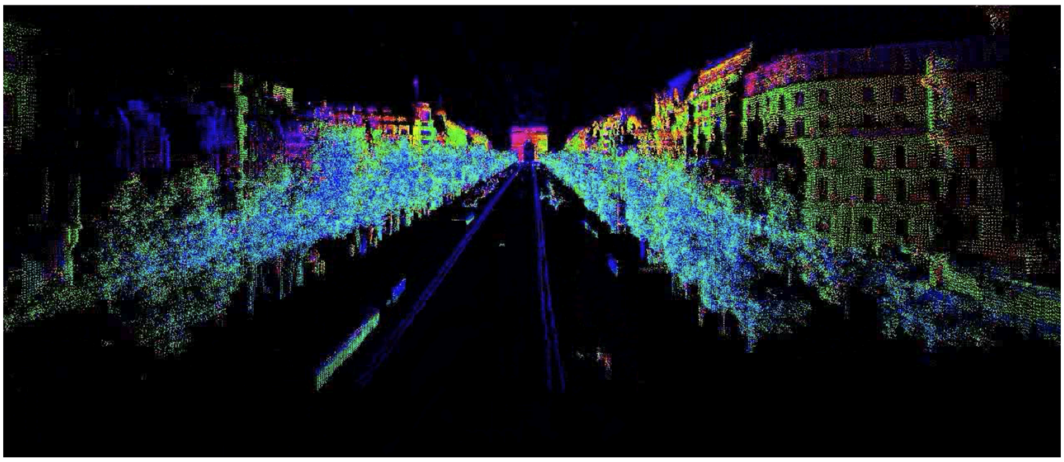

Figure 8 : Point clouds acquired by the two upper RIEGL lidar devices (height colored) over the Champs-Elysées avenue, Paris, France.

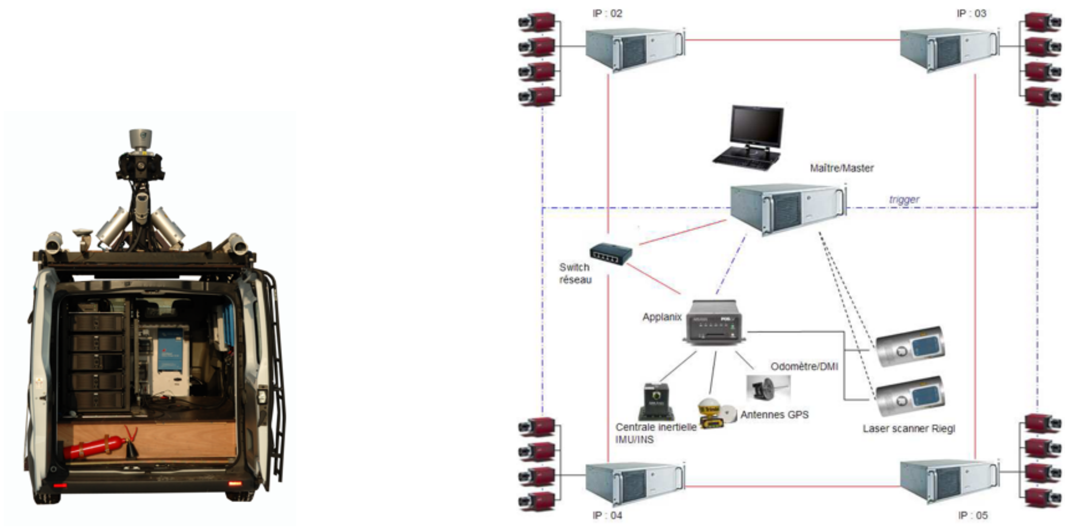

Figure 9 : Left: Back view of the STEREOPOLIS system. One can see the five personal computers (PC) used for processing and storing the acquired imagery. **Left**: The global system architecture of STEREOPOLIS.

matched tie points across the views are used to estimate by bundle adjustment the relative pose of images with respect to the vehicle. Thus, (1) the number of tie points is potentially much higher and (2) tie points sample the image space in a much denser way. This latter method is very efficient and accurate in urban areas but can be unreliable if the environment is not friendly to extract tie points (e.g. in forests, etc.).

The relative pose of the direct georeferencing system is very good (centimetric quality) and so is the absolute orientation (close to 0.04°). Nevertheless, the absolute pose/localisation heavily depends on the GPS configuration, and before all on the GPS gaps and on the drifts of the Inertial Measurement Unit. The absolute georeferencing can be photogrammetrically improved thanks to the observation in the acquired imagery of existing 3D surveyed points (Cannelle et al., 2012) along the survey. The absolute pose of the image node observing one surveyed 3D point can be estimated by space resection because the distance between the observed point and the sensors is known thanks to the lidar point clouds. As the relative poses of our platform are very good, the absolute pose of the re-sected node can thus be propagated along the trajectory.

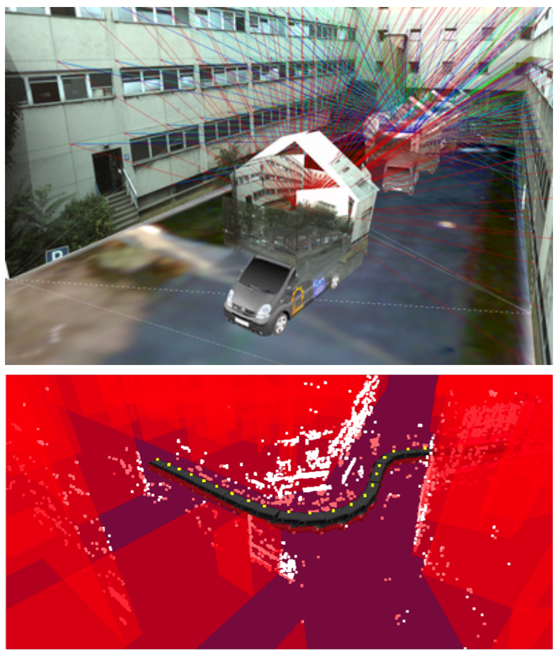

Figure 11 : Calibration of optical sensors with respect to the vehicle reference system. **Above**: Use of a network of surveyed targets fixed on building facades. **Below**: Self-calibration using natural details of the scene along the survey. A 3D model of the scene, displayed in red, is superimposed to the 3D measurements, in white, corresponding to the matched interest points. It qualitatively shows the satisfactory overall calibration procedure.

8. Image-based street navigation and 3D plotting services

Multimedia web-based applications using street-level high resolution images acquired by mobile mapping have been developing fast in the last years. While the quality of images is increasing, the multimedia aspects of such an approach are still to develop. All actors like Google Street-View, Microsoft Live Street-Side offer the possibility to walk through cities in panoramas with click-and-go functions but with a limited degree of interactivity. Nevertheless, users all agree that these systems are very useful to visualise the city from a pedestrian level with an excellent image quality.

Our street view application allows to move through cities from a pedestrian point of view and interact with it in order to create and update precise urban maps,

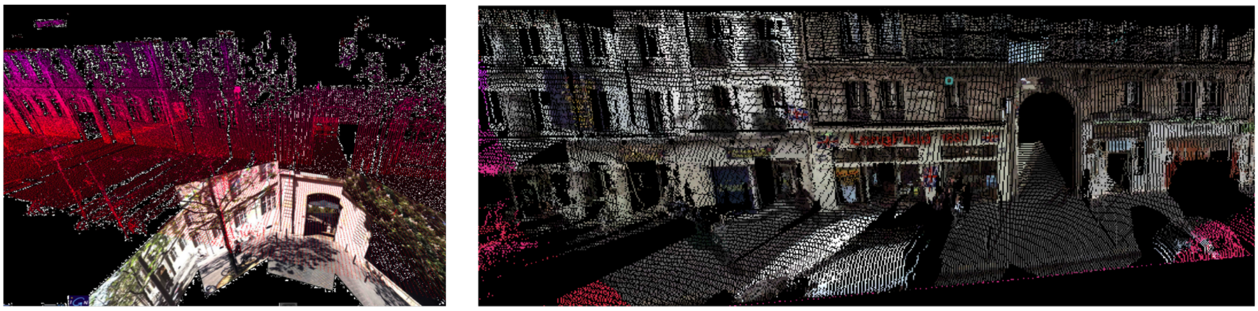

Figure 12 : Relative calibration results of images with respect to the RIEGL lidar point cloud. **Left**: A panorama superimposed with the point cloud. **Right**: A 3D lidar point cloud colored with one of the referenced images.

enabling accessibility diagnostics, services and applications around mobility. We cared about precision in the data acquisition (images and laser) in order to be able to make accurate 3D measurements for professional usages. The user can show large laser clouds in his web-browser, access information on any points, draw 3D bounding boxes and export selections to databases. He can also make measurements from images, integrate 3D models (COLLADA), display OpenData layers, etc. Finally, we use projective texture mapping in real time to texture any mesh with our images

We use multi-planar image panoramas that do not distort the representation of different objects within the scene (straight lines in the images remain straight in the panorama). This eases of course the manual annotation and the 3D measurement of objects. The pre-processing of data is also simpler. We do not need to generate a panorama from the 10 individual full HD images since we directly stream the images to the client navigator using the pose of the cameras and their intrinsic parameters.

An important limitation when creating an application over the Internet is the data transfer. A special effort has been made to transfer efficiently data from the servers to the client. As we work with large images (21 MegaPixels per panoramic) and billions of 3D points we had to find a way to stream images and lidar data fastly enough so that the immersive effect would work.

We use IIPImage fastCGI on the server allowing tile-streaming from JPEG2000 images (using Kakadu library for decompression). A single image source file per camera, not thousands of small separate files. The client can ask a specific part of the original image at a specific resolution with a specific compression. When a client navigates through the city, at every new position the system loads the multi-camera images for this acquisition node to display a panorama view. The loading starts with very low resolution highly compressed images hence taking a short time, then new tiles with better quality appear automatically.

For public applications on the web, we only allow the interaction with the panoramic imagery. The user can then plot objects (points, lines, surface limits) in 2D only within the panorama, and the plottings are then transferred and reconstructed in 3D using the laser point cloud (Figure 13). The user can only access to relative measurements on the objects (heights, surfaces, volumes). For professional applications, it is possible to manipulate directly the laser cloud in the viewer, superposed to the image or alone. While activating only the laser cloud, one can evolve fluidly in the city through 3D points which are streamed dynamically around the position of the user. In this professional mode, the user can plot objects with the same tools as for images but can in addition access to the absolute coordinates of objects in a national or international reference frame.

Figure 13 : Illustration of a 3D measurement service: porch height determination (blue).

Essential for urban maps creation, these tools can also be useful for scientists allowing them to create ground truth databases or to annotate object samples for machine learning techniques (Figure 14).

9. Pedestrian and car plate detection and blurring for privacy preservation

As we are exposing parts of the city life across a multimedia application on the web, we need to be cautious of privacy issues. This means that one needs to blur or remove all identification attributes, especially faces and car plates. People and plates can be anywhere in the panoramas, with varying numbers, sizes, aspects (45°, frontal, profile), with varying light conditions (direct, diffuse, shadows), with often very strong occlusions due to trees, sign posts, cars, etc.

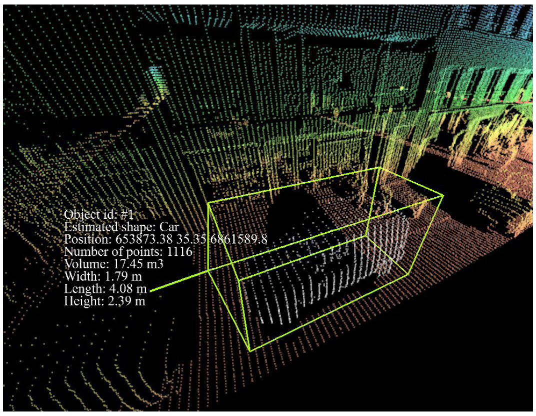

Figure 14 : 3D Bounding Box design and subsequent 3D point extraction (white), and statistics.

Most of pedestrian detectors are designed to detect only standing pedestrians. Indeed, most of the time they are not ideally designed to detect people on bikes, people sitting on a bench or in a car, or lying on the floor, etc., and most often they do not detect them. To deal with these free postures, an addition of a face and profile detector is necessary to increase the completeness of detection. We thus combine a face detector using the famous Viola & Jones algorithm (Viola and Jones, 2001), and a pedestrian silhouette detection algorithm of (Laptev, 2006). Then, we added to those two detections an additional skin detection algorithm in order to eliminate the numerous false positives. The pedestrian system description is showed in Figure 15 and explained in more details in (Devaux et al., 2009). On the 12th district of Paris, we obtained very good results, i.e. a detection rate of 86.2 % and twenty false alarms per panorama. This is not a problem due to the fact that the false positives areas that will be blurred are small and usually in homogeneous areas thus not really visible.

We use also a similar approach for the detection of car plates. Even though car license plates are of rectangular shape with given aspect ratios, they are a difficult object to find in images for all the reasons previously given. Furthermore, in some cases the lack of contrast does not ease the detection of its rectangular shape and, to top it all, the numbers and letters in the plate vary. Nevertheless, the problem of license plate detection is very similar to face detection issue. This is the reason why we also chose the algorithm of Viola and Jones for the detection of license plates. The object appearance has been learnt on a positive image set constituted of 1,500 license plates together with more than 4,000 negative images which are injected to a Haar feature training process (Bradski, 2000) .

The results are also very promising. As for the pedestrian detection we have usually several false alarms per image (Figure 17). This is in fact also not a problem due to the fact that the areas that are detected as plates will be blurred.

10. Conclusion and trends

We have presented a mobile mapping systems allowing to capture a spatial data infrastructure which is compliant with most applications: multimedia and digital globes, surveying and metrology, etc. We have also presented some web services possibly provided with the data that can allow a user to plot 3D objects in order to enrich his own databases or GIS-contents.

The STEREOPOLIS system can yet be enriched by adding other sensors e.g. thermal cameras to address new applications such as measuring energy losses and thermal leaks on facades, and to ensure the follow-up of public policies for the energy reduction bill.

They are two major trends to this work. The first is to reach a centimetric absolute georeferencing of data acquired by STEREOPOLIS in an automatic way by using already georeferenced data such as very high resolution aerial imagery (< 10 cm) (Tournaire et al., 2006) or already existing centimeter georefernced street-view data. The second is to ensure the scaling of the data processing. Indeed, one needs in a production line to process in one night what one has acquired in a day time. Of course using clusters of calculators can help in achieving this objective. Nevertheless, a huge reduction could already be gained by a smart joint processing of lidar and optical images which are very much complementary.

References

Anguelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., Ogale, A., Vincent, L., Weaver, J., 2010. Google Street View: Capturing the world at street level. Computer 43 (6), 32–38.

Belaroussi, R., Foucher, P., Tarel, J.-P., Soheilian, B., Charbonnier, P., Paparoditis, N., 2010. Road sign detection in images: A case study. In: Proc. of the International Conference on Pattern Recognition. IAPR, Istanbul, Turkey, 23-26 August 2010.

Bénitez, S., Denis, E., Baillard, C., 2010. Automatic production of occlusion-free rectified facade textures using vehicle-based imagery. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3A), 275–280.

Bradski, G., 2000. The OpenCV Library. Dr. Dobb's Journal of Software Tools.

Brenner, C., 2010. Vehicle localization using landmarks obtained by a lidar mobile mapping system. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 3A), 139–144.

Cannelle, B., Paparoditis, N., Pierrot-Deseilligny, M., Papelard, J., 2012. Off-line vs. on-line calibration of a panoramic-based mobile mapping system. In: The XXII Congress of the International Society for Photogrammetry and Remote Sensing. Melbourne, Australia, 25 August-1 September 2012.

Demantké, J., Mallet, C., David, N., Vallet, B., 2011. Dimensionality based scale selection in 3D lidar point cloud. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 5/W12), (on CD-ROM).

Devaux, A., Paparoditis, N., Precioso, F., Cannelle, B., 2009. Face blurring for privacy in street-level geoviewers combining face, body and skin detectors. In: MVA IAPR Conference on

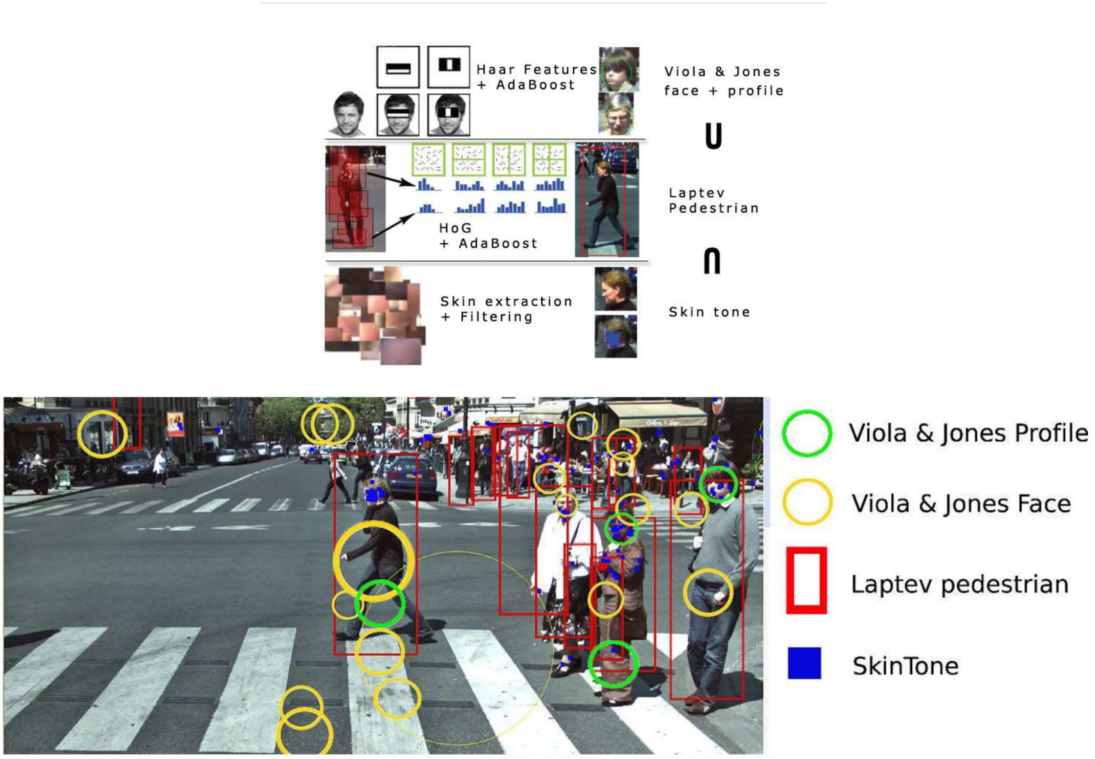

Figure 15 : Pedestrian extraction. **Left**: The pedestrian detection system. **Right**: Results of the four detections on a sample of one image.

Figure 16 : Above: Pedestrian detection and **Below**: face blurring.

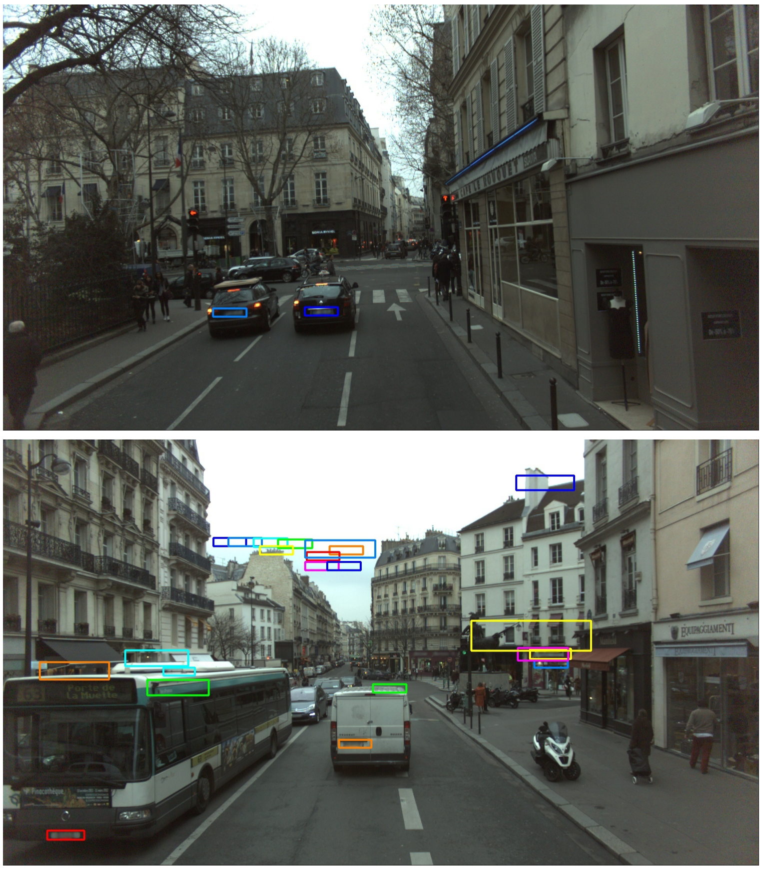

Figure 17 : Results of car plate detection. One can see some false alarms, not yet filtered on a size criteria.

Machine Vision Applications. IAPR, Yokohama, Japan, 13-17 May 2009.

EarthMine, 2012. Accessed 29 February 2012. URL http://http://www.earthmine.com/index

Ellum, C., El-Sheimy, N., 2002. Land-based mobile mapping systems. Photogrammetric Engineering & Remote Sensing 68 (1), 13–17.

GeoInvent, 2012. Accessed 8 May 2012. URL http://www.geoinvent.com.pl/home.html

Goulette, F., Nashashibi, F., Abuhadrous, I., Ammoun, S., Laurgeau, C., 2006. An integrated on-board laser range sensing system for on-the-way city and road modelling. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 1), (on CD-ROM).

Laptev, I., 2006. Improvement of object detection using boosted histograms. In: Proc. British Machine Vision Conference. BMVA, Edinburgh, UK, 4-7 September 2006, pp. 949–958.

Markoff, J., 2010. Google cars drive themselves in traffic. New York Times.

Microsoft, 2012. Microsoft Bing Map StreetSite website. Accessed 1st April 2012. URL http://www.bing.com/maps/

Monnier, F., Vallet, B., Soheilian, B., 2012. Tree detection in mobile laser scans of urban scenes based on accumulations of geometric descriptors. In: The XXII Congress of the International Society for Photogrammetry and Remote Sensing. Melbourne, Australia, 25 August-1 September 2012.

Penard, L., Paparoditis, N., Pierrot-Deseilligny, M., 2005. 3d building facade reconstruction under mesh form from multiple wide angle views. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 5/W17), (on CD-ROM).

Petrie, G., 2010. An introduction to the technology mobile mapping systems. GEOInformatics 13, 32–43.

Pierrot-Deseilligny, M., Paparoditis, N., 2006. A multiresolution and optimization-based image matching approach: An application to surface reconstruction from spot5-hrs stereo imagery. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38 (Part 5/W12), (on CD-ROM).

Soheilian, B., Paparoditis, N., Boldo, D., 2010. 3d road marking reconstruction from street-level calibrated stereo pairs. ISPRS Journal of Photogrammetry and Remote Sensing 64 (4), 347–359.

Tournaire, O., Soheilian, B., Paparoditis, N., 2006. Towards a sub-decimetric georeferencing of ground-based mobile mapping systems in urban areas: matching ground-based and aerial-based imagery using roadmarks. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 1), (on CD-ROM).

Van den Heuvel, F., 2009. Cyclomedia's aerial and ground-based image databases. In: Nederlandse Commissie voor Geodesie 47. Delft, The Netherlands, pp. 53–63.

Viametris, 2012. Accessed 1 May 2012. URL http://viametris.fr/index.php

Viola, P. A., Jones, M. J., 2001. Rapid object detection using a boosted cascade of simple features. In: IEEE Conference on Computer Vision and Patter Recognition. IEEE, Kauai, HI, USA, 8-14 December 2001, pp. 511–518.