

Spezifikation zu Bereitstellung von Offlinedatensätzen aus infra3D

Verwendung von Mobile Mapping Daten für künstliche Intelligenz (KI)

BSc Daniel Bader

Baden-Dättwil, 22.04.2020

Dokumentenversion: 1.2
Kontrolle: Dr. Hannes Eugster

Dokumentenstatus: **VERTRAULICH**

Die Rechte an den in dieser Dokumentation enthaltenen Informationen gehören ausschliesslich der iNovitas AG. Ohne Einwilligung der iNovitas AG dürfen diese nicht publiziert oder anderweitig als für Forschungszwecke genutzt werden.

1.	Einleitung	3
2.	Datenstruktur und Filebeschrieb	3
2.1.	Aufnahmekonfiguration	3
2.2.	Framebasierte Datenablage	4
2.2.1.	Einzelbilder	4
2.2.2.	Disparitätsbilder	4
2.2.3.	Punktwolkenfiles	4
2.2.4.	Metainformationen	4
2.3.	Frameübergreifende Datenablage	5
2.3.1.	Achsdatensatz	5
2.3.2.	Segmentdatensatz	5
3.	Mathematische Grundlagen	6
3.1.	Koordinatensystem und Abbildungsmodelle	6
3.1.1.	Allgemeine Transformationen und Umrechnungen	6
3.1.2.	Perspektivisches Abbildungsmodell	6
3.1.3.	Äquidistantes Modell	6

1. Einleitung

Für die Verwendung von Mobile-Mapping-Daten in Forschungs- und weiterführenden Anwendungen, wird auf Basis dieses Offline-Datensatzes eine statische Datenbasis zur Verfügung gestellt.

Der Datensatz umfasst:

- Aufgenommene Einzelbilder der Stereo-, Mono- und Panoramakameras
- Die dazu gehörenden äusseren und inneren Orientierungsparametern
- Aus den Stereobildern generierte Disparitätskarten mit entsprechenden Metadaten
- Punktwolkenfiles pro Stereobildpaar
- Geodatenfiles zu den verwendeten Achsen und Segmenten

Dieses Dokument spezifiziert den Inhalt der genannten Komponenten und erklärt in den kommenden Kapiteln die Ordnerstruktur, Parameterfiles und mathematischen Modelle im Grundsatz.

2. Datenstruktur und Filebeschreibung

Die Filestruktur des Datensatzes richtet sich nach der verwendeten Datenstruktur der infra3D-Softwarelösung. Im Grundsatz kann eine Strasse («Edge») über mehrere Durchfahrten («Segmente») verfügen und jede Durchfahrt bekommt pro Sensor eine eigene Videosequenz («Stream»). In der Videosequenz gibt es aufeinanderfolgend mehrere Aufnahmestandorte («Frames»), welche aufsteigend ab null «0» implizit die Aufnahmereihenfolge ergeben.

2.1. Aufnahmekonfiguration

In Abb. 1 wird eine mögliche Konfiguration der aufbereiteten Sensorsysteme visualisiert. Die Stereosysteme verfügen jeweils über eine Disparitätskarte und liegen im Perspektivischen Modell vor. Das Monosystem nach hinten folgt wohl dem perspektivischen Modell, verfügt aber über keine Disparitätskarte. Die Einzelbilder der Panoramakamera liegen im äquidistanten Modell vor.

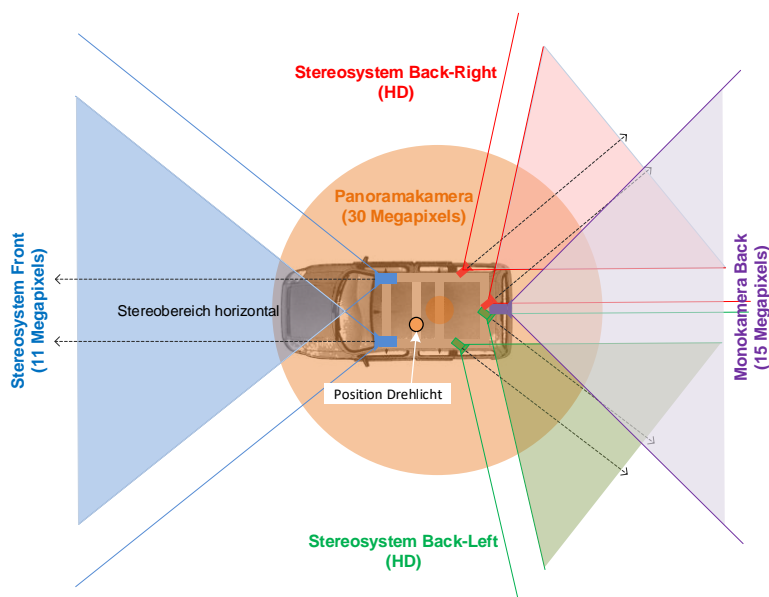


Abb. 1 Beispiel einer Aufnahmekonfiguration

2.2. Framebasierte Datenablage

Die Bild-, Disparitäten- und Punktwolkendateien folgen der Standardordnerstruktur. Unter dem Verzeichnis «Daten» werden alle verfügbaren Streams als Ordnername geführt. Jeder dieser Streams verfügt über Subverzeichnisse, die für das entsprechende Frame stehen, beginnend bei null. Folgend wird die Standardordnerstruktur bis zum Einzelframe illustriert:

- ...\\Daten\\<Stream Nummer>\\<Frame Nummer>\\--framebasierter Datensatz--

Der Framebasierte Datensatz beinhaltet nun alle Einzelbilder, Disparitätskarten, Punktwolken und Metainformationen für den gegebenen Aufnahmestandort. Die Sensornummer wird folgend mit der Raute «#» gekennzeichnet.

2.2.1. Einzelbilder

Die Einzelbilder sind als komprimierte Bilddateien (#.jpg) abgespeichert. Der Dateinamen steht für den jeweiligen Sensor, wobei null für den Referenzsensor steht.

2.2.2. Disparitätsbilder

Die Disparitätsbilder sind als Bilddateien (#.png) abgelegt. Das Disparitätsbild steht nur bei Stereosensorsystemen zur Verfügung und bezieht sich auf den im Dateinamen gegebenen Sensor. Im Falle dieses Datensatzes ist dies immer der Referenzsensor.

Die Disparitätskarte verfügt für jedes Pixel über drei Kanäle (RGB), wobei der Rot- und Grünkanal zusammen den Disparitätswert und der Blaukanal die Qualität verwalten.

$$d_{(x_d, y_d)} = R_{(x_d, y_d)} * 256 + G_{(x_d, y_d)}$$

Der Qualitätswert kann folgende Werte annehmen:

- 192: direkte Pixelzuordnung im Stereobildpaar
- 128: kurz interpolierter Bereich
- 64: lang interpolierter Bereich

2.2.3. Punktwolkenfiles

Die aus den Disparitätsbildern errechnete, ungefilterte Punktwolke liegt als LAZ-File (#.laz) vor und kann mit entsprechenden Bibliotheken gelesen werden. Sie sind mit Dateinamen des Referenzsensors gekennzeichnet, da sie sich auf entsprechenden Sensor beziehen.

2.2.4. Metainformationen

Das File mit allen zum Frame zugehörigen Metainformationen ist als Textfile (#.meta) abgelegt. Den Metainformationen sind alle notwendigen Parameter für die Weiterverarbeitung zu entnehmen. Diese sind wie folgt spezifiziert:

Key	Index	Datentyp	Variable ¹	Beschreibung
Aufnahmezeit	1	Date		Aufnahmezeitpunkt des Frames im Format: YYYY-MM-DDThh:mm:ss.xxxxxx
Sensor	1	Integer	#	Position des Sensors im Sensorsystem Beginnend bei null (=Referenzsensor) Entspricht der Filebenennung
	2	Text		Seriennummer/Kennung des Sensors
	3	Float	c	Kamerakonstante in Millimeter

¹ Variablen werden in Kapitel 3 «Mathematische Grundlagen» verwendet

Key	Index	Datentyp	Variable	Beschreibung
Orientation	1	Float	X ₀	Sensorposition X-Koordinate in Meter
	2	Float	Y ₀	Sensorposition Y-Koordinate in Meter
	3	Float	Z ₀	Sensorposition Z-Koordinate in Meter
	4	Float	ω	Sensorausrichtung Omega in Grad
	5	Float	φ	Sensorausrichtung Phi in Grad
	6	Float	κ	Sensorausrichtung Kappa in Grad
Image	1	Float	p _i	Pixelgrösse in Mikrometer
	2	Integer	w _i	Bild Breite
	3	Integer	h _i	Bild Höhe
DispMap	1	Float	p _d	Pixelgrösse in Mikrometer
	2	Integer	w _d	Disparitätskarten Breite
	3	Integer	h _d	Disparitätskarten Höhe
	4	Integer	n	Nummerischer Disparitätsfaktor
	5	Float	b	Stereobasis

Beispiel aus einer Metainformationsdatei:

Sensor: 0 6064 22.000 → Erläuterung [Key=Sensor: Index0=0 Index1=6064 Index2=22.000]

Der Wert unter «DispMap» kann «None» enthalten, falls für den entsprechenden Sensor keine Disparitätskarte existiert. Ebenfalls ist es möglich das aus gegebener Disparitätskarte keine Punktwolke extrahiert werden konnten, dann wird das entsprechende LAZ-File nicht existieren. Dies kann bei Bildern vom rechten Stereosensor, dem Mono- oder Panoramasystem auftreten.

2.3. Frameübergreifende Datenablage

Nebst der in der Standardnerstruktur abgelegten Originaldaten sind abgeleitete und der Übersicht dienliche Produkte direkt unter dem Hauptverzeichnis abgelegt.

2.3.1. Achsdatensatz

Die für den Export verwendeten Achsen sind als Shape-Datensatz abgelegt und verfügen über einen räumlichen Bezug. Im Unterordner «Shape» befindet sich ein Datensatz mit der Benennung «edges.*» und verfügt über folgende Attribute:

Key	Datentyp	Beschreibung
idedge	Integer	Ident der Achse im topologisch korrekten Kantenmodell
segments	Text	Auflistung der Segmente die dieser Achse zugeordnet sind

2.3.2. Segmentdatensatz

Die für den Export verwendeten Durchfahrten sind als Shape-Datensatz abgelegt und verfügen über einen räumlichen Bezug. Im Unterordner «Shape» befindet sich ein Datensatz mit der Benennung «segments.*» und verfügt über folgende Attribute:

Key	Datentyp	Beschreibung
idsegment	Integer	Ident der Durchfahrt in der Datenbasis
direction	Boolean	Segment verläuft in oder gegen die Achsrichtung
streams	Text	Zugeordnete Streams dieser Durchfahrt
framenummer	Integer	Höchste Framenummer zu dieser Durchfahrt

3. Mathematische Grundlagen

3.1. Koordinatensystem und Abbildungsmodelle

Der Datensatz umfasst grundsätzlich zwei Abbildungsmodelle, das perspektivische und das äquidistante. Diese unterscheiden sich vor allem infolge der Tiefeninformation und der Abbildung im Modellkoordinatensystem.

3.1.1. Allgemeine Transformationen und Umrechnungen

Drehmatrizen

Der Übergang zwischen den Koordinatensystemen basiert auf folgender Gesamtdrehung um mitgedrehte Achsen, Winkleinheit ist Radiant:

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

Umrechnung von Bild- (x_i) in Sensorkoordinaten (x_s)

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

Umrechnung Modell- (x_m) zu Weltkoordinaten (X_n)

$$X_n = X_0 + R * x_m = \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + R * x_m$$

3.1.2. Perspektivisches Abbildungsmodell

Im Fall von Stereobildern können Modellkoordinaten auf zwei Varianten berechnet werden. Entweder ist der Disparitätswert über die Disparitätskarte zu ermitteln oder die Disparität errechnet sich direkt über Beobachtungen in beiden Bildern des Stereopaars. Für Monosysteme wird wohl dieses Abbildungsmodell verwendet, ein Disparitätswert kann jedoch nicht errechnet werden.

Umrechnung von Bild- (x_i) in Disparitätskoordinaten (x_d)

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

Umrechnung eines Disparitätswerts in einen Tiefenwert

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

Erstellung der Sensor- (x_s) und Modellkoordinaten (x_m) für Mono- und Stereosysteme

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

Bei Monosystemen gibt es keinen Tiefenwert deshalb ist: $x_m = x_s$

3.1.3. Äquidistantes Modell

Erstellung der Modellkoordinaten für das Panoramasytem

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