PanoraMIS: An Ultra-wide Field of View Image Dataset for Vision-based Robot-Motion Estimation

Houssem-Eddine Benseddik, Fabio Morbidi, Guillaume Caron

PanoraMIS: An Ultra-wide Field of View Image Dataset for Vision-based Robot-Motion Estimation

Houssem-Eddine Benseddik, Fabio Morbidi*, Guillaume Caron

Abstract

This paper presents a new dataset of ultra-wide field of view images with accurate ground truth, called *PanoraMIS*. The dataset covers a large spectrum of panoramic cameras (catadioptric, twin-fisheye), robotic platforms (wheeled, aerial and industrial robots), and testing environments (indoors and outdoors), and it is well suited to rigorously validate novel image-based robot-motion estimation algorithms, including visual odometry, visual SLAM, and deep learning-based methods. PanoraMIS and the accompanying documentation is publicly available on the Internet for the entire research community.

Keywords
Panoramic cameras, omnidirectional vision, visual-inertial odometry, image-based localization, autonomous robots

1 Introduction

1.1 Motivation and related work

Panoramic or 360° imaging systems have become increasingly popular in the computer vision and robotics communities. In fact, their ability to capture the appearance of a complex scene by a single shot, is extremely attractive for 3-D reconstruction of an unknown environment and for mobile robot navigation. Over the last two decades, a solid theoretical foundation for processing and interpreting spherical images coming from ultra-wide field of view (FoV) sensors has been laid Geyer and Daniilidis (2001); Barreto (2006); Benosman and Kang (2013). The need for a common ground to compare the performance of new vision-based ego-motion estimation algorithms under different operating conditions, has also stimulated continued interest in image datasets and benchmarks. However, to the best of our knowledge, only few databases dedicated to panoramic images and not tailored to a specific application, are currently available for robotic research.

In this paper, we describe a *new* image collection, called PanoraMIS, that our group has recently created to fill this vacuum. PanoraMIS presents several original and distinctive features. It includes real images coming from two ultra-wide FoV sensors mounted on a variety of robot carriers. The image sequences have been recorded indoors and outdoors in man-made and natural environments, and come with accurate time-synchronized ground truth poses from odometry, IMU (Inertial Measurement Unit) and GPS. We release PanoraMIS in an effort to foster reproducibility, and help researchers to rigorously validate and compare new ego-motion estimation and navigation algorithms based on panoramic images. Target applications include (but are not limited to) visual-inertial odometry Li and Mourikis (2013); Matsuki et al (2018), visual SLAM Cadena et al (2016), and visual navigation based on convolutional neural networks Goodfellow et al (2016).

The vast majority of existing robot-vision datasets, includes perspective images taken either by monocular or stereo cameras. In what follows, we will briefly review the most relevant ones for PanoraMIS, with special emphasis on the few databases which also include panoramic images. For the reader's convenience, a summary of the main features of these datasets is reported in Table 1.

Large-scale image collections of dynamic urban environments have been recently created for research on automated or self-driving vehicles. In all cases, car-borne cameras have been installed on the roof or inside the cockpit. For instance, the Ford Campus dataset Pandey et al (2011) provides data acquired by a professional and consumer-grade IMU (with integrated GPS receiver), a Velodyne 3-D LiDAR, two push-broom forward-looking RIEGL LiDARS, and a Point Grey Ladybug 3 omnidirectional camera. The sensor suite was mounted on a pickup truck that the authors drove in the Ford Research Campus and in downtown Dearborn, Michigan. The unprocessed data consist of three main files (containing the omnidirectional images, timestamps, and data from the remaining sensors) for each trial. The provided 3-D point cloud information can also be used for semantic laser/vision labeling or 3-D object recognition.

Similarly to the Ford Campus dataset, AMUSE, proposed in Koschorrek et al (2013), includes synchronized data

The authors are with the MIS laboratory, Université de Picardie Jules Verne, Amiens, France.

Corresponding author:

*F. Morbidi, Robotic Perception group, MIS laboratory, Université de Picardie Jules Verne, 33 rue Saint-Leu, 80039 Amiens Cedex, France.
Email: fabio.morbidi@u-picardie.fr

Table 1. Overview of related image datasets compared to PanoraMIS (the classification is by type of carrier and year of publication).

Dataset	Camera(s)	Carrier	Environment, 3-D point cloud	Sensors, ground truth	Stats
Ford Campus, Pandey et al (2011)	Omnidirectional multi-camera	Car	Outdoors, Yes	3-D/2-D LiDAR, IMU, GPS	2 seqs, ≈ 10 km
AMUSE, Koschorrek et al (2013)	Omnidirectional multi-camera	Car	Outdoors, No	IMU, GPS, height sensors	7 seqs, 24.4 km
KITTI, Geiger et al (2013)	Stereo	Car	Outdoors, Yes	IMU, GPS	22 seqs, 39.2 km
IPDS, Korrapati et al (2013)	Perspective, fisheye, catadioptric	Electric cart	Outdoors, No	IMU, GPS	6 seqs, 18 km
Málaga Urban, Blanco-Claraco et al (2014)	Stereo	Car	Outdoors, Yes	IMU, GPS, LiDAR	15 seqs, 36.8 km
Oxford RobotCar, Maddern et al (2017)	Fisheye, trinocular stereo	Car	Outdoors, Yes	3-D/2-D LiDAR, IMU, GPS	1010 km
Complex Urban, Jeong et al (2019)	Stereo	Car	Outdoors, Yes	3-D/2-D LiDAR, IMU, GPS, 3-axis optic gyro	19 seqs, 191 km
New College, Smith et al (2009)	Omnidirectional multi-camera, stereo	Wheeled robot	Outdoors, Yes	IMU, GPS, 2-D LiDAR	9 seqs, 2.2 km
Rawseeds, Ceriani et al (2009)	Trinocular, perspective, catadioptric	Wheeled robot	In/Outdoors, No	IMU, GPS, 2-D LiDAR, odometry	Up to 1.9 km
Marulan, Peynot et al (2010)	Perspective	Wheeled robot	Outdoors, No	2-D LiDAR, IR camera, radar	40 seqs
TUM RGB-D, Sturm et al (2012)	RGB-D	Wheeled robot, hand held	Indoors, Yes	MoCap	39 seqs
UMich NCLT, Carlevaris-Bianco et al (2016)	Omnidirectional multi-camera	Segway robot	In/Outdoors, Yes	IMU, GPS, 3-D/2-D LiDAR	27 seqs, 147.3 km
EuRoC MAV, Burri et al (2016)	Stereo	Hexarotor MAV	Indoors, Yes	MoCap, IMU	11 seqs, 0.9 km
Zurich Urban MAV, Majdik et al (2017)	Perspective	Quadrotor MAV	Outdoors, Yes	IMU, GPS	1 seq, 2 km
LaFiDa, Urban and Jutzi (2017)	Trinocular fisheye	Helmet	In/Outdoors, No	MoCap, 2-D LiDAR	6 seqs, 190 m
PennCOSYVIO, Pfrommer et al (2017)	Perspective, fisheye, stereo	Hand-held rig	In/Outdoors, No	IMU, visual tags	4 seqs, 0.6 km
TUM VI, Schubert et al (2018)	Stereo w/ fisheye lenses	Hand-held rig	In/Outdoors, No	IMU, MoCap	28 seqs, 20 km
Multi-FoV, Zhang et al (2016)	Perspective, fisheye, catadioptric	None	In/Outdoors	Synthetic data	2 seqs
PanoraMIS (ours)	Catadioptric, twin-fisheye	Wheeled, aerial, industrial robots	In/Outdoors, Yes	IMU, GPS, odometry	8 seqs, 5.6 km

streams recorded by a car equipped with an omnidirectional camera (a Point Grey Ladybug 3), laser-based height sensors, an IMU, a velocity sensor and a GPS receiver, in different testing environments. The KITTI suite Geiger et al (2013) is well-known in the computer vision community. It provides ground truth data for visual odometry, stereo reconstruction, object detection and tracking, optical flow data, and object labels in the form of 3-D tracklets. The Institut Pascal Data Sets (IPDS) by Korrapati et al (2013), consist of multi-sensory data acquired by an electric cart navigating in the Cézeaux university campus in Clermont-Ferrand, France. The video sequences were captured by a standard perspective camera, three synchronized non-overlapping cameras, and

two large FoV color cameras (fisheye and catadioptric). The ground truth is provided by real-time kinematic GPS. In the Málaga Urban Dataset Blanco-Claraco et al (2014), a car was outfitted with a stereo camera and five laser scanners (among other sensors). The dataset, which was designed for large-scale visual navigation applications, contains a single trajectory with centimeter-accuracy GPS ground truth, and high-resolution, high-rate (20 fps) stereo images. A similar setup (including a stereo camera) in a complex urban environment, was recently considered in Jeong et al (2019). The Oxford RobotCar dataset Maddern et al (2017) was collected by a car operated along the same route in Oxford, England, over the course of a year and over 1000 km.

The goal was to provide a good representation of variations in seasons and weather, to enable research investigating long-term localization and mapping for autonomous vehicles. The dataset consists of almost 20 million images coming from six cameras mounted on the car. In particular, the combination of a Point Grey Bumblebee XB3 trinocular stereo camera and three monocular Point Grey Grasshopper2 cameras with fisheye lenses, provides a full 360 degree visual coverage of the scene around the vehicle. Three 2-D/3-D LiDARs and an inertial and GPS navigation system, complete the equipment of the car.

In spite of the large prevalence of car-mounted cameras for image collection, wheeled robots can be found in the literature as well. For example, the NewCollege dataset Smith et al (2009) provides a sequence of panoramic views obtained by stitching together multiple images captured by a Point Grey Ladybug 2 camera. It also includes laser range/reflectance data, 20 Hz stereoscopic imagery and GPS measurements (not available in all scenarios). The timestamped data have been gathered by a mobile robot driving several kilometers through a wooded park and a campus, and it has been widely used for mobile robotic research (e.g. visual odometry, visual topological SLAM). The Rawseeds dataset Ceriani et al (2009) includes sequences of images captured by a trinocular, perspective and catadioptric camera along with the relative IMU, real-time kinematic GPS, 2-D laser (two Hokuyo and two SICK lasers) and sonar information. The ground truth covers both indoor and outdoor environments, and the presence of trajectories with closed loops makes it well suited to evaluate visual SLAM algorithms. In the Marulan dataset Peynot et al (2010), the eight-wheel skid-steering Argo platform was equipped with a monocular color camera together with four 2-D SICK laser scanners, a frequency-modulated continuous-wave radar and a thermal infrared camera. The TUM RGB-D dataset Sturm et al (2012) provides images from a single RGB-depth camera in an office environment and an industrial hall, with an accurate ground truth from a Motion Capture (MoCap) system. The dataset has been widely exploited for the evaluation of RGB-D SLAM and visual odometry algorithms. The North Campus Long-Term (NCLT) dataset Carlevaris-Bianco et al (2016), consists of omnidirectional imagery (Point Grey Ladybug 3 camera), 3-D LiDAR (Velodyne), planar LiDAR (Hokuyo), consumer grade GPS, real-time kinematic GPS, and proprioceptive sensors (IMU, optic gyro) for odometry, collected using a Segway robotic platform on the University of Michigan's North Campus. As for KITTI and TUM RGB-D, a 3-D point cloud of the environment is also provided in the dataset.

As far as aerial robots are concerned, the EuRoC MAV datasets Burri et al (2016) contain sequences of hardware synchronized images captured by a hexarotor Micro Aerial Vehicle (MAV) flying in an indoor cluttered environment, using two stereo cameras. The 6-D pose of the hexarotor is tracked with a Vicon MoCap system and the on-board IMU data, together with the 3-D point clouds of the environment, are available as well. Another dataset for visual-inertial SLAM is the Zurich Urban MAV Majdik et al (2017). The dataset was recorded by camera on-board a MAV flying at low altitudes (5 to 15 m above the ground) in the urban streets of Zurich, Switzerland. The 2 km dataset consists of time-synchronized aerial high-resolution images, GPS and IMU sensor data, and ground-level street view images.

Other unconventional carriers have been used in some recent image datasets. LaFiDa Urban and Jutzi (2017) was recorded by combining a head-mounted trinocular-fisheye camera with a 2-D Hokuyo laser scanner, and the 6-D pose ground truth information is provided by a MoCap (Opti-Track) system. The multi-fisheye camera has been calibrated and a single omnidirectional image is provided as output to the user. The PennCOSYVIO dataset Pfrommer et al (2017), a visual-inertial odometry/SLAM benchmark, was acquired indoors/outdoors with a hand-held rig. It contains synchronized data from a stereo camera and IMU, two Project Tango hand-held devices, and three GoPro Hero 4 cameras. The TUM VI dataset, described in Schubert et al (2018), contains image sequences captured in different indoor and outdoor scenarios for evaluating visual-inertial odometry. Data collection was performed using two monochrome cameras in a stereo setup endowed with fisheye lenses (diagonal FoV, less than 195°) and a microcontroller board with integrated 3-axis IMU, mounted on a hand-held rig. A light sensor between the cameras also delivers an approximate measurement of the illuminance of the environment (for photometric calibration). The stereo camera (working at 20 Hz) and IMU sensor are time-synchronized in hardware, and for trajectory evaluation, an accurate pose ground truth is provided by a MoCap at high frequency (120 Hz).

Finally, in Zhang et al (2016), photorealistic synthetic images have been created with the Cycles raytracing engine implemented in Blender. In particular, three sequences of images of an indoor environment and of outdoor urban environments have been generated using three simulated camera models (perspective, fisheye and catadioptric). The ground truth is available in all cases.

1.2 Original contributions and organization

This paper documents *PanoraMIS*, a new image dataset collected with ultra-wide FoV cameras (catadioptric and twin-fisheye) mounted on wheeled, aerial and industrial robotic platforms. The time-registered dataset, which comprises eight indoor/outdoor sequences (upwards of 13000 images and 16.2 GB), has been created to provide a new challenging testing ground for image-based robot localization thanks to the large variety of urban/natural scenarios, and it responds to the scarcity of generic omnidirectional-vision benchmarks in the literature. In particular, each sequence has been designed to target a specific problem: vision-based heading-angle estimation (Sequences 1, 2, 3 and 5), vision-based attitude estimation (Sequence 4), visual SLAM (Sequences 6, 7 and 8), and visual-inertial odometry (Sequences 7 and 8). In summary, PanoraMIS has at least the following three unique features:

- First dataset to provide calibrated *dual-fisheye images* (Ricoh Theta S camera). This contrasts with the widespread Ladybug-like multi-camera systems, and alleviates the panoramic image-stitching problem,
- Availability of *decoupled* translational and rotational motions with precise ground truth (sub-millimeter/sub-degree 6-D pose accuracy in the case of the Stäubli TX60 robot),

- Data have been collected by keeping in mind the requirements of state-of-the-art visual SLAM and visual odometry algorithms, thus covering several *small* (a few tens of meters) and *large* (up to 1 km) *loops*.

The dataset includes GPS, IMU, and odometry measurements, which are time-synchronized in software with the omnidirectional RGB images. Dense *3-D point clouds* of the environment surrounding the robot, obtained from the dual-fisheye images by photogrammetry, are also available for Sequence 7.

PanoraMIS does not provide a benchmark to evaluate the performance of vision-based localization algorithms. However, in Sequences 6, 7 and 8, the users are recommended to adopt the Absolute Trajectory Error (ATE) and the Relative Pose Error (RPE), which are popular metrics to assess the quality of an estimated trajectory with respect to the ground truth (see Geiger et al (2013); Pfrommer et al (2017); Schubert et al (2018) and the references therein).

This paper builds upon our previous works Morbidi and Caron (2017) and Caron and Morbidi (2018), where part of the omnidirectional images of PanoraMIS were used to validate a new direct visual compass and gyroscope, respectively. It is finally worth pointing out here that the collection of the outdoor sequences was made particularly challenging by the spherical nature of the images: in fact, unable to hide from the camera, we did our best to minimize the impact of our presence and that of moving cars and pedestrians, on the images.

The remainder of this paper is organized as follows. Sect. 2 provides details about the omnidirectional cameras and robotic platforms used in our dataset. In Sect. 3, the data collection process and the eight sequences of PanoraMIS are presented. Finally, Sect. 4 is devoted to data organization and format, and Sect. 5 draws conclusions.

The full dataset and the supplementary material (videos, code, 3-D point clouds) are publicly available at the address: mis.u-picardie.fr/~panoramis

2   Material: cameras and robots

2.1   Catadioptric camera

The catadioptric camera consists of a VStone VS-C450MR-TK objective screwed on top of an IDS uEye UI-3370CP-C-HQ CMOS camera. The objective (diameter 59 mm, height 136 mm) includes a hyperbolic mirror, which allows to capture ultra-wide FoV images when combined with the perspective camera (see Fig. 1(a)): in fact, the catadioptric system has a horizontal FoV of $360°$, a vertical FoV of $75°$ ($15°$ "upper side" + $60°$ "lower side"), and the RGB uEye camera has a resolution of 620 pixels × 620 pixels, fitted to mirror border. For the *intrinsic calibration* of the catadioptric camera, we leveraged Barreto's unified central projection model (Barreto 2006) and used HySCaS (Caron and Eynard 2011). This yielded the calibration matrix,

$$\mathbf{K} = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 233.1999 & 0 & 304.7138 \\ 0 & 233.7360 & 308.7574 \\ 0 & 0 & 1 \end{bmatrix},$$
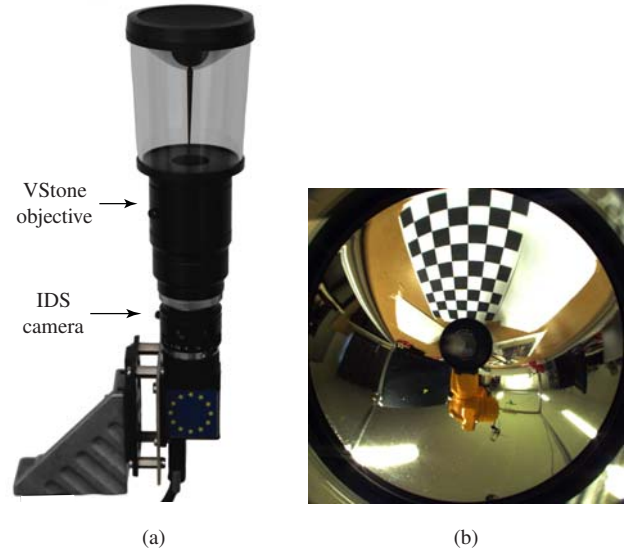

Figure 1. *Catadioptric camera:* (a) The camera consists of a VStone objective with hyperbolic mirror screwed on an IDS uEye RGB camera; (b) The first of the six images of the checkerboard pattern used for the intrinsic calibration of the camera (the camera is held by an industrial robot).

and the related mirror-shape parameter $\xi = 0.9701$. In the matrix above, α_u and α_v denote the focal lengths in pixels in the horizontal and vertical direction, respectively, and (u_0, v_0) are the coordinates of the principal point in pixels. The image coordinates of the corners of a checkerboard pattern observed by the camera at six different locations are provided in PanoraMIS. The first of these images is shown in Fig. 1(b).
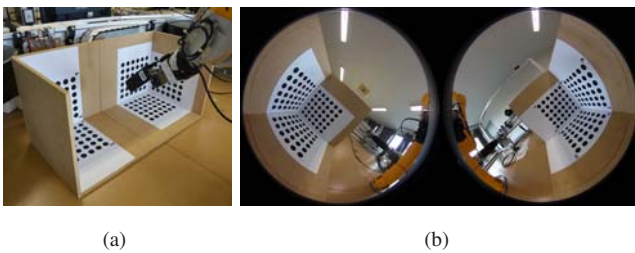
2.2   Twin-fisheye camera

The *Ricoh Theta S* is an innovative omnidirectional camera consisting of a twin-lens folded optical system coupled with two photosensitive sensors (1/2.3" 12 MP CMOS). The two fisheye lenses point in opposite directions and provide a spherical FoV (see Fig. 2(a)). Each image generated by the Theta S (1280 pixels × 720 pixels), contains a pair of fisheye images corresponding to opposite hemispheres (max frame rate, 30 fps). The camera weighs 125 g, it has an autonomy of 25 minutes with its internal battery (continuous shooting), and its dimensions are 44 mm (width) × 130 mm (height) × 22.9 mm (depth).

Two sets of *intrinsic parameters* are considered, one for each "fisheye camera", $\mathcal{P}_{c_j} = \{\alpha_{u_j}, \alpha_{v_j}, u_{0_j}, v_{0_j}, \xi_j\}$, $j \in \{1, 2\}$, where α_{u_j} and α_{v_j} are the focal lengths in pixels in the horizontal and vertical direction, respectively, (u_{0_j}, v_{0_j}) are the coordinates of the principal point in pixels, and ξ_j is the distance between the unit sphere's first projection center and the perspective second projection center of fisheye camera j (Barreto 2006). To guarantee the single-viewpoint property, we assumed that the translation vector between the fisheye-camera frames $\mathcal{F}_{c_1}$ and $\mathcal{F}_{c_2}$ is *zero* (a non-zero baseline is shown in Fig. 2(b) for ease of illustration), and that the frame $\mathcal{F}_c$ of the Theta S coincides with $\mathcal{F}_{c_1}$. The extrinsic parameters of the Theta S, which describe the orientation between the optical axes

(a)

(b)
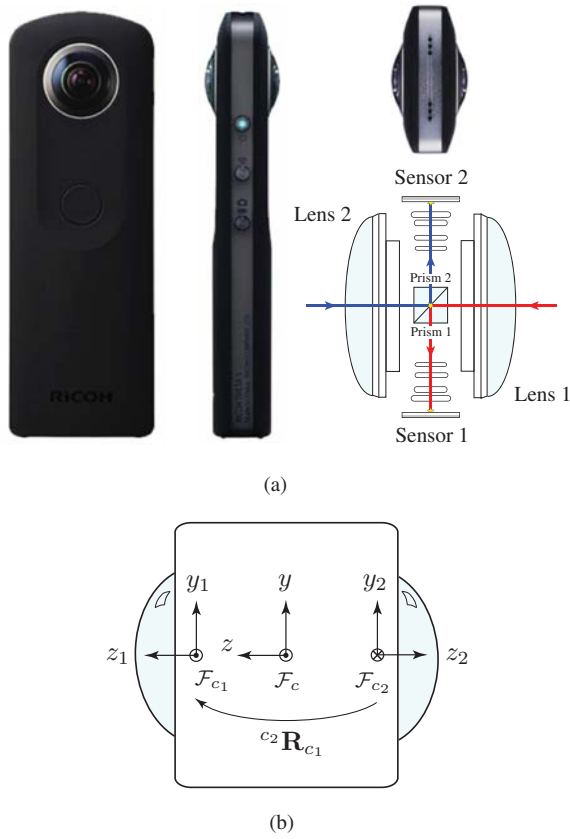
Figure 2. *Twin-fisheye camera:* (a) Front, side and top view of the Ricoh Theta S camera (image courtesy of Ricoh), and schematic of the optical system with the two fisheye lenses, prisms and CMOS sensors; (b) Coordinate frames of the camera and fisheye lenses (top view).

of the two fisheye cameras, are given by $\mathbf{r}_{1,2}$ (axis-angle representation). Rodrigues' formula is used to retrieve the rotation matrix $^{c_2}\mathbf{R}_{c_1}$ between the reference frames of the fisheye cameras, corresponding to the unit vector and rotation angle in the axis-angle representation. Overall, ten intrinsic and three extrinsic camera parameters have then been simultaneously estimated. For the calibration, the Theta S was modeled as a fisheye stereo camera with a single center of projection (Barreto 2006). The corners extracted from multiple images of a calibration rig consisting

(a)   (b)

Figure 3. *Calibration of the twin-fisheye camera:* (a) The calibration rig consists of six checkerboard patterns attached to the inside surface of two half cubes; (b) Sample dual-fisheye image of the calibration rig.

of six checkerboard patterns glued inside two half cubes, were used as input (see Fig. 3(a)). These images (one of which is shown in Fig. 3(b)) are available to the end-users of PanoraMIS. A variant of the calibration algorithm proposed in Caron and Eynard (2011) was employed: we set the translation parameters between the two fisheye lenses to zero and estimated the rotation parameters only (see Caron and Morbidi (2018)). This yielded $\mathcal{P}_{c_1} = \{577.7741, 576.1130, 958.6632, 316.8989, 1.9878\}$, $\mathcal{P}_{c_2} = \{567.8953, 565.1663, 321.5507, 319.4833, 1.9392\}$, and $\mathbf{r}_{1,2} = [-0.0082, 3.1319, -0.0108]^T$ rad.

2.3 Industrial robot

The *Stäubli TX60* is a 6-axis industrial robot with anthropomorphic arm and spherical wrist (see Fig. 4). With a nominal payload of 3.5 kg, the robot guarantees small positioning error and high repeatability ($\pm$ 0.02 mm). The price to pay for the accurate odometry (our ground truth) is the relatively small work envelope: the maximum reach of the robot between axis 1 and 6, is 670 mm. The manipulator, equipped with a CS8C controller, is located in a laboratory room of size 10.05 m $\times$ 7.03 m $\times$ 2.70 m (cf. Sequences 1 and 3 in Sect. 3.1 below).

2.4 Wheeled robots

The *Pioneer 3-AT*, manufactured by Adept MobileRobots[1], is a four-wheel programmable mobile robot intended for research and education (see Fig. 5(a)). The platform measures 50.8 cm in length, 38.1 cm in width and 27.7 cm in height, and weighs 12 kg. The wheels are driven by four reversible DC motors in a skid-steering drive mechanism, which allows zero-radius turning (maximum forward speed, 0.7 m/s). Each motor is equipped with a high-resolution optical quadrature shaft encoder which provides 34000 counts per revolution for use in the odometry module. The robot has an integrated single-axis (yaw) gyroscopic sensor for increased rotational accuracy, and the controller automatically incorporates the gyroscopic corrections into the odometry-based estimates (from wheel encoders).

Figure 4. *Industrial robot:* Stäubli TX60 inside its cell.

Table 2. Summary of the main features of the 8 sequences in PanoraMIS. The last column reports the target problems of each sequence.

Sequ.	Camera	Robot	Environ.	Ground truth	Calib. data	Stats	Target problems
1	Catadioptric	Industrial	Indoors	Odometry	6 images	3600 images, 2.27 GB, 1-axis rotation	Heading-angle estimation
2	Catadioptric	Wheeled	Indoors	Odometry	8 images	158 images, 66.3 MB, pure rotation	Heading-angle estimation
3	Twin-fisheye	Industrial	Indoors	Odometry	6 images	720 images, 628 MB, 1-axis rotation	Heading-angle estimation
4	Twin-fisheye	Industrial	Indoors	Odometry	6 images	94 images, 83.7 MB, 3-axis rotation	Attitude estimation
5	Catadioptric	Wheeled	Outdoors	Odometry	8 images	156 images, 79.3 MB, pure rotation	Heading-angle estimation
6	Catadioptric	Wheeled	Outdoors	Odometry	8 images	318 images, 155 MB, figure-of-eight trajectory	Visual odometry, visual SLAM
7	Twin-fisheye	Wheeled	Outdoors	Odometry, IMU, GPS	None	L1 : 5442 images, 5.33 GB, 3-D point cloud, 2.41 GB, L2 : 2578 images, 2.76 GB, 3-D point cloud, 1.52 GB	Visual-inertial odometry, visual SLAM
8	Twin-fisheye	Aerial	Outdoors	IMU, GPS	None	6 min 41 s video, 769 MB	Visual-inertial odometry, visual SLAM

(a)   (b)

Figure 5. *Wheeled robots:* (a) Pioneer 3-AT robot equipped with the catadioptric camera; (b) Seekur Jr robot outfitted with the twin-fisheye camera (the external IMU is not shown in the figure and the forward-facing SICK laser range-finder was not used).

The *Seekur Jr* is also a skid-steering four-wheel drive robot manufactured by Adept MobileRobots (see Fig. 5(b)). It was designed to have all the qualities of an all-terrain and all-weather mobile platform, which allows operation on steep slopes and over rough terrain. Equipped with a Trimble AG-372 GNSS receiver, a forward-facing SICK laser range-finder (not used), and an integrated 6-DOF IMU (Analog Devices ADIS16362 *i*Sensor), the platform measures 105 cm in length, 84 cm in width and 50 cm

in height, weighs 77 kg and it has a maximum forward speed of 1.2 m/s. The on-board odometry module is analogous to that on the Pioneer 3-AT. In order to have redundant inertial measurements, we installed an external IMU at the top front of the robot: the Xsens MTi is a miniature, gyro-enhanced Attitude and Heading Reference System (AHRS), which provides drift-free 3-D orientation, calibrated 3-D acceleration, 3-D rate of turn, and 3-D earth-magnetic field data. It has a static orientation accuracy of less than $0.5°$ (roll, pitch) and of less than $1.0°$ (yaw), and a dynamic orientation accuracy of $2°$. The absolute position accuracy of the GNSS receiver with OmniSTAR HP correction, is 10 cm after convergence (provided that at least 5 satellites are visible and that PDOP < 4, cf. Sect. 4.4).

2.5 Aerial robot

The *Parrot Disco FPV* is a fixed-wing unmanned aerial vehicle powered by a brushless DC motor with a folding bi-blades propeller (see Fig. 6). It has a wingspan of 1.15 m,

Figure 6. *Aerial robot*: The twin-fisheye camera was mounted on the Disco using a custom-made expanded polystyrene hull.
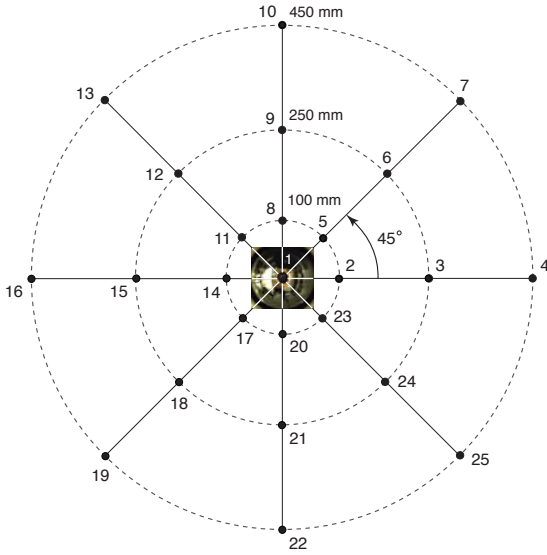
Figure 7. *Sequence 1:* Graphical illustration of the organization of the images acquired with the catadioptric camera on the Stäubli robot: 3600 images were taken at 25 CPs (black disks) lying on 8 rays with an angular spacing of $45°$.

it weighs 0.75 kg, it can reach a top speed of 80 km/h, and it offers 45 minutes of non-stop flight in low-wind speed conditions (i.e. less than 40 km/h). The drone is equipped with a front-facing fisheye camera with a FoV of $185°$, and it integrates an inertial navigation system (3-axis accelerometer, 3-axis gyroscope, 3-axis magnetometer), and a built-in GNSS module (GPS + GLONASS). The gyroscope and the magnetometer have a sensitivity of $0.1°/s$ and 1 μT, respectively, and the GNSS receiver delivers a 167 dBm sensitivity with a precision of $\pm$ 3 m.

3 Data collection

In the forthcoming subsections, each data collection exercise is referred to as a *Sequence*. PanoraMIS comprises 8 Sequences, 4 indoors and 4 outdoors, described in detail below (see Table 2 for a summary). To obtain reproducible results, in the sequences involving the wheeled robots, we synchronized the image acquisition system with the module that logs the pose estimates of the robot via the C++ ARIA library. A new measurement is recorded only if the robot has traveled a distance of at least 10 cm or if it has rotated at least $2°$.

3.1 Indoor environment

3.1.1 Catadioptric camera

Sequence 1: In this first sequence, the catadioptric camera was mounted on the end-effector of the industrial robot. To account for the geometric constraints (maximum arm extension) and kinematic singularities of the manipulator, we limited ourselves to camera poses within a horizontal disk of radius 450 mm around the vertical axis of the manipulator. To sample this disk, located 603.3 mm above the platform where the robot is mounted (see Fig. 7), we chose 25 Collection Points (CPs) on 8 rays with an angular spacing of $45°$: on each ray, three locations, at a distance of 100, 250 and 450 mm, were considered (see the black disks in Fig. 7). At

each CP, the camera was rotated by $360°$ about its optical axis z_c with a step size of $2.5°$, yielding 144 images per CP. For each of the $25 \times 144 = 3600$ PNG images, we recorded the 3-D coordinates of the origin of the camera frame $\{x_c, y_c, z_c\}$ and its orientation with respect to the base frame of the robot. We positioned the manipulator at the 25 CPs by running a VAL3 program (that behaves as a server) on the CS8C controller. Neither master gain nor gamma correction was considered for the camera to reduce image noise. However, to capture images having enough luminance, we set the exposure time to 200 ms, leading to a frame rate of 5 fps. Such a low frame rate was not an issue in practice, since the camera was stopped before every acquisition to avoid image blur. For the *extrinsic calibration* of the camera, let us denote by $^{c_i}\mathbf{T}_p \in \mathrm{SE}(3)$ the rigid transformation between the frame attached to a checkerboard pattern, "p", and the camera's frame at pose "c_i", $i \in \{1, \ldots, 6\}$. As in the first image of the sequence the camera is rigidly mounted on the industrial robot, the end-effector's pose corresponding to each of these camera poses should also be considered. Let $^{e_i}\mathbf{T}_b$ be the rigid transformation between the robot's base frame, "b", and the end-effector's frame at pose "e_i". Matrices $^{c_i}\mathbf{T}_p$ and $^{e_i}\mathbf{T}_b$, $i \in \{1, \ldots, 6\}$, were given as input to the Tsai & Lenz's algorithm (in the ViSP-library's implementation Marchand et al (2005)) for the estimation of the extrinsic camera parameters. The following transformation $^{e}\mathbf{T}_c$ between the camera's frame $\{x_c, y_c, z_c\}$ and the end-effector's frame $\{x_e, y_e, z_e\}$, was thus obtained (cf. Fig. 8):

$$^{e}\mathbf{T}_c = \begin{bmatrix} 0.0247 & 0.9988 & 0.0411 & -0.0646 \\ 0.9992 & -0.0234 & -0.0307 & -0.0024 \\ -0.0297 & 0.0418 & -0.9987 & 0.2206 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

where distances are in meters.
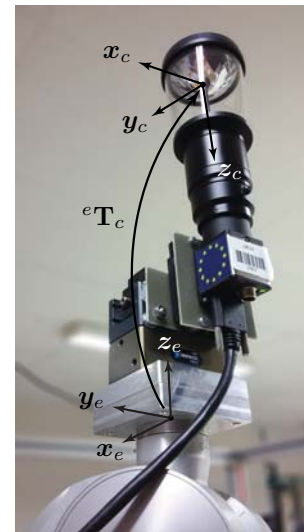
Figure 8. *Extrinsic calibration of the catadioptric camera*: The camera mounted on the end-effector of the industrial robot and coordinate frames superimposed upon it. Here $^{e}\mathbf{T}_c$ denotes the 3-D rigid transformation between the camera's frame $\{x_c, y_c, z_c\}$ and the end-effector's frame $\{x_e, y_e, z_e\}$.

(a)　　　　　　　　　　(b)
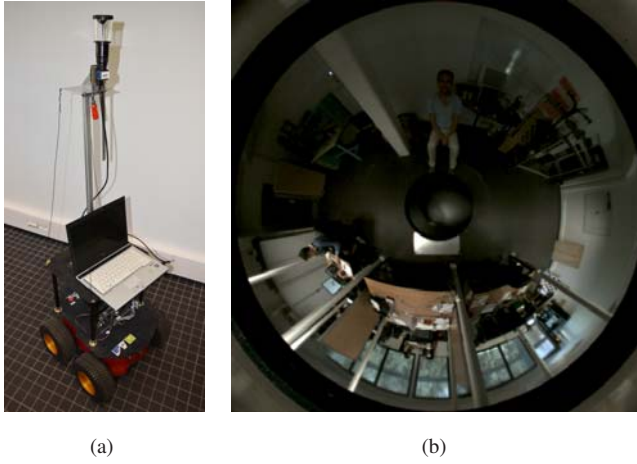
Figure 9. *Sequence 2:* (a) Pioneer 3-AT indoors; (b) First catadioptric image of the sequence.

Sequence 2: In this sequence, we outfitted the Pioneer 3-AT with the catadioptric camera, and considered a pure rotational motion about its vertical axis (see Fig. 9). We applied a constant angular velocity to the skid-steering robot in order to have a regular step. Overall, 160 images were acquired during 49.69 s. In order to have an elevated viewpoint, the camera was placed on top of a mast, 134 cm above the ground (110 cm above the robot). We used a two-axis bubble level to orient the optical axis of the camera perpendicularly to the ground plane. The exposure time of the camera was left free but limited to 50 ms, yielding

Figure 10. *Extrinsic calibration of the twin-fisheye camera:* The camera mounted on the end-effector of the industrial robot and coordinate frames superimposed upon it.

a frame rate of 20 fps. The catadioptric camera was re-calibrated for this sequence, leading to:

$$\mathbf{K} = \begin{bmatrix} 231.4617 & 0 & 319.7041 \\ 0 & 232.4225 & 310.9443 \\ 0 & 0 & 1 \end{bmatrix}, \quad \xi = 0.9583.$$

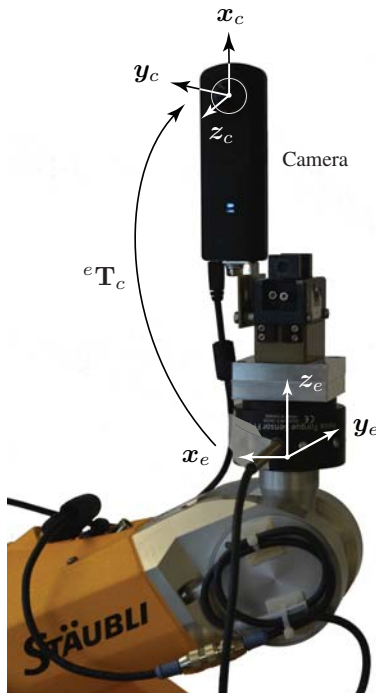The same intrinsic parameters were also used in Sequences 5 and 6 below.

3.1.2 Twin-fisheye camera

Sequence 3: The experimental setup considered for this sequence is similar to that reported in Fig. 7 for Sequence 1, but the catadioptric camera was replaced with the twin-fisheye camera. Five collection points (CP1, CP2, ..., CP5) located on the same line parallel to the ground were considered (the camera's center was 60 cm above the base of the robot, see Fig. 10). At each CP, the camera was rotated of 360° about the vertical axis, with a step size of 2.5°, yielding 144 images. The distance between CP1 and CP2, CP3, CP4 and CP5 is 40, 120, 240 and 400 mm, respectively. Again, we used the Tsai & Lenz's algorithm in the ViSP-library implementation, to extrinsically calibrate the camera with respect to the robot end-effector, i.e. to compute ${}^{e}\mathbf{T}_{c} \in$ SE(3) (see Fig. 10). The calibration rig was observed by the camera from six different poses (cf. Fig. 3(b)), yielding:

$$ {}^{e}\mathbf{T}_{c} = \begin{bmatrix} -0.0136 & 0.9997 & 0.0222 & -0.0401 \\ 0.0142 & -0.0220 & 0.9996 & 0.0000 \\ 0.9998 & 0.0139 & -0.0139 & 0.2372 \\ 0 & 0 & 0 & 1 \end{bmatrix}, $$

where distances are in meters.

Sequence 4: The material used to generate this sequence is the same as that presented in Sequence 3 (twin-fisheye camera mounted on the industrial robot): however, more involved 3-D rotations were considered for the camera. In fact, we took a single collection point of Sequence 3, CP3, and the robot was moved to obtain the maximum number of distinct 3-D orientations of the camera (94 overall), which also complied with the mechanical constraints of the robot.

(a)　　　　　　　　　　(b)
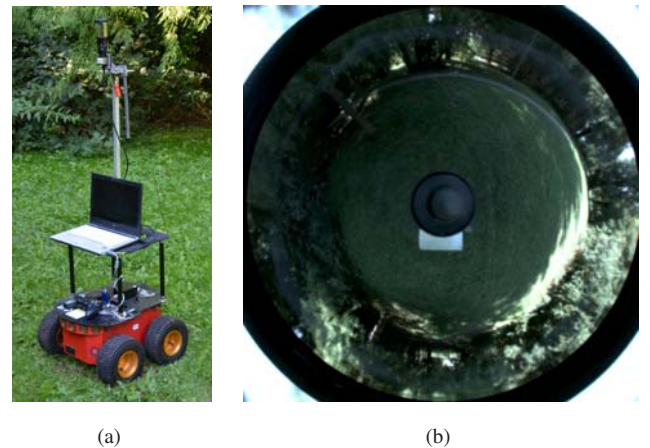
Figure 11. *Sequence 5:* (a) Pioneer 3-AT outdoors; (b) First catadioptric image of the sequence.
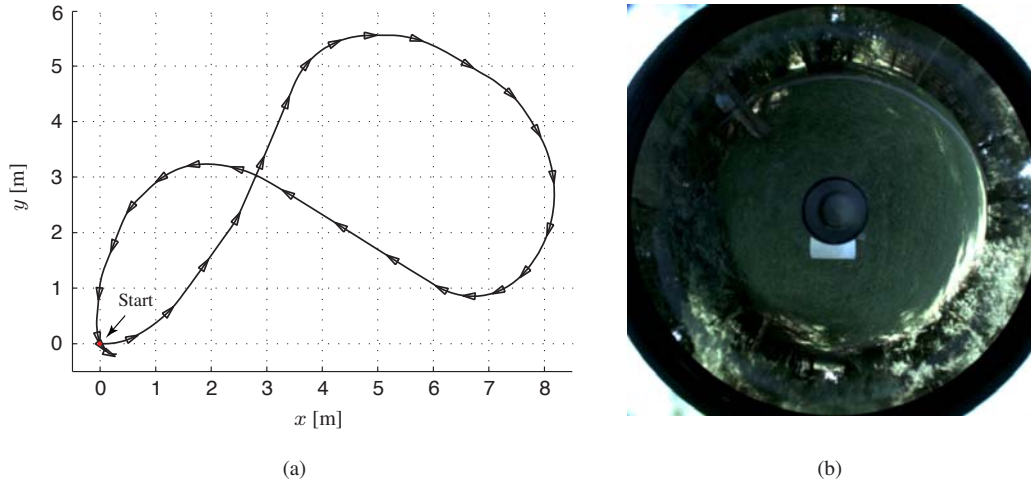
Figure 12. *Sequence 6:* (a) Figure-of-eight trajectory of the Pioneer 3-AT (the robot is marked with a triangle every 10 frames or about every 5 s); (b) First catadioptric image of the sequence.

At these locations, the camera was rotated about its z-axis with a step size of $45°$.

3.2 Outdoor environment

3.2.1 Catadioptric camera

Sequence 5: In this sequence, the catadioptric camera was mounted on the Pioneer 3-AT (cf. Sequence 2). The robot was manually driven with a joystick at Bois Bonvallet, a 4.5 ha public park in Amiens, France. The robot rotated about its vertical axis of $360°$ for 50.22 s, and the camera took 156 images (see Fig. 11 and cf. Sequence 2).

Sequence 6: For this sequence, we considered the same setup and outdoor scenario as Sequence 5 (cf. Fig. 5(a)). The robot was manually controlled with a joystick along a 25.27 m figure-of-eight trajectory (see Fig. 12(a)): its forward and angular velocities were limited to 0.2 m/s and 12 deg/s, respectively. The mean speed along the path was 0.187 m/s. During the 160.99 s trajectory, 318 images were acquired. Note that with the exception of a lamppost and a paved footpath, few recognizable landmarks are visible in the images, making Sequence 6 challenging for vision-based robot localization.

3.2.2 Twin-fisheye camera

Sequence 7: The motivation for this sequence came from the ongoing Interreg VA FCE ADAPT project, which aims at developing a smart and connected electrically-powered wheelchair. In recent years, a growing effort has been devoted to assistive robotic technologies (see e.g. Rebsamen et al (2010); Pasteau et al (2016); Podobnik et al (2017); Morales et al (2018) and the references therein), but the full potential of omnidirectional vision has not yet been fully exploited. In fact, a panoramic camera can improve the perceptual awareness of the wheelchair users and assist them in difficult tasks (e.g. negotiate a narrow doorway or enter an elevator). Having a similar footprint, in Sequence 7 we used the Seekur Jr equipped with the twin-fisheye camera to simulate the motion of a four-wheel electrically-powered wheelchair. As with the Pioneer 3-AT, the camera was

mounted on the top of a mast to have an elevated viewpoint of the environment around the Seekur Jr (150 cm above the ground). The mast was aligned with the vertical axis of the body frame of the robot, as shown in Fig. 5(b).

In order to have a large variety of scenarios for the simulated wheelchair, we considered both unstructured park areas including sidewalk ramps and a narrow timber bridge (Bois Bonvallet, length 500 m) and urban environments (Quai de la Somme, Rue de l'Orée du Bois, a parking lot, for a total distance of 450 m). The full trajectory of the Seekur Jr, which was manually driven with a joystick, is shown in Fig. 13(a) (total time: 21 minutes). Figs. 13(b), 13(c) report 6 snapshots from the data collection campaign. Between waypoints 3 and 4, the reception of GPS signals was locally perturbed by the dense vegetation around the Seekur Jr, which resulted in an inaccurate localization of the robot (see Fig. 14, which corresponds to location B in Fig. 13(a)). However, in the interest of clarity, we did not manually correct the data (which correspond to 13% of the overall trajectory), and the *raw GPS values* are made available in the dataset. Local occlusions in the images due to passing pedestrians and cars, were also experienced. The Seekur Jr returned to the starting point (red disk): hence the 950 m trajectory contains a loop closure. The dataset consists of time-synchronized dual-fisheye images, and GPS, odometry and IMU data recorded by the robot (the external Xsens MTi IMU is the orange box visible in Fig. 13(d)). One of the dual-fisheye images captured at location A, is shown in Fig. 13(e). The forward and angular velocity of the robot along the trajectory was limited to 0.7 m/s and 10 deg/s, respectively.

Agisoft Metashape was used for the 3-D reconstruction of the environment surrounding the robot, by spherical photogrammetry (see Fig. 15(a)). In fact, one every two images of the sequence (in equirectangular format), together with the corresponding GPS coordinates for robust initialization and fast convergence, was given as input to Metashape to generate an accurate and dense 3-D cloud of 89 million points (PLY format). A mask containing a fixed part (the robot) and a variable one (the sky), was applied to the images. As a comparison, Fig. 15(b)
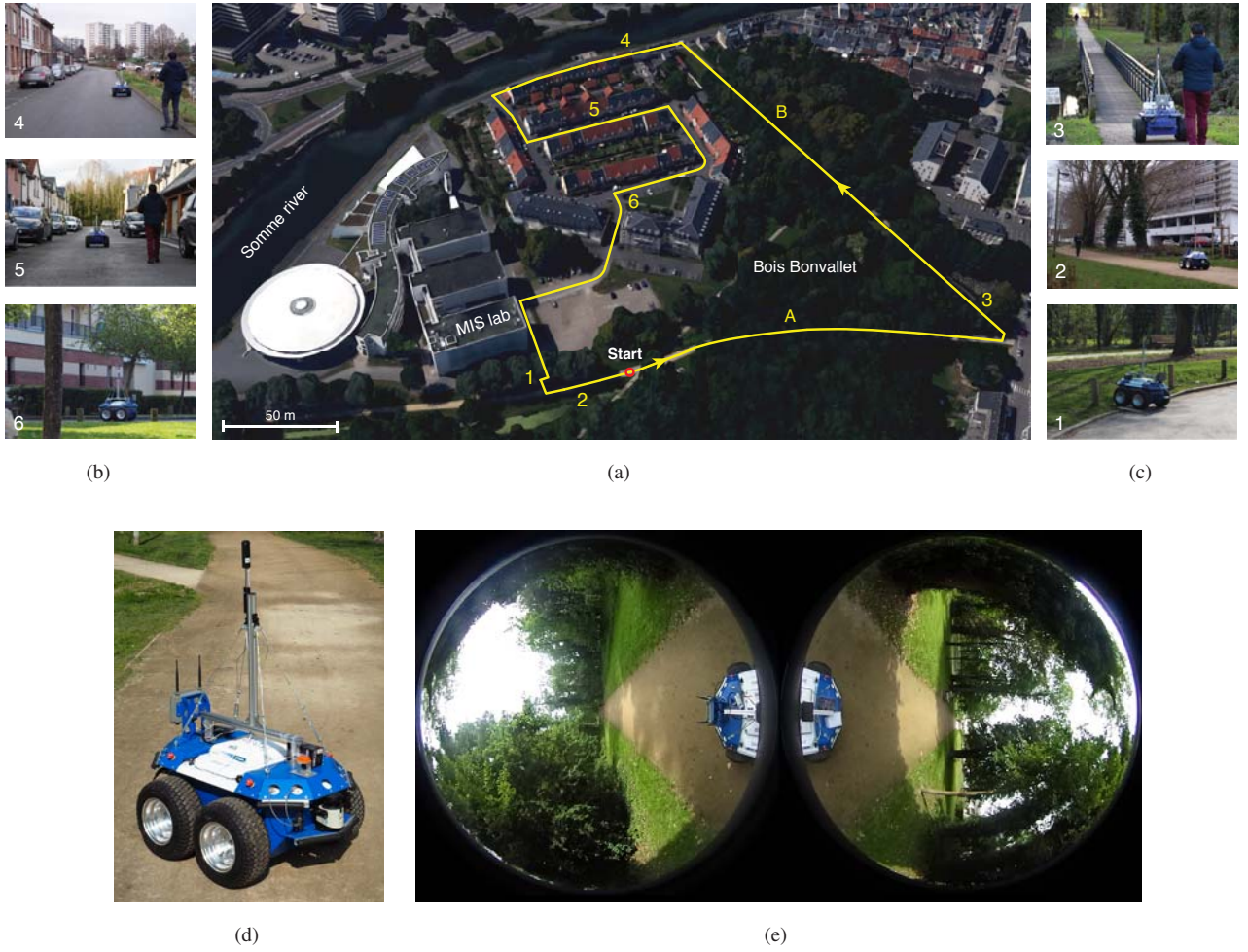
Figure 13. *Sequence 7*: (a) An overview of the trajectory (L1) followed by the Seekur Jr, overlaid atop of an aerial image from Google Maps; (b),(c) Six snapshots from the data acquisition campaign: snapshots 1 and 3 show the robot climbing a curb and crossing a narrow timber bridge, respectively, while snapshots 4-6 show the robot moving in an urban environment; (d) The fully-equipped Seekur Jr at the beginning of the trajectory (red disk in (a)); (e) Dual-fisheye image captured by the Theta S at location A.
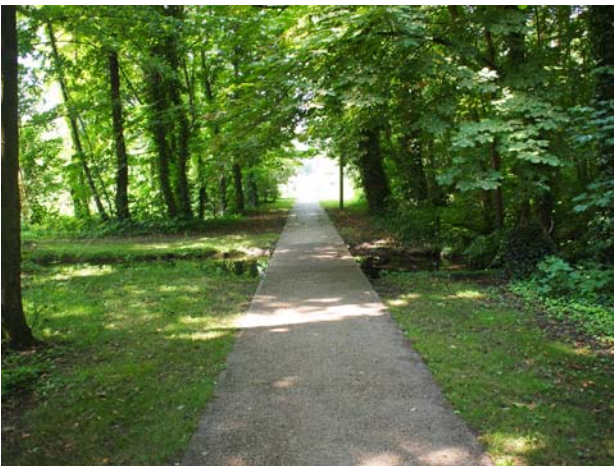
Figure 14. *Sequence 7*: GPS localization was inaccurate in some portions of the trajectory (fewer than 5 satellites tracked), because of the dense canopy ("green tunnel"). Here, the critical location B in Fig. 13(a), is shown.

reports the corresponding 3-D map and trajectory of the robot (green) obtained with OpenVSLAM (Sumikura et al 2019), an indirect visual SLAM approach based on fast binary descriptors and pose-graph optimization to minimize the drift accumulated along the trajectory. A video in equirectangular format, generated from the dual-fisheye images of Sequence 7, was the only input we provided to OpenVSLAM. We set the frame rate to 4 fps (default: 30 fps), and the resolution of the video to 1280 pixels $\times$ 520 pixels (to crop the area containing the chassis of the robot). For reference purposes, the data of a second smaller loop (total length 410 m) recorded the same day on the south-east side of Bois Bonvallet, are also included in Sequence 7 (the two loops are referred to as L1 and L2 in Table 2 and in the dataset website).

Sequence 8: Attitude estimation, localization and navigation algorithms for aerial vehicles could benefit from omnidirectional vision. In response to this need, PanoraMIS includes an MP4 video (30 fps) recorded by the twin-fisheye camera mounted on the Disco (see Fig. 6 and Fig. 16). We chose a front elevated position for the camera, in order to

(a)

(b)
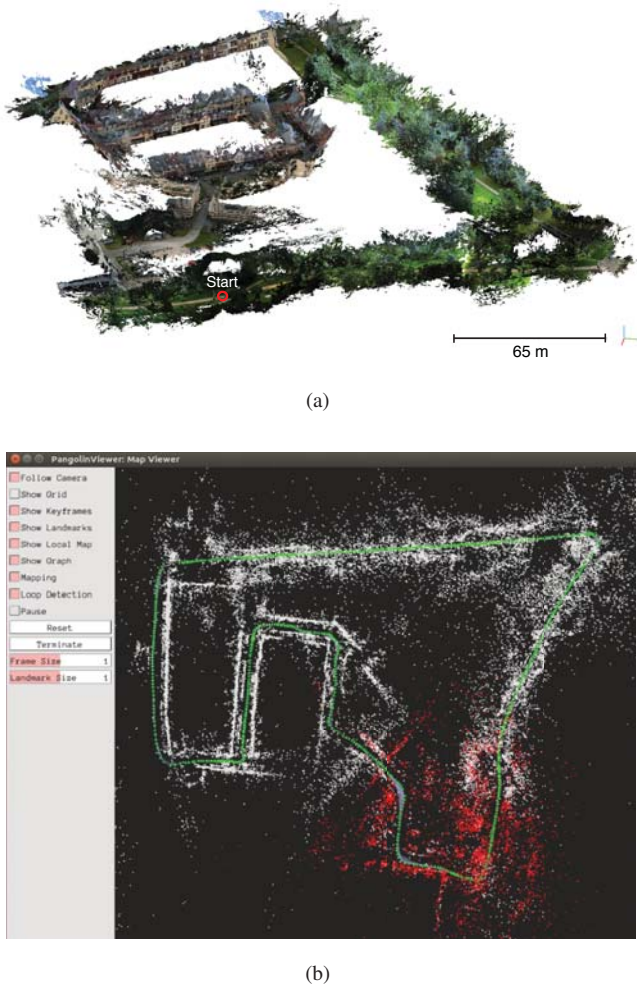
Figure 15. *Sequence 7: mapping of the environment*:
(a) 3-D point cloud obtained by spherical photogrammetry with
Agisoft Metashape; (b) 3-D map and robot trajectory (green)
obtained with OpenVSLAM (Sumikura et al 2019).

have as unobstructed a view of the surrounding environment
as possible, and to not overly perturb the center of mass of
the drone. The Disco was remotely controlled over an open
field, 10 km north of Amiens (Vaux-en-Amiénois), France.
It reached a maximum altitude of 106 m and a maximum
ground speed of 85.9 km/h. The total flight time was 6 min
41 s, the total distance travelled 4.2 km, and the major
axis of the trajectory measured 383.3 m (see Fig. 16(a)).
All the flight data (GPS, IMU, etc.) were recorded using the
on-board computer of the Disco.

4 Data organization

PanoraMIS is organized in folders and the directory layout
is shown in Fig. 17. Sequences 1-6 have the same data
structure, while Sequences 7 and 8 also include GPS and
IMU measurements. Sequences 1-6 have a folder called
"Calibration", containing the images used for calibration
as well as the estimated internal/external calibration
parameters. Each sequence is available for download as a zip
archive. We will briefly describe below the file formats and
data types.

4.1 Images

All the RGB images of PanoraMIS, whether captured by
the catadioptric or the twin-fisheye camera, have been
stored as lossless compressed PNG files. In Sequences 1, 3
and 4 (industrial robot), we considered the following naming
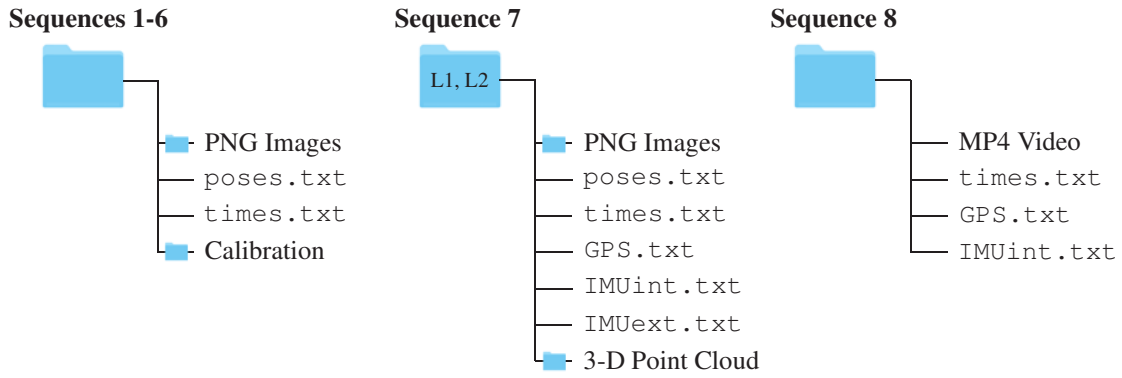convention for the images:

- 1st block of 6 digits: image number,
- 2nd block of 4 digits: angular direction (azimuth) of
  the camera in decidegrees (0-3600),
- 3rd block of 4 digits: angular direction (elevation) of
  the camera in decidegrees (0-3600),
- 4th block of 4 digits: distance from the origin of the
  reference frame at CP1 in millimeters,
- 5th block of 4 digits: rotation of the camera about its
  z-axis in decidegrees (0-3600),
- 6th block of 4 digits: rotation of the camera about its
  y-axis in decidegrees (0-3600),
- 7th block of 4 digits: rotation of the camera about its
  x-axis in decidegrees (0-3600).

On the other hand, in Sequences 2, 5, 6 and 7 (wheeled
robots), the images were labeled according to the following
convention:

(a)

(b)
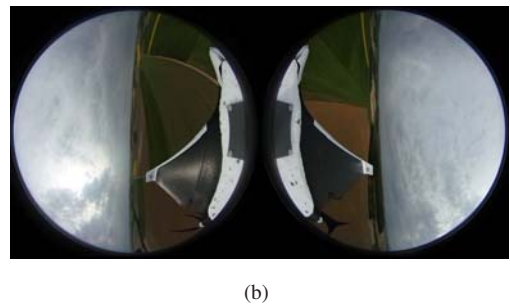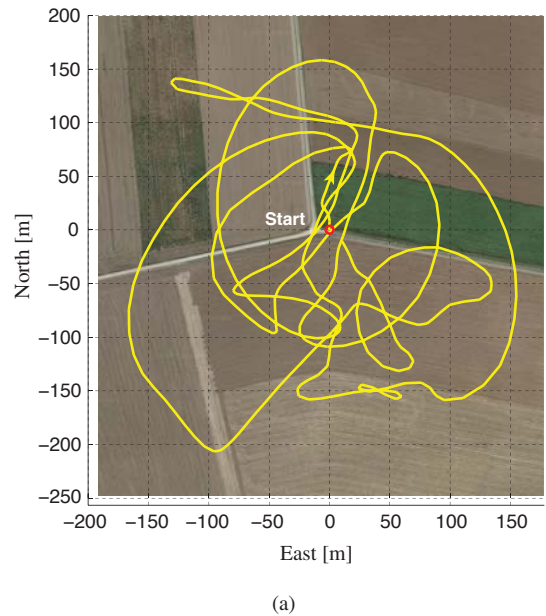
Figure 16. *Sequence 8*: (a) GPS trajectory of the Disco
overlaid atop of an aerial image from Google Maps (local
geodetic frame). The starting point of the drone is marked
with a red disk; (b) An airborne snapshot from the MP4 video.

Figure 17. *Directory layout of PanoraMIS:* Sequences 1-6 have the same data structure, while Sequences 7 and 8 also include GPS and IMU measurements.

- 1st block of 6 digits: image number,
- 2nd block of 4 digits: heading angle of the camera-robot in decidegrees (0-3600),
- 3rd block of 4 digits: distance between the initial and current position of the robot in centimeters,
- 4th block of 4 digits: rotation of the camera about its z-axis in decidegrees (0-3600).

More precisely, if $(x(t), y(t), \theta(t))$ denotes the position (centimeters) and orientation (degrees) of the mobile robot at time t provided by the odometry, the 2nd block of 4 digits corresponds to $10(180/\pi)\mathrm{atan2}(y(t), x(t))$, the 3rd to $\sqrt{x^2(t) + y^2(t)}$, and the 4th to $10\,\theta(t)$. Although this filename convention might seem cumbersome at first glance, it allows the user to easily determine when and where an image was taken without consulting the corresponding pose file (see Sect. 4.2).

4.2   Poses

The odometric data from the robotic platforms were stored in the text file `poses.txt`. Each line corresponds to a time-indexed measurement,

$$[\mathtt{x, y, z, Rotx, Roty, Rotz}],$$

where $[\mathtt{x, y, z}]$ denotes the x, y and z position of the camera in meters with respect to the reference frame of the base of the industrial robot, and $[\mathtt{Rotx, Roty, Rotz}]$ the attitude of the camera in radians (i.e. the rotation angles about the x-, y- and z-axes, respectively). For the wheeled robots, only $[\mathtt{x, y, Rotz}]$ is provided, i.e. the planar position and orientation of the robot with respect to the initial reference frame.

4.3   Times

The text file `times.txt` contains the timestamps in milliseconds provided by the clock of the computer on the robot (the time elapsed from the beginning of data collection, i.e. from $t = 0$ ms). The omnidirectional images and the data stream coming from the sensors on-board the robots have been time-synchronized in software.

4.4   GPS

GPS data are provided for Sequences 7 and 8 (in fact, only the Seekur Jr and the Disco are equipped with GNSS receivers) and stored in the text file `GPS.txt`. Each line in the GPS log file corresponds to a measurement over time. The GPS data from the Disco are organized as follows,

$$[\mathtt{latitude, longitude, number\text{-}of\text{-}satellites}],$$

where the `latitude` and `longitude` coordinates are given in degrees, and `number-of-satellites` indicates the instantaneous number of satellites tracked (generally between 5 and 9, and up to 15). The GPS information on the Seekur Jr is richer, and the following additional fields are available: `altitude`, the altitude of the robot in meters, `speed`, the speed of the robot in m/s, and `timestamp`, the absolute time provided by the GNSS receiver in the HH:mm:ss.SSS format. Information about the geometric "Dilution Of Precision" (DOP), which refers to the accuracy of the data collected by the GPS at the time of use (Farrell 2008), is also provided. In summary, the format of Seekur Jr's GPS log file is the following :

$$\begin{aligned}[\mathtt{latitude, longitude, altitude, speed,}\\ \mathtt{timestamp, number\text{-}of\text{-}satellites,}\\ \mathtt{PDOP, HDOP, VDOP}],\end{aligned}$$

where `PDOP`, `HDOP` and `VDOP` stand for the Position, Horizontal, and Vertical DOP (unitless), respectively. PDOP is a computed measurement of the geometry of satellites above the current location of the receiver: a low PDOP means that the positioning of satellites in the sky is good, and hence good positional accuracy is expected.

4.5   IMU

Two IMUs are installed on the Seekur Jr: one is integrated into the robot (ADIS16362) and the other is external (MTi). Therefore, two different time-indexed text files are provided in Sequence 7: `IMUint.txt` and `IMUext.txt`. The first file contains the following fields,

$$\begin{aligned}[\mathtt{rotX, rotY, rotZ, velX, velY, velZ,}\\ \mathtt{accX, accY, accZ}],\end{aligned}$$

where $[\mathtt{rotX, rotY, rotZ}]$ are the rotation angles of the robot about the x-, y-, and z-axes in radians, $[\mathtt{velX, velY, velZ}]$ are the angular velocities of the robot about the x-, y-, and z-axes in rad/s, and

[accX, accY, accZ] are the linear accelerations of the robot along the x-, y-, and z-axes in m/s^2. The second file, IMUext.txt, only contains the rotation angles of the robot about the x-, y-, and z-axes provided by the MTi IMU, i.e. [rotX, rotY, rotZ]. For Sequence 8, the angular measurements from the IMU on-board the Disco are stored in the text file IMUint.txt, using again the [rotX, rotY, rotZ] format.

5 Conclusion

This paper documented *PanoraMIS*, a new dataset of ultra-wide field of view images captured by catadioptric and twin-fisheye cameras mounted on wheeled, aerial and industrial robots. The data collection covers a large range of indoor and outdoor scenarios with dynamic changes in the scene appearance owing to lighting (direct sun, cast shadows, etc.). It is our hope that PanoraMIS will narrow the gap between theory and practice, and will foster research and encourage reproducible results in the active field of robot vision.

Funding

Supplemental material

The supplementary material accompanying this submission is available for download at the address: mis.u-picardie.fr/~panoramis. The material includes videos of the data collection campaigns, MATLAB scripts for data extraction, processing and visualization, and high-resolution 3-D point clouds (Sequence 7, loops L1 and L2).

Declaration of conflicting interests

The authors declare that they have no conflict of interest.

Notes

1. *Adept MobileRobots* was acquired by *Omron Adept Technology, Inc.* in October 2015.

References

Barreto J (2006) A unifying geometric representation for central projection systems. Comput Vis Image Und 103(3):208–217

Benosman R, Kang SB (2013) Panoramic Vision: Sensors Theory and Applications. Springer

Blanco-Claraco JL, Moreno-Dueñas FA, Gonzalez-Jimenez J (2014) Málaga Urban Dataset: High-rate Stereo and Lidars in a realistic urban scenario. Int J Robot Res 33(2):207–214

Burri M, Nikolic J, Gohl P, Schneider T, Rehder J, Omari S, Achtelik M, Siegwart R (2016) The EuRoC micro aerial vehicle datasets. Int J Robot Res 35(10):1157–1163

Cadena C, Carlone L, Carrillo H, Latif Y, Scaramuzza D, Neira J, Reid I, Leonard J (2016) Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age. IEEE Trans Robot 32(6):1309–1332

Carlevaris-Bianco N, Ushani A, Eustice R (2016) University of Michigan North Campus Long-Term Vision and Lidar Dataset. Int J Robot Res 35(9):1023–1035

Caron G, Eynard D (2011) Multiple Camera Types Simultaneous Stereo Calibration. In: Proc. IEEE Int. Conf. Robot. Automat., pp 2933–2938

Caron G, Morbidi F (2018) Spherical Visual Gyroscope for Autonomous Robots using the Mixture of Photometric Potentials. In: Proc. IEEE Int. Conf. Robot. Automat., pp 820–827

Ceriani S, Fontana G, Giusti A, Marzorati D, Matteucci M, Migliore D, Rizzi D, Sorrenti D, Taddei P (2009) Rawseeds ground truth collection systems for indoor self-localization and mapping. Auton Robot 27(4):353–371

Farrell J (2008) Aided Navigation: GPS with High Rate Sensors. McGraw-Hill

Geiger A, Lenz P, Stiller C, Urtasun R (2013) Vision meets robotics: The KITTI dataset. Int J Robot Res 32(11):1231–1237

Geyer C, Daniilidis K (2001) Catadioptric Projective Geometry. Int J Comput Vision 45(3):223–243

Goodfellow I, Bengio Y, Courville A (2016) Deep Learning. MIT press

Jeong J, Cho Y, Shin YS, Roh H, Kim A (2019) Complex urban dataset with multi-level sensors from highly diverse urban environments. Int J Robot Res 38(6):642–657

Korrapati H, Courbon J, Alizon S, Marmoiton F (2013) The Institut Pascal Data Sets. In: Proc. Journées Francophones des Jeunes Chercheurs en Vision par Ordinateur (ORASIS), [web] http://ipds.univ-bpclermont.fr

Koschorrek P, Piccini T, Oberg P, Felsberg M, Nielsen L, Mester R (2013) A Multi-sensor Traffic Scene Dataset with Omnidirectional Video. In: Proc. IEEE Conf. Comp. Vis. Pattern Recogn. - Workshops, pp 727–734, [web] www.cvl.isy.liu.se/research/datasets

Li M, Mourikis A (2013) High-precision, consistent EKF-based visual-inertial odometry. Int J Robot Res 32(6):690–711

Maddern W, Pascoe G, Linegar C, Newman P (2017) 1 year, 1000 km: The Oxford RobotCar dataset. Int J Robot Res 36(1):3–15

Majdik A, Till C, Scaramuzza D (2017) The Zurich urban micro aerial vehicle dataset. Int J Robot Res 36(3):269–273

Marchand E, Spindler F, Chaumette F (2005) ViSP for Visual Servoing: A Generic Software Platform with a Wide Class of Robot Control Skills. IEEE Rob Autom Mag 12(4):40–52

Matsuki H, von Stumberg L, Usenko V, Stückler J, Cremers D (2018) Omnidirectional DSO: Direct Sparse Odometry with Fisheye Cameras. IEEE Robot Autonom Lett 3(4):3693–3700

Morales Y, Watanabe A, Ferreri F, Even J, Shinozawa K, Hagita N (2018) Passenger discomfort map for autonomous navigation in a robotic wheelchair. Robot Autonom Syst 103:13–26

Morbidi F, Caron G (2017) Phase Correlation for Dense Visual Compass from Omnidirectional Camera-Robot Images. IEEE

Robot Autonom Lett 2(2):688–695

Pandey G, McBride J, Eustice R (2011) Ford Campus vision and lidar data set. Int J Robot Res 30(13):1543–1552

Pasteau F, Narayanan V, Babel M, Chaumette F (2016) A visual servoing approach for autonomous corridor following and doorway passing in a wheelchair. Robot Autonom Syst 75:28–40

Peynot T, Scheding S, Terho S (2010) The Marulan Data Sets: Multi-sensor Perception in a Natural Environment with Challenging Conditions. Int J Robot Res 29(13):1602–1607

Pfrommer B, Sanket N, Daniilidis K, Cleveland J (2017) PennCOSYVIO: A challenging Visual Inertial Odometry benchmark. In: Proc. IEEE Int. Conf. Robot. Automat., pp 3847–3854

Podobnik J, Rejc J, Šlajpah S, Munih M, Mihelj M (2017) All-Terrain Wheelchair: Increasing Personal Mobility with a Powered Wheel-Track Hybrid Wheelchair. IEEE Robot Autonom Lett 24(4):26–36

Rebsamen B, Guan C, Zhang H, Wang C, Teo C, Ang M, Burdet E (2010) A Brain Controlled Wheelchair to Navigate in Familiar Environments. IEEE Trans Neur Sys Reh 18(6):590–598

Schubert D, Goll T, Demmel N, Usenko V, Stückler J, Cremers D (2018) The TUM VI Benchmark for Evaluating Visual-Inertial Odometry. In: Proc. IEEE/RSJ Int. Conf. Intel. Robots Syst., pp 1680–1687

Smith M, Baldwin I, Churchill W, Paul R, Newman P (2009) The New College Vision and Laser Data Set. Int J Robot Res 28(5):595–599

Sturm J, Engelhard N, Endres F, Burgard W, Cremers D (2012) A benchmark for the evaluation of RGB-D SLAM systems. In: Proc. IEEE/RSJ Int. Conf. Intel. Robots Syst., pp 573–580

Sumikura S, Shibuya M, Sakurada K (2019) OpenVSLAM: A Versatile Visual SLAM Framework. In: Proc. 27th ACM Int. Conf. on Multimedia, pp 2292–2295, [web] `http://github.com/xdspacelab/openvslam`

Urban S, Jutzi B (2017) LaFiDa - A Laserscanner Multi-Fisheye Camera Dataset. J Imaging 3(1), article n. 5

Zhang Z, Rebecq H, Forster C, Scaramuzza D (2016) Benefit of Large Field-of-View Cameras for Visual Odometry. In: Proc. IEEE Int. Conf. Robot. Automat., pp 801–808