

7 Kapitel zu 3D-Panoramabildern

Gerhard P. Herbig

1) Das 3D-Panorama Dilemma

Ein stereoskopisches Bildpaar, also zwei zusammen gehörende Flachbilder, speichern genau eine einzige (statische) Blickrichtung. Hierbei beträgt der erfasste Raumwinkel je nach verwendeter Brennweite typischerweise einen Bereich zwischen 30° und 70° (natürlich sind bei extremen Weitwinkel- oder Teleobjektiven noch größere und kleinere Werte möglich, das ist hier aber irrelevant).

Bei einem *Panoramabild* erwartet man dagegen eine *Rundumsicht*, also einen Raumwinkel von 360° . Dies kann durch eine einzige Aufnahme niemals erreicht werden.

Bei Monoaufnahmen kann man rundum mehrere Ansichten fotografieren und diese zu einem großen Panoramabild zusammenrechnen (stitchen). Damit beim Stitchen die Ränder perfekt zusammenpassen, muss die Kamera im Zentrum des optischen Strahlengangs, dem *Nodalpunkt*, rotiert werden. Alle Panorama-Teilbilder und damit auch das resultierende Panoramabild haben damit ein einziges Perspektivzentrum. Erst dadurch ist es überhaupt erst möglich, die komplette 360° -Rundumsicht in einem Bild zu speichern. Halten wir also fest: In einem Panoramabild gibt es immer nur ein einziges Perspektivzentrum, das für alle Blickrichtungen konstant bleibt.

Bei Stereoaufnahmen benötigt man aber *zwei unterschiedliche* Perspektivzentren, eines für das linke Auge und eines für das rechte Auge. Gleichzeitig müssen die optischen Achsen aber in jeder Blickrichtung parallel sein, sonst ist keine Stereowirkung möglich. Dies sind zwei widersprüchliche Forderungen, die man gleichzeitig niemals erfüllen kann: Wir haben ein **3D-Panorama Dilemma**.

In der Praxis zeigt sich dieses Dilemma beim Versuch, ein Stereopanoramabild zu fotografieren: **Entweder** man rotiert die *linke* Kamera im *linken* Nodalpunkt und die *rechte* Kamera im *rechten* Nodalpunkt und wird feststellen, dass die damit erzeugten zwei Panoramabilder zwar in sich fehlerfrei sind (also keine Stitchingfehler haben), aber als Stereobildpaare unbrauchbar, da sich die gewünschte Stereowirkung nur in einem kleinen Teilbereich des Panoramabildes einstellt. **Oder** man rotiert eine komplette Stereokamera (also eine Kamera mit zwei fest verbauten Objektiven) in ihrem Zentrum und erhält als Ergebnis dann perfekte Stereo(teil)bilder, die sich aber nicht fehlerfrei zu einem Panoramabild stitchen lassen, weil die Drehung eben *nicht* im Nodalpunkt erfolgte.

Dieses Dilemma bezieht sich aber glücklicherweise *nur* auf den Versuch, ein 3D-Panoramabild aus mehreren Rundum-Einzelbildern berechnen zu wollen. Es gibt durchaus noch andere Möglichkeiten, eine 3D-Rundumansicht mit fotografischen Mitteln für den späteren Gebrauch zu konservieren.

2.) Panoramabild mit Tiefenmasken

Zum Zeitpunkt dieser Niederschrift sind die Fortschritte der Berechnung von Tiefenmasken mit Hilfe von künstlicher Intelligenz (KI) soweit fortgeschritten, dass sich die Umwandlung von (normalen) Monobildern und -videos mit den Diensten, wie sie z.B. von LeiaPix und LookingGlass zur Verfügung gestellt werden, allgemeiner Beliebtheit erfreuen. Diese Methoden sind im Prinzip auch auf (Mono-)Panoramabilder anwendbar, so dass sie sicherlich bald auch für 3D-Panoramabilder geeignet sein werden. Dazu muss die KI mit den speziellen Panoramaverzerrungen zurecht kommen und entsprechend trainiert werden. Dieser Punkt ist Gegenstand aktueller Forschung; eine gute Übersicht findet man in [1].

Insgesamt erscheint mir dieser Weg im Aufwand sehr effizient und lohnend, da er unmittelbar an die gut beherrschten und breit verfügbaren Methoden der Mono-Panoramabild-Erzeugung anknüpft und technisch fehlerfreie 3D-Panoramabilder erzeugen kann. Die tatsächlich vorhandene Perspektive wird auf diese Weise allerdings nicht wiedergegeben, denn das Mono-Panoramabild besitzt ja nur ein einziges Perspektivzentrum.

3.) Simulation einer natürlichen Rundumsicht

Wie wäre denn das richtige Vorgehen, wenn man technisch perfekt den natürlichen Seheindruck in jede Richtung des Panoramas wiedergeben möchte? Ganz einfach: In winzigen Winkelschritten müssten Stereoaufnahmen angefertigt werden und je nach (späterer) Blickrichtung müsste dann die entsprechende Stereoaufnahme zur Anzeige gebracht werden. Mit einer VR-Brille und den entsprechenden Richtungssensoren ist so eine Vorgehensweise durchaus praktikabel.

Eine Alternative zu einer Unmenge an Einzelaufnahmen wäre die Aufnahme eines *Rundumvideos*, das viel schneller und unkomplizierter angefertigt werden kann. Ein Video mit 30 fps und 5 Minuten Laufzeit für eine komplette 360°-Drehung ergibt 9000 Frames, also 25 Aufnahmen je Grad – eine mehr als nur ausreichende Winkelauflösung. Auch die Wiedergabe so eines Rundumvideos hat seine Vorteile, denn es benötigt lediglich einen Videoplayer, der abhängig von der Blickrichtung schnell genug an die entsprechende Stelle springt und das zur Blickrichtung gehörige Frame darstellt.

4.) Flat Stitching

Im ersten Kapitel wurde erwähnt, dass sich Bilder, die bei der Rotation einer komplette Stereokamera entstehen, nicht fehlerfrei zu einem Panoramabild stitchen lassen (weil die Drehung eben *nicht* im Nodalpunkt erfolgt). Dieser Fehler ist von zwei Faktoren abhängig, einmal vom Sektorenwinkel (das ist der Winkel, um den man die Kamera zwischen zwei Aufnahmen weiterdreht) und zweitens von der Nahpunktweite, also der Entfernung von der Kamera bis zum nächstgelegenen Motivpunkt. Für die Nahpunktweite kann man realistische Annahmen treffen (70 Bogenminuten für den Parallaxwinkel, siehe [2]), denn ein *3D-Panorama* impliziert ja gerade das Vorhandensein von Objekten im Nahbereich. So bleibt uns

noch der Sektorenwinkel übrig, um die Stitchingfehler möglichst klein zu halten: Je kleiner dieser Winkel ist, umso geringer fallen auch die Fehler aus. Mit dieser Frage hat sich die SIG-Panorama der DGS beschäftigt. Dazu gibt es ein Arbeitspapier [3], in dem man nachlesen kann, dass dieser Sektorenwinkel nicht größer als etwa 5.7° sein darf, wenn die dadurch produzierten lokalen Höhenfehler maximal ein Promille der Bildhöhe betragen sollen (also etwa 1 Pixel bei FullHD). Das ergibt ein Minimum von 64 Aufnahmen für den Vollkreis. Abgesehen von der Tatsache, dass dies einen nicht unerheblichen Aufwand darstellt, führt dieses Vorgehen aber immer noch zu einem Restfehler, ist also vom Prinzip her nicht fehlerfrei und daher unbefriedigend. Der anzustrebende Sektorwinkel sollte noch sehr viel kleiner sein!

Erinnern wir uns an die Möglichkeit, anstatt vieler Einzelaufnahmen ein *Rundumvideo* anzufertigen, um winzige Sektorenwinkel zu erreichen und an die von den Smartphones bekannte Panoramafunktion, aus einem Schwenk ein Panoramabild zu erstellen. Diese Methode ist unter dem Stichwort *Flat Stitching* bekannt und ist auch für 3D sehr attraktiv: Mit zwei Rundumvideos (eines von der linken Kamera, eines von der rechten Kamera) lassen sich zwei stereoskopisch perfekte Panoramabilder berechnen, die in jeder Richtung einen sauberen Raumeindruck ermöglicht. Da der Sektorwinkel bei dieser Methode extrem klein ist (im obigen Beispiel nur 0.04° da 25 Frames je Grad), ist das Verfahren nahezu fehlerfrei.

Beim Flat Stitching werden nur die *mittleren* Pixelspalten verwendet und aneinandergereiht, daher auch der Name für diese Methode (wörtlich übersetzt: flache Naht). Da sich die Perspektive beim Drehen der Kamera und damit während des Videos ändert, ist auch die Perspektive innerhalb des Panoramabildes nicht mehr konstant. Dies äußert sich in deutlich sichtbaren geometrischen Verzeichnungen im Bild, alle nicht senkrechten Geraden werden zu Kurven. Die auf Panoramabilder spezialisierten Viewer könnten diese Verzeichnungen zwar wieder korrigieren (darauf sind sie spezialisiert), jedoch ist dies bei 3D-Panorama gerade unerwünscht, da diese Entzerrung lokale Höhenfehler erzeugt. Den Preis für ein technisch fehlerfreies Stereopanorama bezahlen wir also mit der Inkaufnahme dieser Verzeichnungen.

COSIMA unterstützt die Flat Stitching Methode, also konkret die Berechnung eines 3D-Panoramabildes aus einem stereoskopischen Rundum-Videopaar. Bevor im nächsten Kapitel die Vorgehensweise genauer erklärt wird, seien hier nochmals die verschiedenen Methoden mit ihren wesentlichen Eigenschaften gegenüber gestellt:

	Lokale Höhenfehler	Perspektive	Kommentar
Monopanorama plus Tiefenmaske	Systembedingt ohne lokale Höhenfehler	Nur eine Perspektive im ganzen Bild, verzeichnungsfrei	Gegenstand aktueller Forschung, z.B. [1]
Simulation der natürlichen Rundumsicht mit sehr vielen Bildern oder Frames	Fehlerfrei nach geometrischer Bildkorrektur	Beste Lösung bezüglich der Wahrnehmung, in jeder Blickrichtung natürliche Perspektive	Spezieller Viewer erforderlich, Datenbasis mehrere 1000 Bilder bzw. Frames notwendig
Linke und rechte Kamera, Drehung jeweils im Nodalpunkt	Stereoskopisch nicht sinnvoll!		
Einzelbilder, Drehung der kompletten Stereokamera, Sektorwinkel > 5.7°	Nicht mehr akzeptable lokale Höhenfehler	Stufenweise veränderliche Perspektive, deshalb Verzeichnungen	Misslungener Versuch, die Methoden der Monopanorama-Fotografie auf 3D zu übertragen
Einzelbilder, Drehung der kompletten Stereokamera, Sektorwinkel < 5.7°	Nicht höhenfehlerfrei, aber noch akzeptabel	Stufenweise veränderliche bis kontinuierliche Perspektive, deshalb Verzeichnungen	Sehr aufwändig, mindestens 64 einzelne Bilder zu stitchen für 360°
Flat Stitching mit Rundumvideo	Fehlerfrei nach geometrischer Bildkorrektur	Kontinuierlich veränderliche Perspektive, deshalb Verzeichnungen	Vertretbarer Aufwand, wird von COSIMA unterstützt

Tabelle: Vergleich verschiedener 3D-Panorama Methoden

5.) Flat Stitching mit COSIMA

Flat Stitching ist eine einfache und aufwandsgünstige Methode, um technisch perfekte, also höhenfehlerfreie 3D-Panoramabilder zu erzeugen. Deshalb wird sie von COSIMA unterstützt. Dazu werden zwei Rundum-Videos mit möglichst hoher Auflösung benötigt. Geeignet sind z.B. zwei Kameras, die 4k für Video anbieten und in Hochformat betrieben werden, z.B. zwei Sony RX100. Mit einer Brennweite von 24mm (äquivalent) bekommt man in der Höhe einen Blickwinkel von etwa 75°. Wenn das nicht reicht, kann man die Ergebnisse optimieren mit Kameras mit Wechselbajonett und dezidierten Superweitwinkelobjektiven (z.B. zwei Z-Cam E1 oder E2 mit 18mm Festbrennweite). Da während der Aufnahme sowieso keine Bewegungen stattfinden sollten, ist bei diesen Aufnahmen eine *perfekte* Videosynchronisation auch nicht notwendig.

COSIMA sucht am Anfang und am Ende der Videos nach identischen Bildinhalten, um ein schnitt-freies Endlospanoramabild zu erzeugen. Deshalb sollten die Videos

etwas überlappen, empfohlen wird ein Bereich von etwa 370°, also 10° Überlappung. Die Bewegung muss nicht unbedingt schon am Beginn des Videos starten, COSIMA sucht und erkennt den Beginn und das Ende der Bewegung. Typische Ergebnisse sind dann Panoramabilder mit 4000 Pixel in der Höhe und 20000 Pixel in der Breite (entspricht dann 75° vertikaler Blickwinkel und 360° horizontaler Blickwinkel.)

Die Videos werden im AviSynth Scriptgenerator geladen wie bei der normalen Videobearbeitung auch. Ein Deshaken ist zwingend notwendig, also muss die entsprechende Checkbox angekreuzt und die Deshaking-Analyse gestartet werden. Das Deshaken gleicht ungleichmäßige Bewegungen des Panoramamotors aus (einfachere Panoramadrehteller arbeiten wie eine Eieruhr und bewegen die Kamera eher *ruckartig* weiter), darüber hinaus misst das Deshaken die Unterschiede zwischen den Frames und verhindert damit eine Dehnung oder Stauchung des Ergebnisbildes. Auch *Teil*-Panoramas mit Videos, welche *keinen* Vollkreis durchlaufen, sind damit ohne Dehnung oder Stauchung möglich.

Als Drehgeschwindigkeit hat sich ein Wert von etwa 5 Minuten für eine Umdrehung bewährt. Kürzere Werte bezahlt man (unter Umständen) mit mehr Unschärfe, längere Werte sind unpraktikabel. Eine kurze Überschlagsrechnung zeigt, dass bei 5 Minuten Videodauer und 30 fps Framerate genau 9000 Frames entstehen und damit ein Frame für $20000/9000 \sim 2.2$ Pixel verantwortlich ist. Die Belichtung sollte also in diesem Fall nicht länger als $1/100$ s sein. Bei kürzeren Belichtungszeiten sind dann auch schnellere Drehbewegungen möglich. Unabhängig von den Aufnahmebedingungen sind die endgültigen Bildpixel aber in jedem Fall Mittelwerte der Bildinhalte *mehrerer* Frames, das reduziert auch die Artefakte des Videocodecs.

6.) Die Parameter von COSIMA

Die wesentliche (und wichtigste) Parameterwahl ist **Video1/VideoOutput = 10**, damit wird die Erzeugung der 3D-Panoramabilder aktiviert. Die weiteren Settings sind:

- **Input2/RotMirInputLeft = 1** oder **3**, je nach Montage falls Hochkantvideo
- **Input2/RotMirInputRight = 1** oder **3**, je nach Montage, falls Hochkantvideo
- **Output1/LeftRightOutput = 0**, um unnötige Bildausgaben zu unterdrücken
- **Output1/StereoOutput = 0**, um unnötige Bildausgaben zu unterdrücken
- **Output1/AnaglyphOutput = 0**, um unnötige Bildausgaben zu unterdrücken
- **Output2/CroppOutput = 1**, da eine Beschneidung der Frames unerwünscht
- **Video1/SmoothGeometryCorr = 0** für konstante Werte innerhalb des Bildes
- **Video1/SmoothWindowCorr = 0** für konstante Werte innerhalb des Bildes

Es wurde schon erwähnt, dass es notwendig ist, die Videos zu deshaken, allerdings darf die Bewegung in Drehrichtung natürlich *nicht* ausgeglichen werden sondern nur die Analysewerte ausgewertet. Die Empfehlung ist daher:

Video2/DeshakeMethod = 1, keine Dynamik

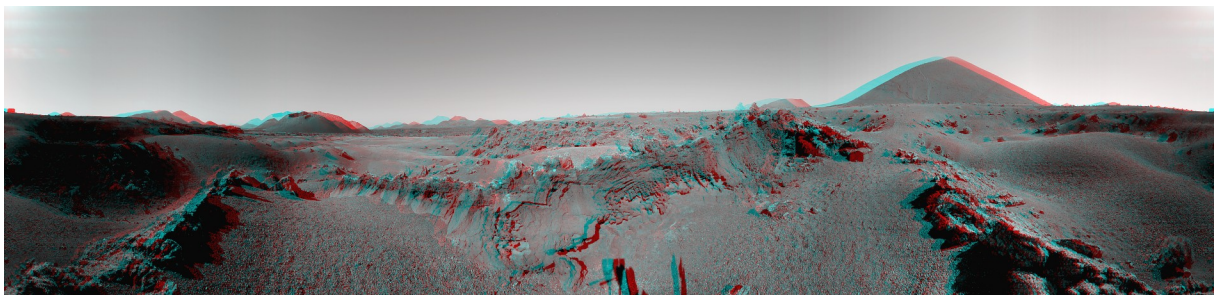
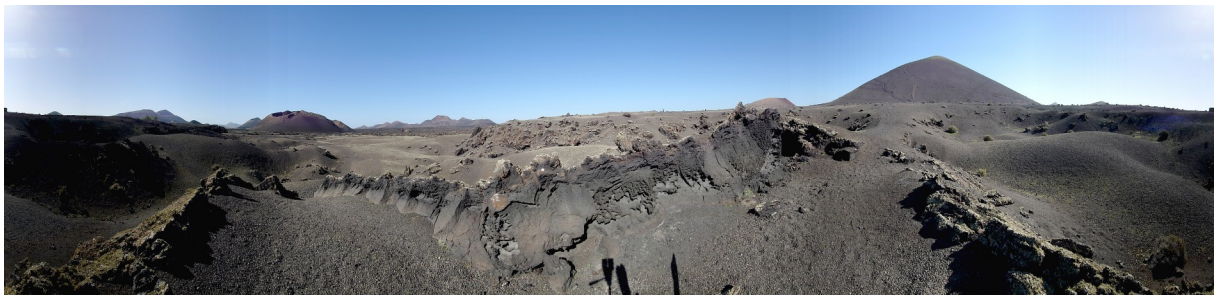
Video2/DeshakeZoomMax = 0, keine Skalierung

Video2/DeshakeFilterX = 1, leichte Korrektur, falls Vibrationen vorliegen

Video2/DeshakeFilterY = 0, keine Korrektur in Drehrichtung (da Hochkant)

Video2/DeshakeFilterR = 5, für Drehungen Defaultwert

Mit diese Parameterwerten wird (wie immer) der Startknopf gedrückt und nach einer gewissen Wartezeit (die Berechnungsdauer ist in etwa genauso groß wie die normale Bearbeitungszeit von Videos) findet man im Unterordner **\cp** vier Bitmaps; das 3D-Panorama als Anaglyphe (ca), im Side-by-Side Layout (cs) und zusätzlich links-rechts-getrennt (cl und cr). Die Ergebnisse sehen dann etwa so aus:

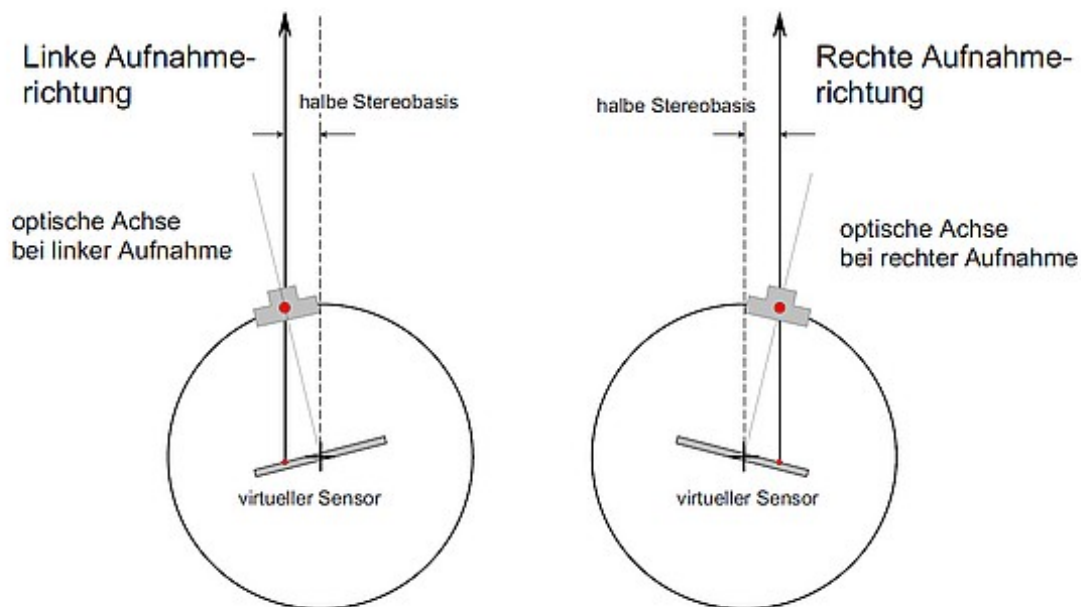


Vulkane auf Lanzarote: Aufgenommen mit zwei TZ10 (hochkant). Oben: linke Ansicht, unten: Anaglyphenbild. Die Störungen unten in der Mitte sind Schatten des Stativs, die bei etwas mehr Sorgfalt vermeidbar gewesen wären.

7.) Sequentielles 3D-Panorama mit *einer* Videokamera

Die Israelis Peleg et. al. haben ein Verfahren vorgestellt [4], wie (ähnlich wie bei der sequentiellen Stereofotografie mit Verschiebetechnik) mit nur *einer Kamera* 3D-Panoramabilder hergestellt werden können. Aufbauend auf der Flat Stitching Technik werden hierbei nicht die mittleren Pixelspalten ausgewertet, sondern – etwas versetzt – die seitlichen. Aus den Pixelspalten der einen Seite wird das linke Panoramabild berechnet, aus der anderen Seite das rechte. Wie bei der *normalen* sequentiellen Stereofotografie darf es auch hier während der Aufnahme keine Bewegungen geben.

In der folgenden Skizze ist schematisch dargestellt, wie sich die Kamera auf einer Kreisbahn bewegt. Die jeweils verwendeten Pixelbereiche sind auf dem Sensor mit roten Punkten markiert:



Sequentielle 3D-Panoramafotografie: Die Kamera bewegt sich auf einer Kreisbahn. Die Skizzen zeigen die beiden unterschiedlichen Zeitpunkte, an denen für einen Motivpunkt die korrespondierenden linken und rechten Pixel belichtet werden.

Mit dem Sonderfall ‚Radius gegen Unendlich‘ schließt diese Methode auch die seitliche Verschiebung der Kamera auf einer Schiene ein. Auch ein Drohnenflug mit senkrecht nach unten gerichteter Kamera ist damit abgedeckt.

Um den Fall dieser sequentiellen 3D-Panoramafotografie mit COSIMA zu bearbeiten, sind *fünf* zusätzliche Parameter notwendig. Diese zusätzlichen Parameter sind aber nicht über die COSIMA Gui zugänglich, sondern müssen in einer zusätzlichen Datei mit dem Namen **parameter.ini** im Verzeichnis **PathofImages** eingetragen werden. Diese fünf Parameter sind:

OneKamera = 1, wird im Fall mit nur einer Kamera auf 1 gesetzt. Im Fall mit zwei Kameras (OneKamera = 0) kann man die Datei parameter.ini auch ganz weglassen.

FocalSensorRatio = 1.0, Verhältnis der Brennweite zur *horizontalen* Sensorbreite, also z.B. 24mm zu 24mm = 1.0. Entweder beide Größen in absoluten Werten oder umgerechnet auf Kleinbild (aber nicht mischen).

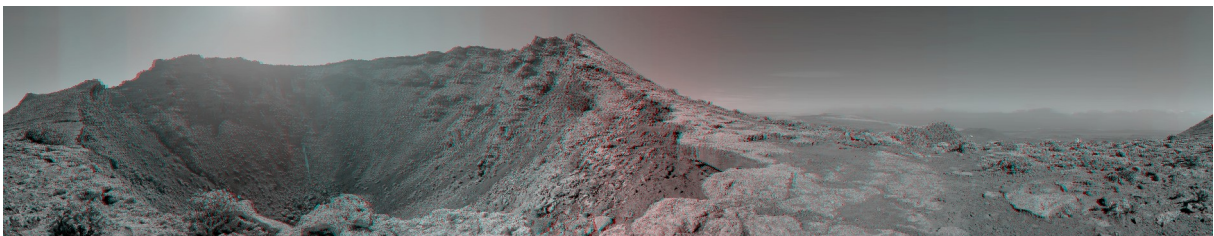
BaseRadiusRatio = 0.75. Hier müssen zwei Fälle unterschieden werden: falls die Kamera im Kreis bewegt wird, ist dies das Verhältnis Stereobasis zu Radius, also z.B. 60mm zu 80mm = 0.75; falls die Kamera auf einer Geraden bewegt wird, ist es das Verhältnis Stereobasis zu Objektstand = $1/30 = 0.033$.

PixelHorizontal = 2160, Anzahl der *horizontalen* Sensorpixel also z.B. 2160 bei 4k

PanoDissolve = 1: falls 360°-Panorama, werden damit die Ränder erkannt und nahtlos aneinandergefügt. Falls das Panorama nur ein Kreissegment abdeckt oder die Kamera nur seitlich bewegt wird (Radius Unendlich oder Drohnenflug), sollte dieser Parameter natürlich zu 0 gesetzt werden.

Die Datei panorama.ini könnte also folgendermaßen aussehen:

```
% Kommentarzeilen mit %, keine Leerzeilen!  
% Diese Inidatei wird nur fuer den Fall "einer Kamera", also fuer OneKamera = 1 beneoetigt  
% Im Fall OneKamera = 0 könnte man diese Datei auch weglassen  
% FocalSensorRatio: Ratio Brennweite zu horizontaler Sensorgroesse, z.B. = 24mm/24mm = 1.0;  
% BaseRadiusRatio: Verhaeltnis Basis zu Radius:  
% Man kann zwei Faelle unterscheiden:  
% - Fall 1: Eine Kamera im Kreis: typisch BaseRadiusRatio = Basis zu Radius = 60mm/80mm = 0.75  
% - Fall 2: Parallele Verschiebung: BaseRadiusRatio = Basis zu Abstand = 1/30 = 0.033  
% PixelHorizontal: Anzahl der Sensor-Pixel, bei 4k hochkant z.B. 2160  
% Genau 5 Parameter, Reihenfolge nicht aendern!  
%  
OneKamera      = 1  
FocalSensorRatio = 1.0  
BaseRadiusRatio = 0.75  
PixelHorizontal = 2160  
PanoDissolve   = 1
```



Vulkane auf Lanzarote: Aufgenommen mit *einer* Z-CAM E1. Oben: linke Ansicht, unten: Anaglyphenbild. Die sichtbaren Raumverzerrungen sind das Ergebnis leider vorhandener Objektivfehler.

Ein Vorteil der sequentiellen 3D-Panoramafotografie ist die *nachträglich wählbare Stereobasis*, in der Parameterdatei panorama.ini mit dem Parameter BaseRadiusRatio als Verhältnis der Stereobasis zum (mechanisch gegebenem) Drehradius.

Eine letzte Bemerkungen, welche bei der sequentiellen 3D-Panoramafotografie noch beachtet werden sollte: Da die Bildinformation unter Umständen weit außerhalb der Mittellinie verwendet wird, ist das Verfahren sehr empfindlich gegenüber Linsenverzeichnungen, die sich in deutlich sichtbaren Raumverzerrungen zeigen. Im Beispielbild oben sind diese Verzerrungen des Raums trotz sehr moderat gewählter Stereobasis (leider) deutlich sichtbar.

[1]: Lu He et al: Rethinking Supervised Depth Estimation for 360° Panoramic Imagery
https://openaccess.thecvf.com/content/CVPR2022W/OmniCV/papers/He_Rethinking_Supervised_Depth_Estimation_for_360deg_Panoramic_Imagery_CVPRW_2022_paper.pdf

[2]: 70-Minuten Bedingung und ihre Folgen,
<http://www.herbig-3d.de/download/70-minuten.pdf>

[3]: Stereoskopischer Stitchingfehler
http://www.cosima-3d.de/download/tn_stereoskopischer_stitchingfehler.pdf

[4]: Panoramic Imaging with Horizontal Stereo,
<https://www.cs.huji.ac.il/~peleg/papers/book00-OmniStereo.pdf>