Specification for the provision of offline data sets from infra3D

Use of mobile mapping data for artificial intelligence (AI)

BSc Daniel Bader

Baden-Dättwil, April 22nd, 2020

Document version: 1.2

Control: Dr. Hannes Eugster

Document status: CONFIDENTIAL

1 Introduction

For the use of mobile mapping data in research and advanced applications, a static database is created based on this offline data record.
The data set includes:
• Recorded single images of the stereo, mono, and panorama cameras
• The associated external and internal orientation parameters
• Disparity maps generated from the stereo images with corresponding metadata
• Point cloud files per stereo image pair
• Geodata files for the axes and segments used

This document specifies the content of the named components and explains them in the following chapters the folder structure, parameter files, and mathematical models in principle.

2. Data structure and file description

The file structure of the data record depends on the data structure used by the infra3D software solution.
In principle, a street («edge») can have several passages(«Segments») and each passage gets its own video sequence for each sensor(«Stream»). In the video sequence, there are several shooting locations in succession(«Frames»), which in ascending order from zero «0» implicitly result in the recording sequence.

2.1. Recording configuration

In Fig. 1, a possible configuration of the processed sensor systems is visualized. The stereo systems each have a disparity map and are in the perspective model. The mono system to the rear probably follows the perspective model but has about no disparity map. The individual images of the panorama camera are equidistant models.

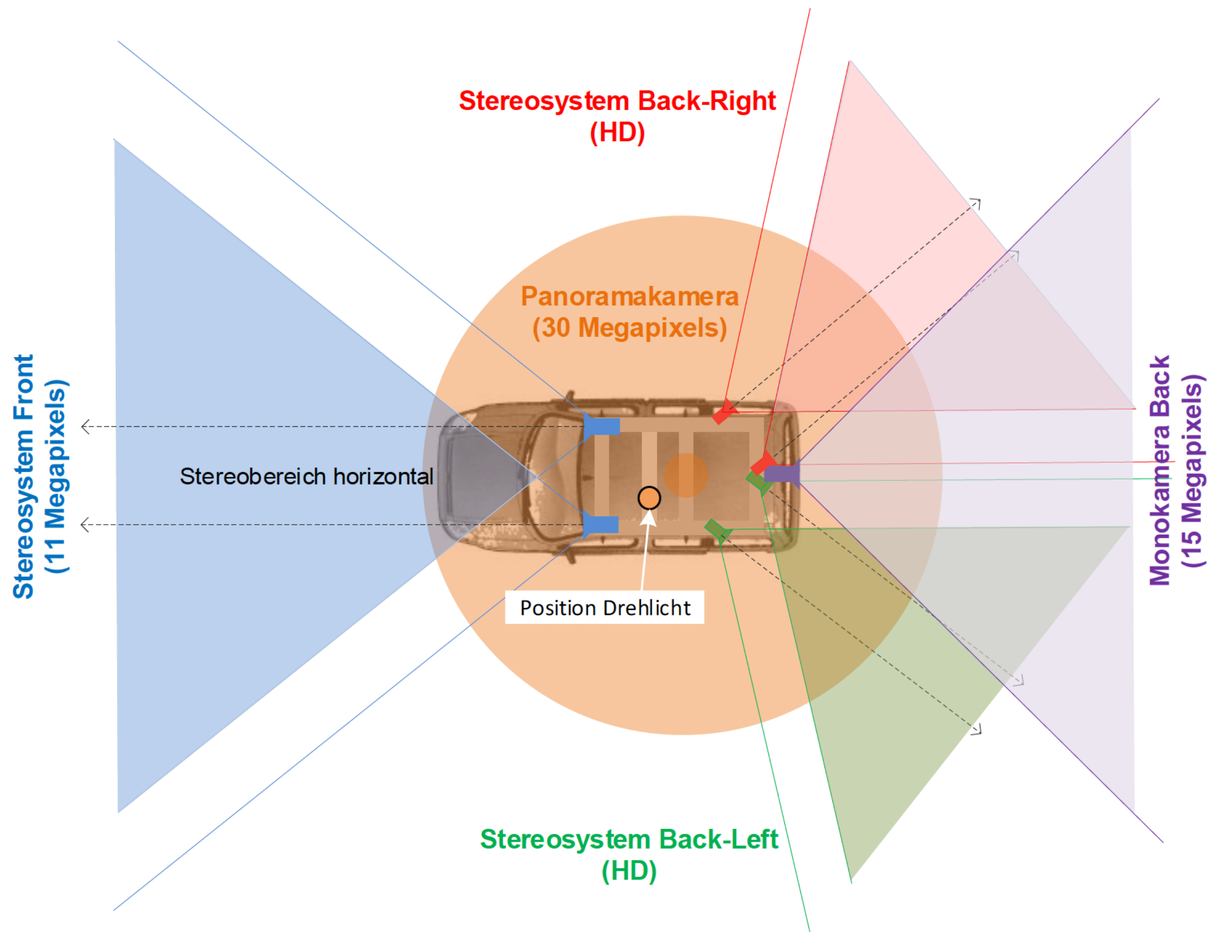

Abb. 1 Beispiel einer Aufnahmekonfiguration

2.2. Frame-based data storage

The image, disparity, and point cloud files follow the standard folder structure. All available streams are listed as folder names under the "Data" directory. Each of these streams has subdirectories that represent the corresponding frame, starting with zero. The standard folder structure down to the single frame is illustrated below:

•… \ data \ <stream number> \ <frame number> \ - frame-based data record--

The frame-based data set now contains all single images, disparity maps, point clouds, and meta information for the given recording location. The sensor number is then marked with the hash «#».

2.2.1. Single images

The individual images are saved as compressed image files (# .jpg). The file name stands for the respective sensor, whereby zero stands for the reference sensor.

2.2.2. Disparity images

The disparity images are stored as image files (# .png). The disparity image is only available with stereo sensor systems and relates to the sensor given in the file name. In the case of this data set, this is always the reference sensor.
The disparity map has three channels (RGB) for each pixel, the red and green channels together managing the disparity value and the blue channel managing the quality.

$$d\,(xd,\,yd) = R\,(xd,\,yd) * 256 + G\,(xd,\,yd)$$

The quality value can have the following values:
• 192: direct pixel assignment in the stereo image pair
• 128: briefly interpolated area
• 64: long interpolated area

2.2.3. Point cloud files

The unfiltered point cloud calculated from the disparity images is available as a LAZ file (# .laz) and can be read with the appropriate libraries. They are marked with the file name of the reference sensor, as they relate to the corresponding sensor.

2.2.4. Meta information

The file with all the meta-information associated with the frame is stored as a text file (# .meta).
All the necessary parameters for further processing can be found in the meta-information. These are specified as follows:

Key	Index	Data Type	Variable	Description
Recording time	1	Date		Recording time of the frame in the format: YYYY-MM-DDThh: mm: ss.xxxxxx
Sensor	1	Integer	#	Position of the sensor in the sensor system Starting at zero (= reference sensor) Corresponds to the filename
	2	Text		Serial number/identifier of the sensor
	3	Float	c	Camera constant in millimeters

Key	Index	Data Type	Variable	Description
Orientation	1	Float	X_0	Sensor position X coordinate in meters
	2	Float	Y_0	Sensor position Y coordinate in meters
	3	Float	Z_0	Sensor position Z coordinate in meters
	4	Float	ω	Sensor orientation Omega in degrees
	5	Float	ϕ	Sensor orientation Phi in degrees
	6	Float	κ	Sensor orientation Kappa in degrees
Image	1	Float	p_i	Pixel size in micrometers
	2	Integer	w_i	Image width
	3	Integer	h_i	Image height
DisMap	1	Float	p_d	Pixel size in micrometers
	2	Integer	w_d	Disparity maps width
	3	Integer	h_d	Disparity maps height
	4	Integer	n	Numerical disparity factor
	5	Float	b	Stereo base

Example from a meta-information file:
Sensor: 0 6064 22,000 → Explanation [Key = Sensor: Index0 = 0 Index1 = 6064 Index2 = 22,000]

The value under “DispMap” can contain “None” if there is no disparity map for the corresponding sensor. It is also possible that no point cloud could be extracted from the given disparity map, then the corresponding LAZ file will not exist. This can occur with images from the right stereo sensor, the mono, or panorama system.

2.3. Cross-frame data storage

In addition to the original data stored in the stand folder structure, derived products that are useful for the overview are stored directly under the main directory.

2.3.1. Axis data set

The axes used for the export are stored as a shape data record and have a spatial reference. In the "Shape" subfolder there is a data record with the name "edges. *" And has the following attributes.

Key	Data Type	Description
idedge	Integer	Ident the axis in the topologically correct edge model segments
segments	Text	List of the segments that are assigned to this axis

2.3.2. Segment record

The passages used for the export are stored as a shape data record and have a spatial reference. In the "Shape" subfolder there is a data record with the name "segments. *" And has the following attributes:

Key	Data Type	Description
idsegment	Integer	Ident the passage in the database
direction	Boolean	Segment runs in or against the axial direction
streams	Text	Associated streams of this passage
framenumber	Integer	Highest frame number for this passage

3. Mathematical basics

3.1. Coordinate system and mapping models

The data set basically comprises two imaging models, the perspective and the equidistant. These differ mainly due to the depth information and the mapping in the model coordinate system.

3.1.1. General transformations and conversions

Rotary dies
The transition between the coordinate systems is based on the following total rotation around co-rotated axes, the angle unit in radians:

$$\boldsymbol{R} = \begin{bmatrix} \cos\varphi\cos\kappa & -\cos\varphi\sin\kappa & \sin\varphi \\ \cos\omega\sin\kappa + \sin\omega\sin\varphi\cos\kappa & \cos\omega\cos\kappa - \sin\omega\sin\varphi\sin\kappa & -\sin\omega\cos\varphi \\ \sin\omega\sin\kappa - \cos\omega\sin\varphi\cos\kappa & \sin\omega\cos\kappa + \cos\omega\sin\varphi\sin\kappa & \cos\omega\cos\varphi \end{bmatrix}$$

Conversion of image (x_i) into sensor coordinates (x_s)

$$x_s = \left(x_i - \frac{w_i}{2}\right) * p_{i[mm]} \qquad y_s = \left(\frac{h_i}{2} - y_i\right) * p_{i[mm]}$$

Conversion of model (x_m) to world coordinates (X_n)

$$\boldsymbol{X}_n = \boldsymbol{X_0} + \boldsymbol{R} * \boldsymbol{x_m} = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + \boldsymbol{R} * \boldsymbol{x_m}$$

3.1.2. Perspective imaging model

In the case of stereo images, model coordinates can be calculated in two ways. Either the disparity value is to be determined using the disparity map or the disparity is calculated directly from observations in both images of the stereo pair. This mapping model is used for mono systems, but a disparity value cannot be calculated.

Conversion of image (x_i) into disparity coordinates (x_d)

$$\boldsymbol{x_d} = \begin{bmatrix} x_d \\ y_d \end{bmatrix} = \boldsymbol{x_i} * \frac{w_d}{w_i} = \begin{bmatrix} x_i \\ y_i \end{bmatrix} * \frac{w_d}{w_i}$$

Conversion of a disparity value into a depth value

$$d_{[\boldsymbol{m}]} = \frac{b_{[m]}}{p_{x0} * p_{[mm]}} = \frac{65535 * b_{[m]}}{d_{(x_d, y_d)} * n * p_{d[mm]}} = \frac{b_{[m]}}{(p_{x1} - p_{x2}) * p_{i[mm]}}$$

Creation of the sensor (x_s) and model coordinates (x_m) for mono and stereo systems

$$\boldsymbol{x_s} = \begin{bmatrix} x_s \\ y_s \\ -c \end{bmatrix} = \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} \qquad \boldsymbol{x_m} = d_{[\boldsymbol{m}]} * \boldsymbol{x_s} = \begin{bmatrix} x_m \\ y_m \\ z_m \end{bmatrix}$$

With mono systems, there is no depth value, therefore: $\boldsymbol{x_m} = \boldsymbol{x_s}$

3.1.3. Equidistant model

Creation of the model coordinates for the panorama system

$$r = \sqrt{{x_s}^2 + {y_s}^2} \qquad \theta = \frac{r}{-c} \qquad k = \frac{\sin\theta}{r} \qquad \boldsymbol{x_s} = \boldsymbol{x_m} = \begin{bmatrix} x_s * k \\ y_s * k \\ cos\,\theta \end{bmatrix} = \begin{bmatrix} x_s \\ y_s \\ z_s \end{bmatrix} = \begin{bmatrix} x_m \\ y_m \\ z_m \end{bmatrix}$$