BONNER METEOROLOGISCHE ABHANDLUNGEN

Heft 88 (2020) (ISSN 0006-7156)
Herausgeber: Andreas Hense

Christoph Beekmans

3-D Cloud Morphology and Evolution Derived from Hemispheric Stereo Cameras

BONNER METEOROLOGISCHE ABHANDLUNGEN

Heft 88 (2020) (ISSN 0006-7156)
Herausgeber: Andreas Hense

Christoph Beekmans

3-D Cloud Morphology and Evolution Derived from Hemispheric Stereo Cameras

3-D Cloud Morphology and Evolution Derived from Hemispheric Stereo Cameras

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Dipl. Geoinformatiker
Christoph Beekmans
aus
Stadtlohn

Bonn, Mai, 2019

Diese Arbeit ist die ungekürzte Fassung einer der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn im Jahr 2019 vorgelegten Dissertation von Christoph Beekmans aus Stadtlohn.

This paper is the unabridged version of a dissertation thesis submitted by Christoph Beekmans born in Stadtlohn to the Faculty of Mathematical and Natural Sciences of the Rheinische Friedrich-Wilhelms-Universität Bonn in 2019.

Anschrift des Verfassers: Address of the author:

Christoph Beekmans
Institut für Geowissenschaften, Abt. Meteorologie
Universität Bonn
Auf dem Hügel 20
D-53121 Bonn

1. Gutachter: Prof. Dr. Clemens Simmer,
Rheinische Friedrich-Wilhelms-Universität Bonn
2. Gutachter: Prof. Dr. Roel Neggers,
Universität zu Köln

Tag der Promotion: 11. Oktober 2019

Danksagung

Als Erstes möchte ich mich bei meinem Doktorvater Prof. Dr. Clemens Simmer bedanken, der mir die Möglichkeit gegeben hat, an einem faszinierenden Thema in der Meteorologie zu arbeiten und zu promovieren. Er hat mich während dieser Zeit immer unterstützt und wichtige Impulse gegeben. Auch möchte ich der Promotionskommission für die Begutachtung der Dissertation danken.

Diese Arbeit enstand im Rahmen der ersten Phase des BMBF-Projektes *High Definition Clouds and Precipitation for Advancing Climate Prediction*, kurz HD(CP)^2, für dessen Finanzierung ich mich bedanken möchte.

Mein besonderer Dank gilt Martin Lennefer, dessen technische Expertise und Erfahrung die Feldexperimente erst ermöglicht hat. Er war integraler Bestandteil dieses Projektes und hat – angefangen von Standortsuche und Bau der Kameras bis hin zu deren Installation sowie Instandhaltung – entscheidend zu den Ergebnissen dieser Arbeit beigetragen.

Ein weiterer entscheidender Beitrag kam von Thomas Läbe sowie Dr. Johannes Schneider vom Institut für Photogrammetrie der Universität Bonn, die mir insbesondere bei der Kalibrierung der Kameras sehr geholfen haben. Bei Prof. Dr. Wolfgang Förstner und Prof. Dr. Cyrill Stachniss möchte ich mich für die interessanten Diskussionen und den wertvollen Rat bedanken.

Für seine aktive Unterstützung unserer Arbeit im Forschungszentrum Jülich möchte ich Dr. Birger Bohn sehr herzlich danken. Zudem gilt mein Dank den Verantwortlichen von RWE (Tagebau Hambach), sowie der Volkshochschule Jülich und des Gymnasiums Zitadelle Jülich für die Erlaubnis zum Betrieb der Kameras auf deren Gelände.

Großen Nutzen habe ich aus den Diskussionen mit Dr. Pablo Saavedra Garfias und Dr. Victor Venema gezogen, um die Resultate der Arbeit und deren Interpretation besser einzuordnen und kritisch zu hinterfragen. Im Zuge dessen möchte ich mich auch bei Dr. Emilio Orlandi, Dr. Kerstin Ebell und Dr. Jan Schween bedanken, die mir bei Erstellung und Zugang von Vergleichsdaten des Wolkenradars, des Lidar-Ceilometers sowie des Wind-Lidars geholfen haben. Bei Prof. Dr. Roel Neggers bedanke ich mich für die interessanten Diskussionen und wichtigen Impulse, die dieser Arbeit zugutegekommen sind.

Auch möchte ich mich bei Sophie und Florian bedanken, die mir wichtige Kommentare für den Feinschliff dieser Arbeit haben zukommen lassen.

Zu guter Letzt möchte ich mich noch bei meiner Familie bedanken, insbesondere bei meinen Eltern, die mich immer unterstützt haben.

Abstract

Clouds play a key role in the Earth-atmosphere system as they reflect incoming solar radiation back to space, while absorbing and emitting longwave radiation. A significant challenge for observation and modeling pose cumulus clouds due to their relatively small size that can reach several hundreds up to a few thousand meters, their often complex 3-D shapes and highly dynamic life-cycle. Common instruments employed to study clouds include cloud radars, lidar-ceilometers, (microwave-)radiometers, but also satellite and airborne observations (in-situ and remote), all of which lack either sufficient sensitivity or a spatial or temporal resolution for a comprehensive observation. This thesis investigates the feasibility of a ground-based network of hemispheric stereo cameras to retrieve detailed 3-D cloud geometries, which are needed for validation of simulated cloud fields and parametrization in numerical models. Such camera systems, which offer a hemispheric field of view and a temporal resolution in the range of seconds and less, have the potential to fill the remaining gap of cloud observations to a considerable degree and allow to derive critical information about size, morphology, spatial distribution and life-cycle of individual clouds and the local cloud field.

The technical basis for the 3-D cloud morphology retrieval is the stereo reconstruction: a cloud is synchronously recorded by a pair of cameras, which are separated by a few hundred meters, so that mutually visible areas of the cloud can be reconstructed via triangulation. Location and orientation of each camera system was obtained from a satellite-navigation system, detected stars in night sky images and mutually visible cloud features in the images. The image point correspondences required for 3-D triangulation were provided primarily by a dense stereo matching algorithm that allows to reconstruct an object with high degree of spatial completeness, which can improve subsequent analysis.

The experimental setup in the vicinity of the Jülich Observatory for Cloud Evolution (JOYCE) included a pair of hemispheric sky cameras; it was later extended by another pair to reconstruct clouds from different view perspectives and both were separated by several kilometers. A comparison of the cloud base height (CBH) at zenith obtained from the stereo cameras and a lidar-ceilometer showed a typical bias of mostly below 2% of the lidar-derived CBH, but also a few occasions between 3-5%. Typical standard deviations of the differences ranged between 50 m (1.5 % of CBH) for altocumulus clouds and between 7% (123 m) and 10% (165 m) for cumulus and strato-cumulus clouds. A comparison of the estimated 3-D cumulus boundary at near-zenith to the

sensed 2-D reflectivity profiles from a 35-GHz cloud radar revealed typical differences between 35 - 81 m. For clouds at larger distances (> 2 km) both signals can deviate significantly, which can in part be explained by a lower reconstruction accuracy for the low-contrast areas of a cloud base, but also with the insufficient sensitivity of the cloud radar if the cloud condensate is dominated by very small droplets or diluted with environmental air.

For sequences of stereo images, the 3-D cloud reconstructions from the stereo analysis can be combined with the motion and tracking information from an optical flow routine in order to derive 3-D motion and deformation vectors of clouds. This allowed to estimate atmospheric motion in case of cloud layers with an accuracy of 1 ms^{-1} in velocity and 7° to 10° in direction. The fine-grained motion data was also used to detect and quantify cloud motion patterns of individual cumuli, such as deformations under vertical wind-shear.

The potential of the proposed method lies in an extended analysis of life-cycle and morphology of cumulus clouds. This is illustrated in two show cases where developing cumulus clouds were reconstructed from two different view perspectives. In the first case study, a moving cloud was tracked and analyzed, while being subject to vertical wind shear. The highly tilted cloud body was captured and its vertical profile was quantified to obtain measures like vertically resolved diameter or tilting angle. The second case study shows a life-cycle analysis of a developing cumulus, including a time-series of relevant geometric aspects, such as perimeter, vertically projected area, diameter, thickness and further derived statistics like cloud aspect ratio or perimeter scaling. The analysis confirms some aspects of cloud evolution, such as the pulse-like formation of cumulus and indicates that cloud aspect ratio (size vs height) can be described by a power-law functional relationship for an individual life-cycle.

Zusammenfassung

Wolken haben einen maßgeblichen Einfluss auf den Strahlungshaushalt der Erde, da sie solare Strahlung effektiv reflektieren, aber von der Erde emittierte langwellige Strahlung sowohl absorbieren als auch ihrerseits wieder emittieren.

Darüber hinaus stellen Cumulus-Wolken wegen ihrer verhältnismäßig kleinen Ausdehnung von wenigen hundert bis einigen tausend Metern sowie ihres dynamischen Lebenszyklus nach wie vor eine große Herausforderung für Beobachtung und Modellierung dar. Gegenwärtig für deren Erforschung im Einsatz befindliche Instrumente wie Lidar-Ceilometer, Wolkenradar, Mikrowellenradiometer oder auch satellitengestützte Beobachtungen stellen die für eine umfassende Erforschung dieser Wolken erforderliche räumliche und zeitliche Abdeckung nicht zur Verfügung. In dieser Arbeit wird untersucht, inwieweit eine bodengebundene Beobachtung von Wolken mit hemisphärisch projizierenden Wolkenkameras geeignet ist detaillierte 3-D Wolkengeometrien zu rekonstruieren um daraus Informationen über Größe, Morphologie und Lebenszyklus einzelner Wolken und des lokalen Wolkenfeldes abzuleiten.

Grundlage für die Erfassung der 3-D Wolkengeometrien in dieser Arbeit ist die 3-D Stereorekonstruktion, bei der eine Wolke von jeweils zwei im Abstand von mehreren Hundert Metern aufgestellten, synchron aufnehmenden Kameras abgebildet wird. Beidseitig sichtbare Teile einer Wolke können so mittels Triangulation rekonstruiert werden. Fischaugen-Objektive ermöglichen das hemisphärische Sichtfeld der Wolkenkameras. Während die Positionsbestimmung der Kameras mit Hilfe eines Satelliten-Navigationssystems durchgeführt wurde, konnte die absolute Orientierung der Kameras im Raum mit Hilfe von detektierten Sternen bestimmt werden, die als Referenzpunkte dienten. Die für eine Stereoanalyse wichtige relative Orientierung zweier Kameras wurde anschließend unter Zuhilfenahme von Punktkorrespondenzen zwischen den Stereobildern verfeinert.

Für die Stereoanalyse wurde primär ein Bildanalyse-Algorithmus eingesetzt, welcher sich durch eine hohe geometrische Vollständigkeit auszeichnet und auch 3-D Informationen für Bildregionen mit geringem Kontrast liefert. In ausgewählten Fällen wurden die so rekonstruierten Wolkengeometrien zudem mit einem präzisen Mehrbild-Stereo-Verfahren verglichen. Eine möglichst vollständige 3-D Wolkengeometrie ist vorteilhaft für eine darauffolgende Analyse, die eine Segmentierung und Identifizierung einzelner Wolken, deren raum-zeitliche Verfolgung oder die Ableitung geometrischer Größen umfasst.

Der experimentelle Aufbau im Umfeld des Jülich Observatory for Cloud Evolution (JOYCE) umfasste zuerst eine, später zwei Stereokameras, die jeweils

mehrere Kilometer entfernt installiert wurden um unterschiedliche Wolkenpartien rekonstruieren zu können. Ein Vergleich zwischen Stereorekonstruktion und Lidar-Ceilometer zeigte typische Standardabweichungen der Wolkenbasishöhendifferenz von 50 m (1.5 %) bei mittelhoher Altocumulus-Bewölkung und 123 m (7 %) bis 165 m (10 %) bei heterogener Cumulus- und Stratocumulus-Bewölkung. Gleichzeitig wich die rekonstruierte Wolkenbasishöhe im Durchschnitt meist nicht weiter als 2 %, in Einzelfällen 3-5% vom entsprechenden Wert des Lidars ab. Im Vergleich zur abgeleiteten Cumulus-Morphologie aus den 2-D Reflektivitätsprofilen des Wolkenradars, zeigten sich im Zenit-Bereich typische Differenzen zwischen 35 und 81 m. Bei weiter entfernten Wolken (> 2 km) können sich Stereorekonstruktion und Reflektivitätssignal stark unterscheiden, was neben einer abnehmenden geometrischen Genauigkeit der Stereorekonstruktion in kontrastarmen Bereichen insbesondere mit einer oftmals unzureichenden Sensitivität des Radars bei kleinen Wolkentröpfchen erklärt werden kann, wie man sie an der Wolkenbasis und in den Randbereichen von Wolken findet.

Die Kombination von Stereoanalyse und der Bewegungsinformation innerhalb einer Bildsequenz erlaubt die Bestimmung von Wolkenzug- und -deformationsvektoren. Neben der Verfolgung einzelner Wolkenstrukturen und der Erfassung von Wolkendynamik (beispielsweise der Deformation von Wolken durch Windscherung), kann im Fall von stratiformen Wolken Windgeschwindigkeit und -richtung abgeschätzt werden. Ein Vergleich mit Beobachtungen eines Wind-Lidars zeigte hierfür typische Abweichungen der Windgeschwindigkeit von 1 ms^{-1} und der Windrichtung von 7° to 10°.

Ein besonderer Mehrwert der Methode liegt in einer tiefergehenden Analyse von Morphologie und Lebenszyklus von Cumulus-Wolken. Dies wurde anhand zweier exemplarischer Fallstudien gezeigt, in denen die 3-D-Rekonstruktionen zweier entfernt aufgestellter Stereokameras kombiniert wurden. Im ersten Fall wurde ein sich unter vertikaler Windscherung entwickelnder Cumulus von zwei Seiten aufgenommen, was eine geometrische Erfassung des stark durch Scherung geneigten Wolkenkörpers ermöglichte. Kennwerte wie Vertikalprofil, Neigungswinkel der Wolke und Durchmesser einzelner Höhenschichten wurden abgeschätzt. Der zweite Fall zeigte eine statistische Analyse eines sich entwickelnden Cumulus über seinen Lebenszyklus hinweg. Dies erlaubte die Erstellung einer Zeitreihe mit relevanten Kennzahlen wie äquivalenter Durchmesser, vertikale Ausdehnung, Perimeter oder abgeleitete Größen wie Aspektrate oder Perimeter-Skalierung. Während die Analyse bisherige Ergebnisse aus Simulationen und satellitengestützten Beobachtungen bestätigt, erlaubt diese aber eine Erweiterung auf die Ebene individueller Wolken und der Ableitung funktionaler Zusammenhänge wie zum Beispiel dem Verhältnis von Wolkendurchmesser und vertikaler Dimension.

Contents

1 Introduction

1.1 Motivation

The role of clouds in the Earth-atmosphere system is manifold. They regulate the Earth's energy budget by reflecting incoming light from the sun, while absorbing and emitting radiation in the infrared spectrum. The interplay between clouds, radiation and atmospheric circulation introduces a high complexity to climate models (Figure 1.1). Accordingly, a well-founded understanding of cloud processes, spatial distribution and evolution is required to reduce uncertainties in climate models. Among the great variety of clouds, cumulus is a particularly challenging category for both modeling and observation. The underlying convective and turbulent processes take place at much smaller scales than is usually resolved by numerical models. The typical size of fair weather cumulus clouds ranges between several hundred meters and a few kilometers, whereas the grid-cell size of large-scale numerical models reaches several kilometers (numerical weather prediction) up to tens or hundreds of kilometers (global climate model). Parametrizations replace an explicit simulation of convection and radiative transfer. This, however, requires detailed knowledge of macro- and micro-physical properties of cloud populations and life-cycles of individual clouds under various atmospheric conditions. Deficits, however, still persist in the ability to adequately monitor these types of clouds, their geometric characteristics, spatial distribution and evolution due an insufficient sampling or sensitivity of current instrumentation. It is desirable to investigate new approaches to complement existing observations in this regard.

Convection plays an important role by transferring energy between the planetary boundary layer and the free atmosphere (Siebesma et al., 2003; Sherwood et al., 2010). On a global scale, shallow convection promotes tropical deep convection as it significantly moistens the sub-tropic air in the trade-wind-layer before it reaches the inner-tropical convergence zones (Stevens et al., 2001; Tiedtke, 1989).

Cloud size, height and evolution are important indicators of the net vertical transport and distribution of moisture and heat. Cumulus clouds are strongly linked to the thermodynamic properties and evolution of the convective boundary layer (Zhang and Klein, 2013), while their life-cycle is influenced by turbulent and dynamical mixing of cloud air with environmental air (Blyth, 1993). Several formulations have been proposed to describe this *entrainment* process adequately, suggesting a dependence on cloud size, height of entrainment, relative humidity and magnitude of buoyancy (de Rooy et al., 2012).

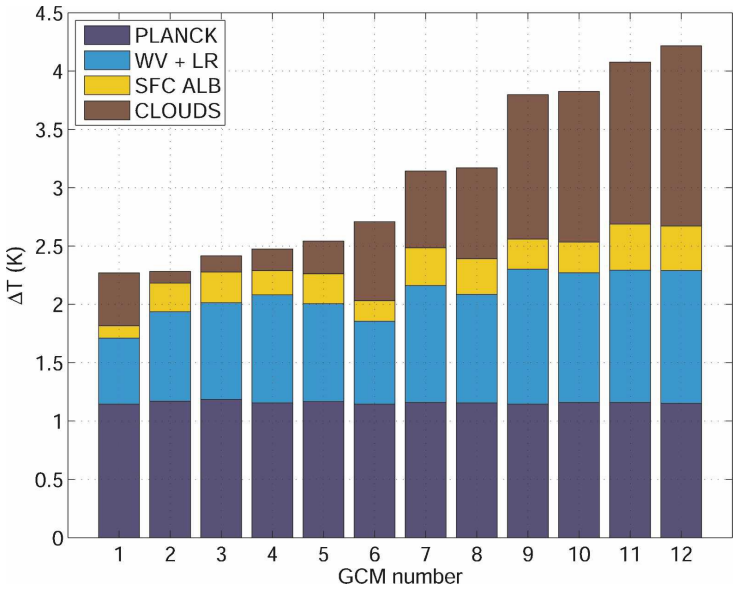


Figure 1.1: Surface temperature change caused by doubling of CO_2 computed by an ensemble of 12 different climate models (GCM number), decomposed into: equilibrium temperature change due to surface heating and emission (PLANCK), absorption and re-emission by water vapor and change of the atmospheric lapse rate (WV+LR), a change of surface albedo (SRF ALB) and cloud-radiation feedback (CLOUD). Adapted from Dufresne and Bony (2008).

Simulations suggested that smaller cumulus clouds tend to moisten the environmental air along their vertical extent, while larger cumuli moisten primarily the cloud top regions and have a net warming effect near the base (Zhao and Austin, 2003).

Radiative transfer depends explicitly on the geometric properties of cumulus populations and individual clouds (Davies, 1978). Radiative transfer routines in general circulation models (GCM), such as the ECHAM-6 or ICON model (Stevens et al., 2013; Giorgetta et al., 2018), generally assume horizontally homogeneous clouds in order to keep computational complexity manageable. Calculations are performed for individual atmospheric columns, while ignoring the horizontal net transport of radiation between the columns (Cahalan et al., 1994; Chambers et al., 1997). Compared to their horizontal size, cumulus clouds often reach high vertical dimensions (hundreds of meters up to several kilometers) and their occurrence in cloud populations complicate radiative transfer in several ways (Figure 1.2). 3-D effects like mutual shadowing, photon leakage through the sides of the cloud and interception of solar radiation due to high vertical extents thus become relevant (Hinkelman et al., 2007; Evans et al., 2001). Satellite remote sensing of cloud micro-physical properties also depends explicitly on the cloud shape (Vant-Hull et al., 2007).

One approach to derive cumulus population statistics and investigate cloud

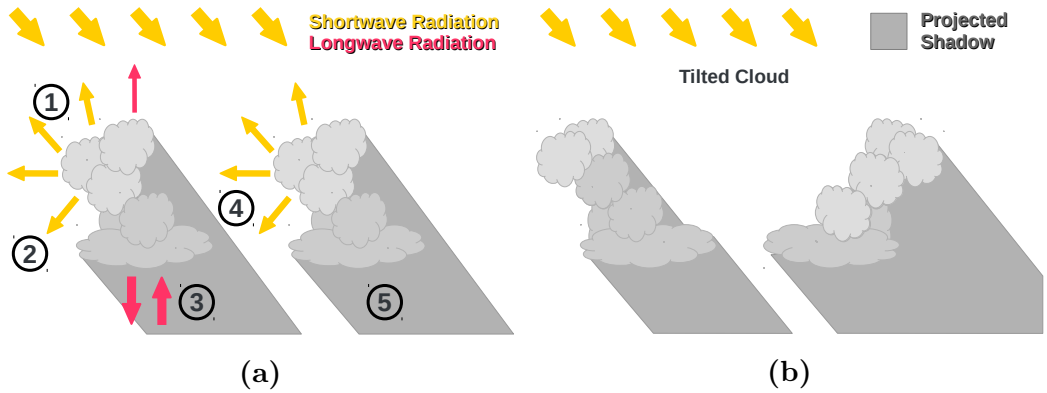


Figure 1.2: Radiative impact of cumulus clouds (a) depends strongly on cloud size, geometry and spacing due to simultaneous reflection (1), horizontal scattering (2), absorption/emission (3), mutual illumination (4) and shadowing (5). Cloud shadowing can vary significantly if the clouds are tilted (b).

life-cycle is based on an extensive analysis of data produced by Large-Eddy-Simulations (LES) (Brown et al., 2002; Siebesma et al., 2003; Neggers et al., 2003; Zhao and Austin, 2005; Heus et al., 2009; Heus and Jonker, 2008; Brast et al., 2016). Thanks to a comparatively small grid size of about 25 m, relevant turbulent and dynamic processes can be resolved. Key information, such as liquid water content, virtual potential temperature or the individual terms of the vertical momentum equation, is directly accessible and allows to identify and track individual clouds over their life-cycle (Griffith et al., 2005).

LES are, however, still models and thus require observational validation, which is in part conducted by ground-based instruments, such as cloud radars, microwave radiometers and lidar-ceilometers (Löhnert et al., 2015; Stevens et al., 2016; Corbetta et al., 2015; Oue et al., 2016). Millimeter-wavelength cloud radars are designed to sense cloud droplets, which generally reach diameters of only a few up to some tens of micrometers. The transparency of clouds at these wavelengths allows to retrieve the internal cloud structure by sensing the reflected radar signal along a pencil beam. The high sensitivity on droplet size (to the sixth power), however, also renders cloud radars often ineffective, especially at the cloud base and the edges, where condensation of sub-cloud air and mixing with environmental air produces relatively small droplets. Moreover, a single cross-section scan takes, depending on the signal integration time, between 15 to 60 seconds, yields only a single slice through the full hemisphere and thus cannot provide a detailed and instantaneous view on the local cloud field. Scanning schemes were developed that exploit cloud motion direction to sense passing clouds slice by slice (Kollias et al., 2014; Lamer and Kollias, 2015; Fielding et al., 2013), but could not overcome the limited sensitivity as a source of uncertainty. Also, current cloud radars still rely on moving mechanical parts, which makes a permanent, intensive scanning

problematic.

Similar to cloud radars, lidar-ceilometers emit radiation along a pencil beam and are used to detect clouds and derive cloud base height. Instead of millimeter waves, the laser beam of a lidar emits light in the ultra-violet (UV), visible (VIS) or near infra-red (NIR), which allows a reliable detection of cloud boundaries due to an efficient backscatter. This comes, however, at the cost of a vastly diminished beam penetration depth as clouds are often opaque at such wavelengths. The lidar-ceilometer is generally aligned to a fixed direction and spatial coverage is limited to a point measurement, which makes them ineffective in monitoring cumulus clouds and reconstruction of the small-scale morphology.

Air- and space-borne cloud observations offer a wider view on geometric aspects of clouds and cloud fields, compared to their ground-based counterparts. In-situ and/or remote sensing of clouds from aircraft provides direct estimates of droplet size distributions, vertical velocity or geometric size, but is expensive and thus limited to field campaigns (Plank, 1969; Damiani et al., 2006; Alexandrov et al., 2016). Satellites allow to monitor cloud fields and populations globally. Geostationary satellites like *Meteosat* or *GOES* always observe the same area, but at the cost of a comparatively low temporal resolution (e.g. 5 to 15 minutes for Meteosat-11/10) and a rather coarse spatial resolution of one or more kilometers due to the large distance to Earth of 36.000 km. Polar-orbiting satellites orbit Earth at a height of several hundred kilometers and offer finer resolutions. Radiometric imagers, such as the ASTER instrument on-board the NASA satellite *Terra*, are able to record images of up to 15 m resolution; the multi-angle instrument MISR of up to 250 m. The high resolution comes at the cost of a much smaller repeat coverage (time to next visit of the same region) of 9 days in case of MISR. Although the photogrammetric method is well applied in satellite remote sensing and provides information about cloud top height and motion (Muller et al., 2007; Seiz and Davies, 2006; Genkova et al., 2007; Mueller et al., 2017), the low temporal coverage of satellites and aircraft make them inappropriate for the monitoring of cumulus life-cycle.

Ground-based cloud photogrammetry has the potential to close the gap of current cloud observations to a large degree. Hemispheric sky cameras capture images of the visible sky in high detail and at arbitrary time intervals. Employed in a stereo setup, they are capable to monitor cloud formation via 3-D stereo reconstruction and tracking instantaneously, which can be used to derive statistics about the cloud field and individual clouds, such as 3-D morphology, life-cycle or spatial organization. Meanwhile, sky imagers are deployed at many measurement sites, but their dominant application is limited to the derivation of cloud cover (Tuominen and Tuononen, 2017) and cloud type classification (Heinle et al., 2010). Given the large uncertainties related to clouds and cumulus clouds in particular, it is desirable to investigate the potential of this cost-efficient method and its incorporation into existing ob-

servation sites. Previous experimental studies, which employed stereo pairs of cameras with wide-angle or fisheye lenses, have already shown that cloud photogrammetry is able to provide key information about individual clouds, such as cloud base height (Seiz et al., 2007a) or cloud top height (Öktem and Romps, 2015).

An attempt to exploit the potential of ground-based 3-D vision to reproduce and quantify the irregular morphology, size and spatial distribution of individual cumulus clouds was, however, not explored so far. In this study, it is assumed that the method is well-suited to validate LES models and help understanding the 3-D radiative effect under cloudy conditions more comprehensively. Both research applications require further geometric characteristics besides cloud base and top height, namely cloud area (diameter), perimeter, shape and vertical overlap, volume and boundary surface area. A successful survey of individual clouds then should allow to estimate cloud population statistics like size distributions and nearest-neighbor-spacing of clouds. A major research goal of this thesis is to derive these geometric quantities using ground-based hemispheric stereo cameras.

1.2 Previous Work on Ground-Based Stereo Reconstruction

Ground-based photogrammetry in atmospheric sciences dates back to Koppe (1896) who was using two theodolites to obtain viewing orientation angles for cloud height and location estimation. An early survey of towering cumulus clouds was conducted by Malkus and Ronne (1954) who estimated the size and growth rate of cumulonimbus clouds. Orville and Kassander (1961) installed a pair of analog K-17 aerial reconnaissance and mapping cameras separated three miles from each other and oriented towards the Santa Catalina Mountains near Tucson, Arizona, with the aim to observe orographic convection. An early application of hemispheric cameras was done by Bradbury and Fujita (1968) who derived cloud base height and motion from stereo image sequences. In all these studies, the orientation of the cameras and the registration of image features had to be done manually, e.g. by using a stereoplanimetric plotter or a tripod with a gimbal on an airplane.

The advent of digital consumer cameras and growing computational capacities greatly accelerated applications of cloud photogrammetry by offering automated high-quality image acquisition and stereo analysis. The growing number of recent publications in this area shows the interest in a reliable and affordable instrumentation that is able to fill gaps of current cloud observations.

More recent studies employed stereo cameras with wide-angle lenses and performed a feature-based stereo analysis, i.e. the matching of distinctive interest points, lines or other geometric entities across the images. Seiz et al. (2002)

applied an automated least-squares matching of interest points to establish correspondences between the stereo images. They used two cameras with a view field of 100° and a spatial distance of 800 meters to derive the cloud base height of passing clouds with errors of well below 5%. The cloud top height was derived from satellite images, which allows for a more complete view on the cloud body if combined. A cloud motion analysis was discussed, but not conducted. Zehnder et al. (2006) continued the work of Orville in the frame of the CuPIDO project (Damiani et al., 2008). Similar to Seiz et al. (2002) they performed a feature-based stereo analysis with automatically detected, but manually matched image points. Hu et al. (2010) provided a detailed description of absolute orientation estimation of the camera system using visible landmarks. They compared the resulting 3-D point cloud with airborne measurements of liquid water content that took place simultaneously, but an empirical validation of their 3-D reconstructions is not available. Öktem et al. (2014) applied a pair of wide-angle cameras to observe marine convection and estimated the updraft of individual cloud thermals (Öktem and Romps, 2015; Romps and Öktem, 2015). Their stereo system near Biscayne Bay, Florida, with a baseline (i.e. the distance between the cameras) of 873 m offered sufficient accuracy to compute reliable height values of convective turrets several kilometers away. Like Seiz et al. (2002) and Hu et al. (2010), they applied a stereo analysis based on image features and validated their reconstructions against a co-located lidar-ceilometer, revealing errors less than 2% for shallow clouds and 8% for high cirro-cumulus. They also derived cloud motion for stratiform clouds, which was validated against radiosonde measurements. Location and alignment of the cameras towards the ocean permitted the use of the visible horizon for camera orientation estimation.

Studies involving hemispheric sky imagers were mainly focused on the derivation of cloud base height. Allmen and Kegelmeyer (1996) used two Whole Sky Imagers (WSI) with a baseline of 5 km and found that an ordinary stereo analysis was difficult due to the large distance between the cameras. They reported that in 50% of the cases the error was below 5% compared to the heights obtained by the nearby lidar-ceilometer. Kassianov et al. (2005) performed a stereo analysis of image data obtained from a virtual camera pair with stochastically generated cloud images, and a real pair of sky imagers situated at the Atmospheric Radiation Measurement (ARM) site in the southern Great Plains. They used a baseline length of 540 m and reported typical errors of about 10% for low-level clouds up to 2 km height. Nguyen and Kleissl (2014) derived the cloud base height using two sky imagers separated by a baseline of 1230 m. 3-D reconstruction was accomplished by determining the height at which a maximum correlation between two projected image patches was achieved. Veikherman et al. (2015) investigated the use of five Raspberry Pi modules with fisheye camera adapters to derive the external and internal structure of shallow cumulus clouds. They used a variant of the space carving technique (Kutulakos and Seitz, 1999), which iteratively carves a 3-D volume

until a photometrically consistent object appearance is reached in all cameras. They argued in favor of a very large camera network of cheap devices in order to exploit the visual information given from multiple views. A comparison with a radiosonde sounding showed good agreement of increased moisture in the same region where clouds had been reconstructed. Crispel and Roberts (2017) used a pair of mobile hemispheric cameras to georeference clouds during a measurement campaign. They also derived the horizontal displacement of the cloud field via image tracking.

In this thesis, up to four hemispheric sky imagers are employed to obtain complementary 3-D information of cumulus clouds from different perspectives. The underlying 3-D retrieval technique is the well-established stereo reconstruction. The images taken by a synchronously recording pair of cameras observing an object are used in a stereo analysis and provide a set of corresponding image points that represent the same 3-D object point. Once the internal projective properties of the cameras are computed in a calibration procedure, the position and orientation of the cameras in space can be estimated, which then allows to convert a pair of corresponding image points to 3-D projection rays to obtain the 3-D object point via triangulation. The main difference to previous approaches lies in the combination of the large field-of-view of hemispheric sky imagers and a dense stereo reconstruction approach to obtain more complete 3-D geometries, which is favorable when deriving morphological information, post-processing, classification or scientific visualization. The approach allows for a generally improved exploitation of the image content provided by the cameras and often also resolves low-contrast regions (e.g. at the cloud base) reasonably well. Alongside the dense stereo approach, a more accurate multi-view reconstruction technique is applied, which permits an empirical comparison between the sparse, feature-based approaches used in previous studies and the dense stereo method applied in this thesis.

A novelty is the comparison of the stereo reconstructions to the reflectivity profiles from a local cloud radar, which allows to evaluate the accuracy of the obtained cloud morphologies in certain situations. A direct comparison with the cloud boundary height observations from a lidar-ceilometer is used to assess the accuracy of the method.

The high temporal resolution of camera system in the range of seconds and less further allows to compute 3-D cloud motion vectors by tracking clouds over an image sequence (*optical flow*) and to combine this with the 3-D cloud geometries obtained from the stereo method. This permits derivation of atmospheric motion, and detection and classification of distinctive motion patterns by analyzing the vertical profile of velocity and direction, which was not done so far. The cloud motion vectors are compared to observations from a co-located wind-lidar. Case studies are presented for convective boundary-layer clouds and stratiform clouds, including an example of vertical wind-shear. For cloud layers, the motion field is compared to observations by a wind-lidar.

The main contribution of this thesis is to show that the proposed method

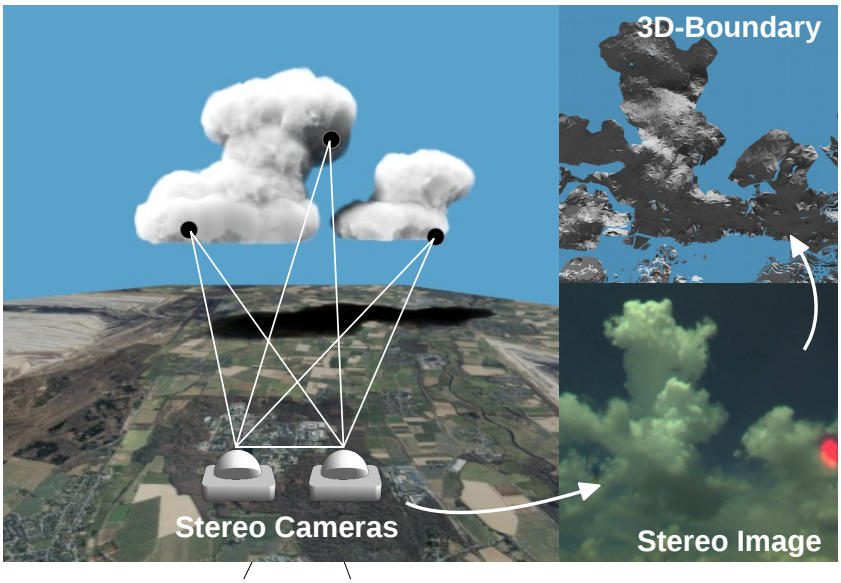


Figure 1.3: 3-D reconstruction of cloud boundaries with ground-based hemispheric stereo cameras. **Left**: mutually visible parts of a cloud are identified and matched across the images, and then reconstructed via triangulation. **Right**: The result is a dense, approximate 3-D cloud boundary that can be used to estimate size, height, shape, spacing and evolution of clouds. **Lower Right**: Detail view of a cloud. **Upper Right**: Corresponding shaded 3-D cloud boundary mesh made with *Blender* (Blender-Foundation, 2016). Map data: *Land NRW (2017) Datenlizenz Deutschland - Namensnennung - Version 2.0 (www.govdata.de/dl-de/by-2-0).*

allows to derive cloud morphology and life-cycle information, which is relevant for LES validation, studies of cloud evolution and cloud radiative effect. Besides cloud top and base height, which was already addressed in previous studies, this includes cloud tilting angle, perimeter, equivalent diameter, height and further geometric statistics.

1.3 Outline

This thesis is organized as follows. Chapter 2 first introduces relevant aspects of digital image processing and camera calibration of hemispheric sky cameras. This includes the working principles of digital cameras, image processing routines, a description of the sky imager system used in this thesis and an algorithmic calibration procedure for both the internal projection parameters of cameras with fisheye lenses and the orientation of the camera in space using stars as reference. Basic geometric relations relevant for 3-D stereo reconstruction, such as the relative orientation (epipolar geometry) between two cameras and its estimation using stereo images, are described in Chapter 3. The prin-

ciples of stereo analysis and related image transformation are explained and techniques are discussed that establish image point correspondences automatically. In Chapter 4 the reconstructed cloud morphologies are evaluated against cloud radar and lidar-ceilometer observations. The case studies are designed to examine the reliability and accuracy of the reconstructions both qualitatively and empirically. The case of a towering cumulus cloud is used to compare two employed stereo methods. Chapter 5 covers the estimation of cloud motion vectors via combination of image tracking and the obtained 3-D geometries from the stereo reconstruction. The derived cloud motion is compared to observations from a wind-lidar and is discussed for a developing cumulus under vertical wind-shear. The value of 3-D stereo reconstruction of clouds for LES validation and 3-D radiative transfer is addresses in Chapter 6. The 3-D reconstructions obtained from two stereo pairs, separated by several kilometers, are used to obtain more complete 3-D cloud geometries, which are then used to conduct a statistical life-cycle analysis of a cloud. A detailed view is provided on the 3-D cloud morphology by analyzing a developing cumulus under wind-shear. The last chapter, Chapter 7, summarizes the results of this thesis and develops a perspective towards future sky-imager-based cloud monitoring.

2 Technical Aspects of Hemispheric Sky Imagers

This chapter introduces fundamentals of digital image processing, including the working principles of modern digital cameras in Sec. 2.1 and image preprocessing routines used to monitor the cloudy sky and tailor 3-D cloud reconstructions in Sec. 2.3. The camera system that was part of the experimental campaigns is described in detail in Sec. 2.2 and compared to the design implemented in previous studies. Sec. 2.4 encompasses the photogrammetric aspects of an individual camera system. A mathematical formulation is provided for the internal orientation (projective properties) and external orientation (localization and orientation in space), followed by a description of the numeric routine to compute the model parameters.

2.1 Digital Cameras

The two fundamental components of a digital camera are the sensor and the lens optics. The sensor of a camera acts as a projection plane and incorporates numerous sensor elements (semi-conductors) on its surface aligned in a regular grid. Based on the photoelectric effect, incoming light (photons) dissociates electrons and creates an electric charge in the semiconductor that is quantified and converted to discrete intensity values.

To obtain spectral information (color), the sensor is usually covered by a repeating pattern of filters (*Bayer filter*), each covering a 2x2 grid of sensor elements. Such a filter consists of three types of spectral filters (1 $\times$ red, 2 $\times$ green, 1 $\times$ blue) that provide the typical three-channel color information suited for human vision. The two green filters account for the higher sensitivity of the human eye in this spectral range. The additional color information, however, reduces the effective resolution of the sensor to 25 percent, because full color information (red, green and blue) is only available within each 2x2 block of sensor elements. A full resolution color image is obtained by interpolating the missing color information from neighboring segments.

Sensor resolutions range from a few to tens of megapixels (10^6 sensor elements) for current consumer cameras to hundreds of megapixels, as used for aerial photography. A single sensor element acts as a collector of incoming photons: the stronger the light intensity the more photons hit the sensor and are converted to an electric charge. Since a larger photo-diode is capable to catch

more photons, the result is a higher charging rate and thus an increased sensitivity. Applications that demand high image quality in low-light situations thus employ sensors with larger elements at the cost of resolution. Further aspects that influence the sensors sensitivity to light are the transmittance of the Bayer filter and the quantum efficiency of the photo-diode, i.e. the probability of the diode to convert an incoming photon into electrical charge (Luhmann, 2018). While a low sensitivity may be compensated by signal amplification, this also has the undesirable effect of an amplified noise signal. Among the different types of image noise, photon noise is probably the most critical noise type as it randomly disturbs the image signal and effectively deteriorates low-contrast image regions that are typical for clouds. Photon noise is well described by a Poisson process: for a large number of photons hitting the diode, the measured intensities follow approximately a normal distribution, resulting in a signal-to-noise-ratio (SNR) that is proportional to the square-root of the light intensity. Hence, longer integration times generally lead to a better SNR if adequately supported by the sensor.

The *dynamic range* (DR) of intensities resolved by the sensor is defined as the ratio between maximum and minimum intensity given a certain *exposure time*, i.e. the time interval during which the sensor is exposed to light, and is measured in decibels[1]. A small dynamic range may cause over- or under-saturated image regions, which then appear homogeneously bright or dark. In sky photography, the direct sun light, a bright cloud top or the strong forward-scattering by optically thin clouds often cause over-saturation, which may be compensated by lowering the exposure time. This might, however, result in other areas appear too dark, making them vulnerable to image noise due to a lower SNR in those regions. There are hardware- and software-based solutions aimed at increasing the dynamic range. One option is to record a series of images with varying exposure times and combine them to a high-dynamic-range (HDR) image as each exposure contributes information related to a high-, medium- or low-intensity image region. The downside of this technique is a motion blur in case of dynamic objects (such as moving clouds). Hardware solutions of HDR include, among others, larger capacities of the sensor elements (expensive), locally adaptive exposure times or logarithmic intensity mapping. For more detailed information, see El Gamal (2002).

Modern sensors are either based on the CCD- (Charged Coupled Devices) or CMOS-technology (Complementary metal–oxide–semiconductor), which differ in the effective size of the photo diode, signal amplification, quantum efficiency and dynamic range (Sonka et al., 2007). The very high *fill factor* of CCD sensors, i.e. the area of a sensor element used to collect photons, results in a comparatively high sensitivity in low-light conditions. The recent developments in CMOS sensor technology, such as the application of micro-lenses (IDS, 2013), reduced these deficit of CMOS significantly. CMOS is currently

[1] $\mathrm{DR} = 20 \log_{10}(I_{max}/I_{min})[\mathrm{dB}]$

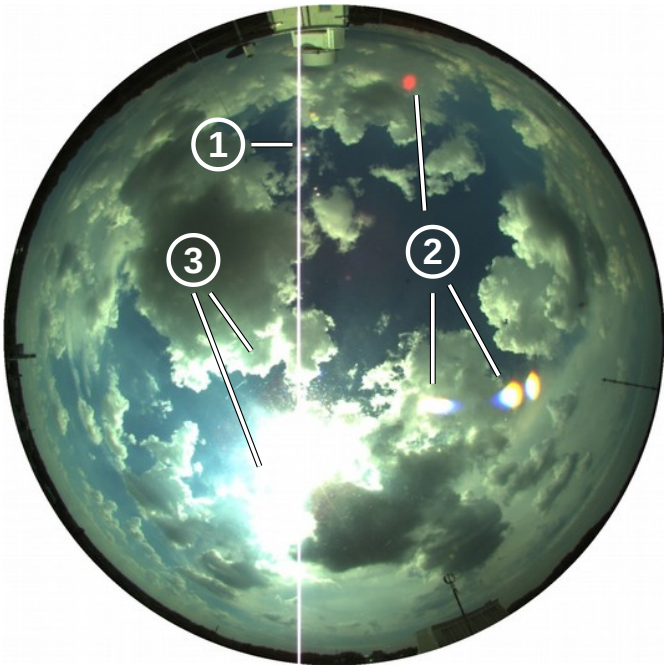


Figure 2.1: Common deteriorations of the sky image: (1) over-saturation of sensor elements and the resulting *blooming effect* due to high radiant intensity by the sun, (2) specular reflections by the lens optics and (3) intense scattering of clouds that also results in over-saturated sensor elements.

considered a state-of-the-art sensor technology due to a large dynamic range, individually accessible sensor elements and fast read-out speed.

The lens of a camera captures a cone of light rays originating from an object and focuses it on the camera sensor . Assuming a pinhole camera with a small aperture on one side and a small patch on the projection plane on the other side, the diameter of the aperture regulates both the amount of light hitting the patch and the sharpness of the resulting image. A larger aperture increases the cone of light hitting the patch, thus reducing the effective spatial resolution. Based on the refraction of light, a lens allows to collect more light from an object, while maintaining a decent aperture size (Forsyth and Ponce, 2012). The lens optics determines the way light rays are projected onto the sensor. Besides the specifications of the sensor and its elements, information about the lens type is crucial for photogrammetric applications as it determines the mathematical model of the projection (Sec. 2.4.2). Since the working principle of a lens is based on the refraction of light, its spectral dependency may cause deviations in the projection for different wavelengths (*chromatic aberration*). The difference bias is lowest near the projection center and may increase to a few pixels towards the borders, depending on the lens quality.

In addition to signal noise or dynamic range limitations, some application-specific deteriorations of the image may occur (Figure 2.1). Specular reflec-

tions, which are caused by the complex interplay of the different parts of the lens, usually only affect the image quality locally. The housing of the cameras used in this thesis included an additional glass dome through which the camera observed the sky; reflections of the outer case or parts of the interior are visible at some occasions, especially at clear sky and sunny conditions. Also, a strong over-saturation of the CCD-sensor elements, which results in a spill-over of electric charge, can saturate neighboring elements. As CCD sensors perform a line-by-line readout, the additional charge spreads along the readout lane, resulting in a vertically saturated line. The sun - if visible - permanently produces such a *blooming effect* due to its extreme intensity.

2.2 Sky Imager Design

Traditional application fields of sky imager systems are the estimation of cloud cover and cloud type. Beyond this scope, some experimental studies investigated the retrieval of sky radiance (Tohsing et al., 2013; Román et al., 2012), cloud optical depth (Olmo et al., 2008) or short-term prediction of solar irradiance (Urquhart et al., 2015, 2016; Schmidt et al., 2016).

Most systems agree on the principal construction designs, such as a protective case, a glass dome on top, an additional computer for control and data storage, external power supply, network connection and a radiator to guarantee operational temperatures and prevent condensation on the glass dome. In some cases, the imager was installed within the local infrastructure, which made an additional computing unit obsolete. For some applications, a commercial, panoramic surveillance camera may be a viable choice as it largely avoids a costly and time consuming development and produces images with reasonable quality. A different solution is realized for example by the *Total Sky Imager* (TSI) manufactured by Yankee Environmental Systems (YES) that uses a catadioptric design with a conventional perspective camera looking downward on a hyperbolic mirror. One design disadvantage for sky observation is that camera and mounting are always visible in the images, thereby leaving parts of the sky unobserved. A completely different and minimalistic approach was pursued by Veikherman et al. (2015) who used Raspberry Pi modules with manually fixed fish-eye adapters. Each module was solar-powered - a solution that is not generally applicable - and maintained a connection to a central database to which images were frequently uploaded.

The sky imager systems used in this study are hosted by a waterproof, protective and portable plastic case, which has been customized to provide an upward view through the case by an acrylic glass dome, connectors for power supply and a network cable. Stable mounting of the camera device is guaranteed by a metal installation within the case. In order to prevent condensation at the glass dome, the case is equipped with a fan, a radiator and a drying agent. The employed camera device is a network camera of type

Figure 2.2: **Left**: Internal design of the hemispheric sky imager used in this study, including attached fan, metal mounting for the camera device, radiator and electric supply. **Right**: Sky imager on a roof within the Forschungszentrum Jülich GmBH.

IDS uEye GigE UI-2280SE equipped with a $2/3''$ CCD sensor by Sony, which offers a maximum resolution of 2448×2048 pixels. As lens optics serves a Fujinon FE185C057HA-1 C-Mount Fisheye adapter, designed for sensors of 5 megapixel resolution, with a fixed focus of f = 1.8 mm and a maximum view field of 185°.

A self-developed application based on the IDS C++ Software Development Kit (IDS, 2013) controls the camera and is hosted by a compact computer, which also provides storage for the image data. Connection between camera and computer is realized via a network cable using the GigE Vision standard. Synchronization of the recordings is done by regular system time updates using the time signal of a conventional GPS receiver of type Navilock NL-302U. Although the serial interface of the USB connection induces a - compared to the precision of the time signal - relatively large uncertainty, the absolute error lies in the range of some tens of milliseconds, which is sufficiently precise for this application. Before using GPS devices for synchronization, the system time was updated via requests to an NTP server, which sometimes failed to answer the requests, resulting in an unchanged system time and accumulated time differences of up to a few seconds. The effect of such an asynchronous recording on the 3-D reconstruction can be substantial and is discussed in Sec. 3.7.

Because the cameras are usually distributed over a large area and installed on various types of buildings and a local internet connection - wired or wireless - is often not available, the current camera network is decentralized, which also requires a higher effort regarding hardware and software, compared to a centralized alternative with server-controlled recordings. According to Veikherman et al. (2015), future operational sky imager networks may have a more centralized architecture to overcome the need of synchronization. Regular data transfer and centralized storage should also reduce the danger of data losses,

which can be critical even if only one camera drops out. A centralized solution, however, requires a stable connection and sufficient transfer volume.

2.3 Image Preprocessing

Image processing routines are used to optimize image quality and extract information required for further analysis. Modern computer vision algorithms are designed to be robust to a number of image deteriorations, such as image noise, radiometric differences or geometric transformations. Hence, much of the overall processing in 3-D vision is included within individual routines. The most important application-specific image processing steps are described below.

2.3.1 Cloud Mask

A cloud mask is the result of an image classification that assigns to each pixel the category *thin cloud*, *opaque cloud* or *sky*, or simply *cloud/sky*. Cloud masks are primarily used to provide automated estimations of cloud cover (Heinle et al., 2010). The classification is often based on the ratio or the difference between the intensity of the red and blue color channel. While being quite effective, such a simple discrimination function cannot cope with all sky conditions and visual effects caused by scattered light. Rayleigh scattering, for example, causes the clear sky to appear brighter near the sun and towards the horizon, or produces a red-shift during dusk and dawn, which has significant impacts the classification. More sophisticated approaches incorporate neural networks to cope with the spatial, temporal and condition-specific complexity of the task (Onishi and Sugiyama, 2017).

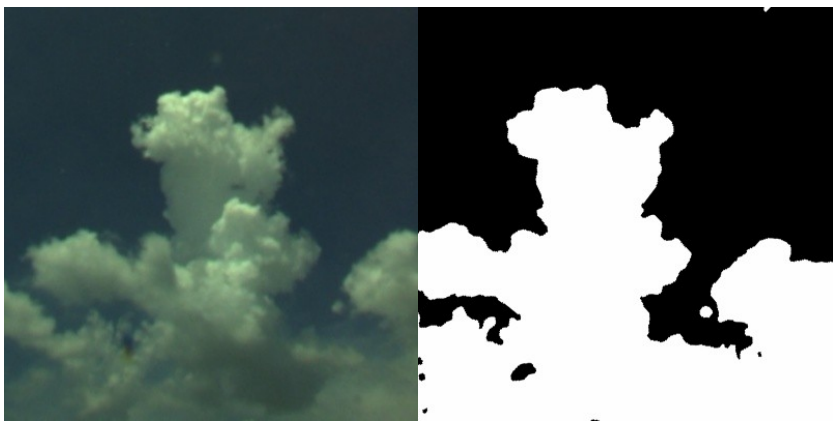


Figure 2.3: A cloud mask (right image) discriminates between cloudy (white) and non-cloudy (black) pixels based on the intensity in the color channels of the original (left image).

This study makes use of cloud masks to filter out artifacts from the stereo analysis and prevent "false" reconstructions in areas where no cloud was detected. Also, cloud masks related to different camera locations (views) provide effective visibility constraints to filter and regulate the resulting 3-D reconstruction.

A simple ratio-based cloud mask is used, which provides reasonable results in most situations (Figure 2.3). Given an image pixel at $\mathbf{x} = [u, v]^\top$, it is classified as cloudy or cloud-free according to

$$f_{cloud}(u,v) = \begin{cases} \text{cloud} & \text{for} \quad (I_b(u,v) - I_r(u,v))/(I_b(u,v) + I_r(u,v)) \geq 0.3 \\ \text{sky} & \text{for} \quad (I_b(u,v) - I_r(u,v))/(I_b(u,v) + I_r(u,v)) < 0.3 \end{cases}$$

Here I_r and I_b denote the image intensity of the red and the blue channel, respectively.

2.3.2 Radiometric Equalization

Radiometric differences between the stereo images, such as an overall brightness or color difference, can affect subsequent image analysis like cloud mask computation or stereo analysis. During the experimental stage of this study, differences in illumination are hard to avoid because the different locations and views on the sky scenery and the automatic exposure adjustment often lead to moderate differences in the radiometric conditions (e.g. one camera is exposed to direct sunlight, while the other camera is not). Radiometric differences of the two stereo images are corrected by an intensity value based on the difference between the grey level distributions (histograms).

Let $h_1(v)$ and $h_2(v)$ be the histograms of two stereo images I_1 and I_2. Both histograms are constructed from the circular region of each fisheye image. To avoid the influence of over-saturated pixels on the resulting histogram, only intensity values between the empirically determined thresholds $I_{min} = 10$ and $I_{max} = 235$ are included.

Both images are then corrected according to the differences between their median intensity values:

$$I_{1,new} = I_{1,old} + 0.5\ (\text{median}(h_2) - \text{median}(h_1))$$
$$I_{2,new} = I_{2,old} + 0.5\ (\text{median}(h_1) - \text{median}(h_2))$$

2.3.3 Contrast Enhancement

A frequent problem in the stereo reconstruction of clouds is due to the poor contrast in cloudy image areas. In many situations, however, clouds do generally not appear completely homogeneous and leave some potential for contrast enhancement to better exploit small variations in the image signal. A standard method to enhance contrast is histogram equalization, which is particularly

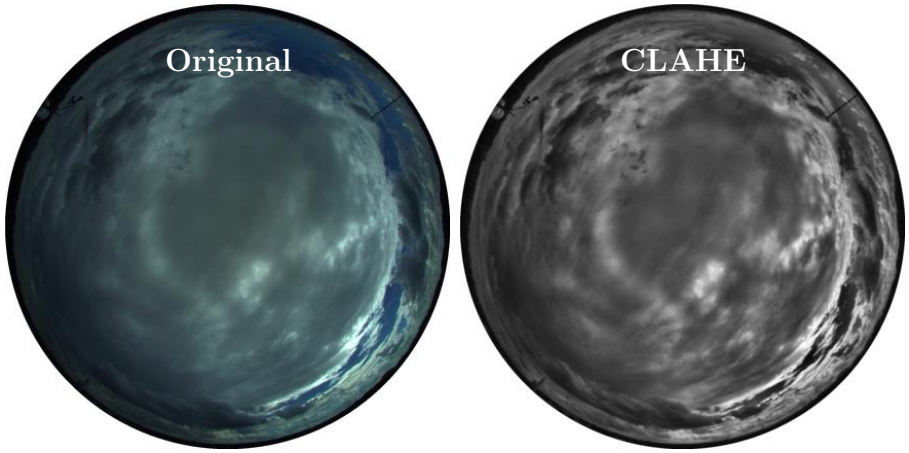


Figure 2.4: Left: Original fisheye image before application of CLAHE. **Right**: Image after application of CLAHE.

useful if the image signal only occupies a small range of intensity values. In this case, the histogram can be stretched to cover the full intensity range, with the consequence that small variations in the original image are enhanced. The premise of a relatively small range of occupied intensities is, however, rarely given in sky photography, except for overcast situations. In general, the low-intensity areas of the sky or cloud bases contrast with the high-intensity areas of reflecting cloud tops, making histogram equalization quite useless.

This problem can be solved by limiting the equalization to small regions of the image in which the desired conditions hold. *Contrast-Limited Adaptive Histogram Equalization* (CLAHE) subdivides an image into a regular grid and performs a local histogram equalization (Pizer et al., 1990). The resulting intensity value of each pixel is computed via bilinear interpolation between each box. Figure 2.4 shows the application of CLAHE to an image of a cloud layer. The low contrast in the original image is boosted and increases the available contrast substantially. To limit the amplification of image noise in homogeneous areas, the original histogram is cut at a predefined value and equally re-distributed in the cumulative histogram used for the redefinition of the intensity values. The *OpenCV* implementation (Bradski, 2000) of CLAHE was used extensively in this thesis and often led to a significant increase in valid matches during stereo analysis (Sec. 3.5).

2.4 Geometric Camera Calibration

Geometric camera calibration provides a functional relation for the mapping between a 3-D object point $\mathbf{X}$ and its corresponding projection $\mathbf{x}$ in the image. A *camera model* defines a set of parameters and geometric relations to formulate the projection, which includes the *internal* and *external* orientation of the camera. While the external orientation is related to the spatial location and

rotation of a camera in space, the internal orientation describes the projection in terms of the lens optics and the sensor. This section addresses geometric model and calibration for a single camera. The projective relationships between two cameras are described in Chapter 3.

2.4.1 Projective Space and Homogeneous Coordinates

The concept of a *projective space* is closely related to the projection conducted by a camera. A camera performs a *central-projection*, i.e. each 3-D object point is projected along a projection ray through the camera center. Hence, any point that lies on a projection ray is mapped onto the same image point, making them projectively equivalent. A projective space is an extension of a real space as it introduces the concept of an equivalence relation or equivalence class that relates projectively equivalent elements to each other.

The point $\tilde{\mathbf{x}}$ denotes the *homogeneous* form and $\mathbf{x}$ the *inhomogeneous* form of a point. Both are related to each other via:

$$\tilde{\mathbf{x}} = [x_1, \ldots, x_n, 1]^\top = [\mathbf{x}, 1]^\top$$

Two homogeneous points $\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2 \in \mathbb{P}^n$ are projectively equivalent, i.e. represent the same point $\mathbf{x} \in \mathbb{R}^n$, if

$$\tilde{\mathbf{x}}_1 \equiv \tilde{\mathbf{x}}_2 \quad \Leftrightarrow \quad \tilde{\mathbf{x}}_1 = \lambda\, \tilde{\mathbf{x}}_2 \qquad \forall \lambda \in \mathbb{R}_{\neq 0}.$$

The use of homogeneous coordinates allows to formulate projective transformations as matrix-vector product. More important, homogeneous coordinates can be used to handle *points at infinity*, i.e. points at infinite (or very large) distance, which can improve the estimation of camera orientation, as was shown in Schneider et al. (2012). Incorporation of points at infinity is possible by setting the last coordinate to zero

$$\mathbf{X}_\infty \longrightarrow \tilde{\mathbf{X}}_\infty = [x''_1, ..., x''_n, 0]^\top,$$

where $\mathbf{x}'' = [x''_1, ..., x''_n]^\top$ is a vector indicating the direction of the projection ray related to $\mathbf{X}_\infty$.

2.4.2 Internal Orientation

The intrinsic camera model defines two basic components, the *optical axis* and the *image plane*, as shown in Figure 2.5a for a hemispheric camera. The optical axis is perpendicular to the image plane and passes through the camera center $\mathbf{C}$. The orientation of the image plane and the optical axis define the camera reference system $\mathbf{\Omega_C}$, located at $\mathbf{C}$. If $\mathbf{X_C}$ denotes an object point in coordinates relative to the camera reference system, its projection to the image point $\mathbf{x}$ depends on the refractive properties of the lens optics and the

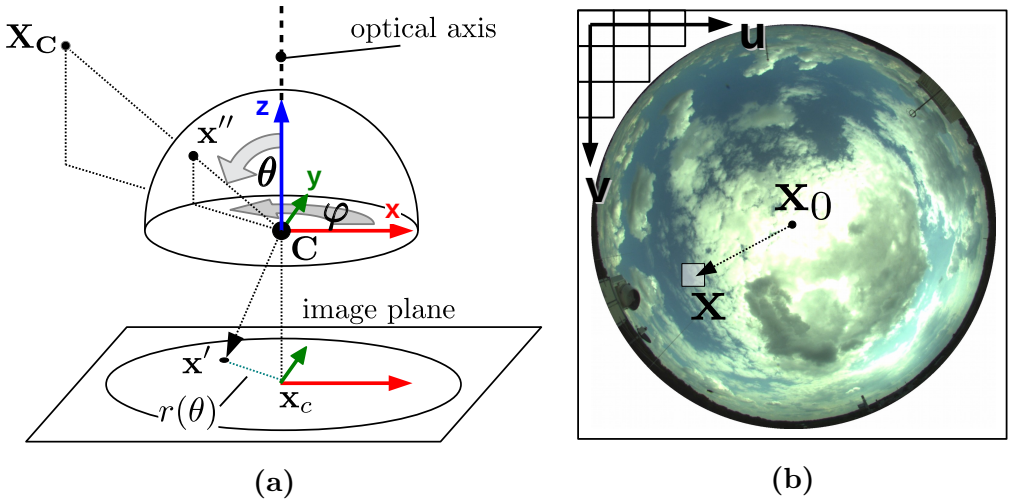


Figure 2.5: Left: Camera model of a hemispheric camera with object point $\mathbf{X_C}$ and its projection $\mathbf{x}''$ on the image hemisphere, which is mapped to $\mathbf{x}'$ on the image plane, given in normalized image coordinates. The angles φ and θ are based on the camera coordinate system defined by the orientation of the image plane and the optical axis, while $r(\theta)$ is only a function of θ. **Right**: An affine transformation then maps $\mathbf{x}'$ to $\mathbf{x}$ in the sensor coordinate system. Adapted from Abraham and Förstner (2005).

technical specifications of the sensor. Hence, one may distinguish between a mapping of $\mathbf{X_C}$ to $\mathbf{x}'$ on the image plane, and a subsequent mapping of $\mathbf{x}'$ to $\mathbf{x}$ on the sensor plane, given in pixel coordinates.

The central projection of a camera maps each object point $\mathbf{X_C}$ along the corresponding projection ray to $\mathbf{x}''$ on the unit (hemi-)sphere, which yields

$$\mathbf{x}'' = \begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix} = \begin{bmatrix} \cos\varphi \sin\theta \\ \sin\varphi \sin\theta \\ \cos\theta \end{bmatrix} . \tag{2.1}$$

The spherical coordinates given in Eq. (2.1) provide the azimuth angle φ, i.e. the angle defined by the vertical projection of $\mathbf{x}''$ on the x-y-plane of $\mathbf{\Omega_C}$, and the zenith angle θ enclosed by the optical axis and the projection ray. Note that throughout the rest of this study the unit vector $\mathbf{x}''$ refers to the respective projection ray of the same direction.

The point $\mathbf{x}'$ on the image plane is given by

$$\mathbf{x}' = \begin{bmatrix} \cos\varphi\, r(\theta) \\ \sin\varphi\, r(\theta) \end{bmatrix}, \qquad \text{where} \qquad \begin{aligned} \varphi &= \operatorname{atan2}(y'', x'') \\ \theta &= \operatorname{acos}(z'') \end{aligned} \tag{2.2}$$

and atan2 is defined as

$$\operatorname{atan2}(x,y) = \begin{cases} \arctan\frac{y}{x} & \text{for } x > 0, \\ \arctan\frac{y}{x} + \pi & \text{for } x < 0 \wedge y \geq 0, \\ \arctan\frac{y}{x} - \pi & \text{for } x < 0 \wedge y < 0, \\ \operatorname{sgn}(y) \cdot \frac{\pi}{2} & \text{for } x = 0 \end{cases}. \tag{2.3}$$

The point $\mathbf{x}'$ is determined by the azimuth angle φ and the *projection function* $r(\theta)$. The projection function, which depends only on the angle θ, is responsible for the radial distance of $\mathbf{x}'$ to the *principal point* $\mathbf{x}_\mathrm{C}$ on the image plane and is different for a fisheye lens compared to an ordinary pinhole-type lens. Following Abraham and Förstner (2005), the most common projection functions are:

$$\begin{aligned} r(\theta) &= \tan\theta && \text{perspective} \quad \text{(pinhole cameras)} \\ r(\theta) &= \tan\frac{\theta}{2} && \text{stereo-graphic} \\ r(\theta) &= \theta && \text{equidistant} \quad \text{(used in this thesis)} \\ r(\theta) &= \sin\theta && \text{orthogonal} \\ r(\theta) &= \sin\frac{\theta}{2} && \text{equisolid angle.} \end{aligned}$$

The perspective projection function describes the mapping of the classical pinhole camera model, i.e. the direct projection of an object point on the image plane along its projection ray. An important limitation of the perspective model becomes apparent if the zenith angle θ of the projection ray approaches 90°. In this case, the point $\mathbf{x}'$ is mapped at infinite distance; the image becomes severely distorted. The projection of a fisheye lens can often be described by the *equidistant* mapping, which is also used for the sky imagers used in this thesis. In some cases, however, the equisolid angle model better models the projection for θ close to 90°.

The point $\mathbf{x}'$ on the image plane is given in *normalized image coordinates* of the corresponding point $\mathbf{x} = [u, v]^\top$ in the actual image, meaning that it abstracts from any specific camera realization w.r.t. focal length and sensor specifications. An affine transformation describes the mapping between $\mathbf{x}'$ in the (Euclidean) reference system of the image plane and $\mathbf{x}$ in the (affine) reference system of the sensor plane. Using homogeneous coordinates allows its formulation as matrix-vector product:

$$\tilde{\mathbf{x}} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{K}\,\tilde{\mathbf{x}}' = \begin{bmatrix} ac & s & u_0 \\ 0 & c & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix}. \tag{2.4}$$

The *calibration matrix* $\mathbf{K}$ encapsulates key parameters that must be estimated during camera calibration. The *camera constant* c denotes the product of the

focal length f, i.e. the distance between camera center $\mathbf{C}$ and the image plane, and the number of sensor elements per unit length of the image plane. The skewness parameter s accounts for a relative horizontal shift between individual rows of a sensor, but is often be ignored ($s = 0$). The aspect ratio a is the ratio of the horizontal and vertical extent of each sensor element, and $\mathbf{x}_0 = [u_0, v_0]^\top$ represents the principal point $\mathbf{x}_\mathrm{C}$ in pixel coordinates. Based on this idealized model, the mapping of $\mathbf{X_C}$ to $\mathbf{x}$ can then be written as

$$\mathbf{x} = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a\,c\ \cos\varphi\ r(\theta) \ + \ s\ \sin\varphi\ r(\theta) \ + \ u_0 \\ c\ \sin\varphi\ r(\theta) \ + \ v_0 \end{bmatrix}. \tag{2.5}$$

The radial projection function $r(\theta)$ of real lenses, however, does usually not match this idealized model exactly. A certain degree of additional distortion, which is often referred to as *barrel* or *pincushion* distortion, affects the projection. These are included into the model by adding polynomials to the coordinates of $\mathbf{x}$ during the calibration procedure. One distinguishes between symmetric distortion, also known as radial distortion, and asymmetric distortion. Symmetric distortion is modeled by even-powered polynomials and is generally more significant than asymmetric distortion, which can be modeled by odd-powered polynomials. In this application, however, asymmetric distortion can also be significant due to an additional refraction by the glass dome that protects the camera from environmental effects. Asymmetric distortion may occur when the optical axis is not correctly aligned to the image sensor, so that the projection rays are slightly skewed (*decentering distortion*).

The additional correction terms used in this study then read as

$$\begin{aligned} \triangle u &= \sum_{i=1}^{3} A_i\,\hat{u}\,\hat{r}^{2i} + B_1(\hat{r}^2 + 2\hat{u}) + B_2 2\hat{u}\hat{v} \\ \triangle v &= \sum_{i=1}^{3} A_i\,\hat{v}\,\hat{r}^{2i} + B_2(\hat{r}^2 + 2\hat{v}) + B_1 2\hat{u}\hat{v} \end{aligned} \qquad \text{with} \quad \begin{aligned} \hat{u} &= u - u_0, \\ \hat{v} &= v - v_0, \\ \hat{r} &= \sqrt{\hat{u}^2 + \hat{v}^2} \end{aligned} \quad . \tag{2.6}$$

The coefficients A_i control the magnitude of the even-powered polynomials for the radial-symmetric distortion, while B_1 and B_2 are the coefficients related to the radial-asymmetric and tangential distortion (Brown, 1966). Including these terms into Eq. (2.5) gives the formulation for the distorted image coordinates

$$\mathbf{x}^* = \begin{bmatrix} u^* \\ v^* \end{bmatrix} = \begin{bmatrix} a\,c\ \cos\varphi\ r(\theta) \ + \ s\ \sin\varphi\ r(\theta) \ + \ u_0 + \triangle u \\ c\ \sin\varphi\ r(\theta) \ + \ v_0 + \triangle v \end{bmatrix}. \tag{2.7}$$

Throughout the rest of this study, it is assumed that $\mathbf{x}$ is corrected for this kind of distortion and thus relates to the camera model according to Eq. (2.5). Eq. (2.7) is only relevant for camera calibration rather than for actual 3-D reconstruction. It is further assumed that images are corrected for any skewness or aspect ratio, i.e. $s = 0$ and $a = 1$.

The inverse mapping of an image point $\mathbf{x}$ to its corresponding projection ray $\mathbf{x}''$ requires normalization of $\mathbf{x}$ according to

$$\mathbf{x}' = \begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} (v - v_0)/c \\ (u - u_0)/c \end{bmatrix}. \tag{2.8}$$

The projection ray is then given by

$$\mathbf{x}'' = \begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix} = \begin{bmatrix} \frac{u'}{r}\sin(r) \\ \frac{v'}{r}\sin(r) \\ \cos(r) \end{bmatrix} \qquad \text{with} \quad r = \sqrt{(u')^2 + (v')^2}. \tag{2.9}$$

The model presented here is *one* way to formulate the projection geometry of a fisheye camera. Alternative approaches make use of polynomials, for example Chebycheff (Abraham and Hau, 1997), or as a central-projection of light rays onto a model parabola and a subsequent perspective projection of that parabola (Scaramuzza, 2008; Micušık, 2004). The aim is to provide a more generic model that can handle various lens projections and that may deviate considerably from a standard model, such as the equidistant projection. They allow a direct modeling of different kinds of lenses, such as dioptric, catadiotptric and fisheye, without switching to a different projection function $r(\theta)$. In most cases, however, a standard physically-based model as the one described in this section is very well suited and less vulnerable to over-fitting during calibration.

2.4.3 External Orientation

The extrinsic parameters define the *absolute* location and orientation of a camera w.r.t. a global reference system $\mathbf{\Omega_W}$. The object point $\mathbf{X_C}$, related to the camera reference system, can be rewritten in terms of the point $\mathbf{X}$ in the world reference frame, according to

$$\mathbf{X_C} = \mathbf{R}^\top(\mathbf{X} - \mathbf{C}),$$

where $\mathbf{C}$ is the projection center of the camera within the world reference frame. The rotation matrix $\mathbf{R}^\top$ represents the orientation of the camera device, i.e. the mapping from world coordinates to camera coordinates. Thus, a projection ray $\mathbf{x}''$ within the camera reference frame relates to its correspondence within the world reference system through $\mathbf{x}''_\mathbf{W} = \mathbf{R}\,\mathbf{x}''$. The camera orientation has three degrees of freedom and can be written as the product of three consecutive rotations around the (fixed) coordinate system axes of $\mathbf{\Omega_W}$ by the angles α, β and γ:

$$\mathbf{R} = \mathbf{R}_z(\gamma)\,\mathbf{R}_y(\beta)\,\mathbf{R}_x(\alpha),$$

where

$$\mathbf{R}_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & -\sin\gamma \\ 0 & \sin\gamma & \cos\gamma \end{bmatrix}, \qquad \mathbf{R}_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix},$$

$$\mathbf{R}_z(\gamma) = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

2.4.4 Parameter Estimation: Internal Orientation

The first task is to collect a set of correspondences between real world 3-D features of known geometry and their corresponding 2-D projections in the images. Ideally, the calibration data set contains pairs of 2-D/3-D point correspondences. Once a camera model is chosen that adequately describes the principal projection (e.g. the equidistant model of hemispheric cameras), the next step is to derive a first approximation of the parameters before refining them iteratively and minimizing the 2-D geometric error between the observed and predicted image coordinates. The calibration is successful if the magnitude of the residual error distribution reflects the accuracy of the data set. In the following, a more detailed description of the two steps is given.

The acquisition of reference data requires some 3-D features that can be imaged, detected and identified in the recorded images in order to establish a relation between the projections and the real world 3-D coordinates. There are two common ways to achieve this: The first option is to use a sophisticated setup in a laboratory as in Seiz (2003), which provides a number of fixed, clearly detectable and identifiable patterns within a uniformly illuminated room. Unfortunately, such a setup is also comparatively costly and does not allow to calibrate cameras that are fixed and located outdoors.

The alternative is to use a mobile calibration test field, with (approximately) known geometry, which allows a manual placement in the view field of the camera. This can be a 2-D chessboard pattern printed on a flat underground (Zhang, 2000; Scaramuzza, 2008) or - as in this study - a 3-D test field consisting of flat circular markers attached to the interior of an open half-cube, as shown in Figure 2.6 (also called *Tsai grid*). Recording a few dozen images with the marker pattern clearly visible in various poses and different parts of the view field is usually sufficient to obtain a reliable and accurate estimation of the camera geometry over the whole hemisphere. The elliptic projections of the circular markers in the image are detected by a contour algorithm, e.g. a Hough transform. Attached to each side of the cube is also a special group of markers that allows to identify each side and consequently each individual marker on that side. The exact procedure for the correct identification of each marker can be found in Abraham and Hau (1997).

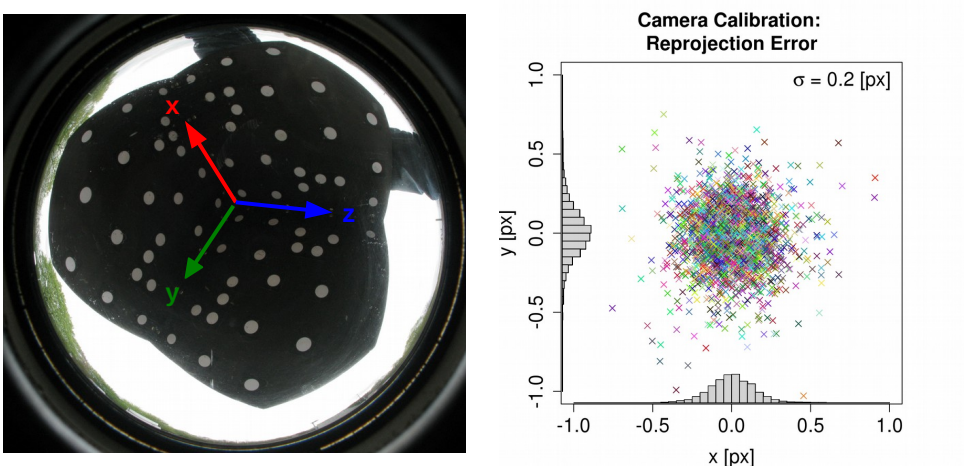


Figure 2.6: Left: Camera calibration with *Cube* test pattern where each white marker defines a 3-D point within the coordinate system of the cube. **Right**: Distribution of the reprojection error in pixels after calibration.

Once the obtained set of 2-D/3-D point correspondences is available, the camera parameters can be estimated. As the complete projection is highly non-linear due to the rotation and the polynomial distortion terms, one needs to find an approximate solution of the intrinsic and extrinsic parameters before refining the parameters iteratively by minimizing the reprojection error between the detected markers in the image and the predicted marker positions by the model parameters.

An initial estimate of both the intrinsic and extrinsic camera parameters can be found by solving for the entries of the camera matrix $\mathbf{P}$, which - like the calibration matrix $\mathbf{K}$ - is a homogeneous entity and is thus only defined up to scale. It encapsulates the calibration matrix $\mathbf{K}$, the rotation matrix $\mathbf{R}^\top$ and the camera center $\mathbf{C}$:

$$\tilde{\mathbf{x}}_i^p = \underbrace{\left[\mathbf{K}\mathbf{R}^\top \mid -\mathbf{K}\mathbf{R}^\top\mathbf{C}\right]}_{3\times 4} \tilde{\mathbf{X}}_i = \mathbf{P}\,\tilde{\mathbf{X}}_i. \tag{2.10}$$

The linear relationship between image and object points allows to obtain an approximate solution by solving for the entries of the camera matrix $\mathbf{P}$; subsequent decomposition then yields $\mathbf{K}$, $\mathbf{R}$ and the camera center $\mathbf{C}$.

The hemispheric camera model does not offer such a linear solution and requires prior knowledge of the intrinsic parameters before orientation and location can be derived. To obtain a reasonable initial estimate, intrinsic parameters like the aspect ratio of the sensor elements or the skewness factor may be neglected at this stage and set to default values, e.g. $sk = 0$, $a = 1$. The principal point $\mathbf{x}_0$ is generally assumed to be located in the image center. The camera constant c can vary considerably among different cameras and

Algorithm 1 Computation of the camera geometry

1. Generate a calibration data set of correspondences between 3-D reference points and 2-D image points:
   (a) Place the test field (Cube) in front of the camera and record 10 to 30 images with the marker pattern clearly visible, in various locations and orientations so that markers are visible over the whole hemispheric view field.
   (b) Apply an image processing routine to detect and identify the marker pattern.
2. Compute the model parameters using the obtained marker-projection-correspondences:
   (a) Use a rough estimate of the camera constant to form the projection rays $\mathbf{x}_i''$ of each identified marker and project it to $\mathbf{x_i}^p$ on the image plane.
   (b) Compute a linear estimate of the camera matrix $\mathbf{P}$ via DLT.
   (c) Decompose $\mathbf{P}$ to obtain $\mathbf{K}$, $\mathbf{R}$ and $\mathbf{C}$.
   (d) Start the bundle adjustment using the initial parameters as starting point.

camera configurations and should not be assigned a default value. Given that the view field of the lens corresponds well to the visible area in the resulting image, a reasonable starting value for the camera constant may be derived, which reproduces the size of the circular view field in the image, as proposed by Kannala and Brandt (2006). The method used in this study simply tries out a number of values, e.g. $c \in \{200, 400, 800, 1600\}$ and keeps the value that produces the least geometric errors.

Given a set of image points and corresponding object points $\mathbf{X}_i$ a linear estimate of $\mathbf{P}$ can be computed via Direct Linear Transformation and decomposed to yield the calibration matrix $\mathbf{K}$, the rotation $\mathbf{R}$ and the camera center $\mathbf{C}$ (Förstner and Wrobel, 2016, pp. 498-500; Hartley and Zisserman, 2000, Sec. 5.2.4).

Once an initial estimate of the parameters is computed, it can be used as a starting point to refine the parameters of the full camera model, which is non-linear in the respective parameters. For this thesis, a *bundle adjustment* was used to compute a least-squares estimate, which simultaneously estimates both the 3-D marker positions in the cube coordinate system and the intrinsic and extrinsic parameters (Abraham and Hau, 1997; Förstner and Wrobel, 2016, Chapter 15). During calibration, it is assumed that the intrinsic parameters remain constant, while the extrinsic parameters have to be estimated with

each additional image. Under the assumption that the localization error of the detected markers follows a normal distribution, the minimization of the *reprojection error*

$$\epsilon^2 = \sum_i \sum_o \|\mathbf{x_{i}}_o - \mathbf{f}(\hat{\mathbf{p}}_\mathbf{I}, \hat{\mathbf{p}}_\mathbf{E}, \hat{\mathbf{X}}_i)\|^2 \tag{2.11}$$

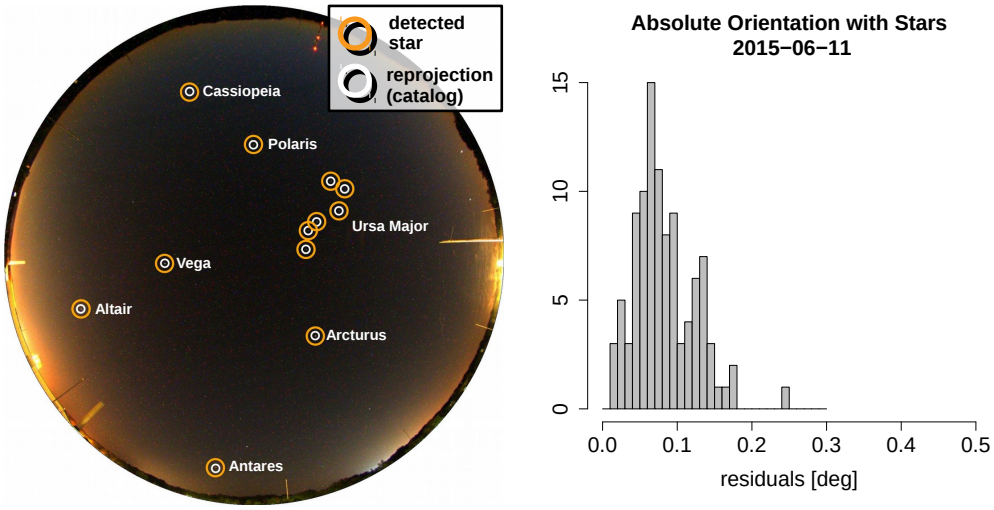
yields an optimal solution. Here, $\hat{\mathbf{p}}_\mathbf{I}$ and $\hat{\mathbf{p}}_\mathbf{E}$ contain the estimated intrinsic and extrinsic parameters of the camera, while $\hat{\mathbf{X}}_i$ is the estimated 3-D location of the marker and $\mathbf{f}$ is the function that maps $\hat{\mathbf{X}}_i$ to its reprojected image point $\hat{\mathbf{x}}_i$ using $\hat{\mathbf{p}}_\mathbf{I}$ and $\hat{\mathbf{p}}_\mathbf{E}$. $\mathbf{x_{i}}_o$ denotes the location of detected marker o in image i.

Following Eq. (2.6), the calibration of the sky imagers in this study included parameters for radial-symmetric distortion terms up to the second degree (A1 and A2) and a fist-degree term for radial-asymmetric distortion (B1), which models the additional refraction caused by the glass dome.

2.4.5 Parameter Estimation: External Orientation

Once the intrinsic parameters are available, location and orientation of the camera within the world reference frame can be derived.

Satellite navigation systems like GPS, GLONASS or the future GALILEO allow to measure the position of each camera with a typical accuracy of 3 to 5 meters using ordinary receivers, 2 m in case of multiple averaged measurements. Additional correction information provided by reference stations with accurately known coordinates can lead to accuracies in the range of 10 to 20 cm, but require special antennas and are commonly not free of charge. The estimation of the absolute camera orientation within the world reference system $\mathbf{\Omega_W}$ must be derived from the image content alone and with sufficient accuracy (e.g. 0.1°) to avoid a bias in a 3-D stereo reconstruction. Since sky imager cameras are designed for sky observation, the identification of potential landmarks in the image is often complicated due to occluding trees or buildings. Three solutions to this problem were proposed and implemented in previous studies. Hu et al. (2010) were able to identify nearby landmarks in combination with a geographic information system (GIS) and successfully obtained accurate orientation parameters. Öktem et al. (2014) was able to use the horizon as reference during their study of marine convection. Seiz (2003) used night sky imagery to detect and identify stars that could be used as reference. During clear sky conditions stars can be recorded and extracted using image processing techniques, while star catalogs provide information about the position of each star in the local night sky. The camera sensor must, however, be able to sense the faint star light via automatic or manual adjustments to the ISO sensitivity, exposure time and aperture. Note that changing the aperture might change the projective properties of the camera, which makes an additional calibration necessary. Another approach is to use the sun as orientation reference, as done by Veikherman et al. (2015). The remainder of this section



(a) **(b)**

Figure 2.7: Left: Long-exposure night-sky images for absolute orientation estimation of a sky imager, including the result of a parameter estimation for 11 June 2015 at 22:00 UTC, JOYCE (meteorological tower visible at the image top). **Right**: Angular reprojection error and parameter uncertainty after the least-squares estimation.

describes the semi-automatic calibration procedure used in this study based on sensed stars.

Given a number of night sky images taken over a time span in a certain interval (e.g. 15 minutes), the stars must be detected and localized. Depending on sensor and recording configuration, the images may be affected by noise, which impairs the reliability to detect and localize stars. The impact of background illumination and fixed noise patterns can be reduced to a minimum by subtracting the median image of the sequence. This avoids a significant deterioration of the signal from the recorded stars, which continue to move throughout the image sequence (Figure 2.7a).

The identification of stars is carried out in two stages. First, a threshold is applied to the median-filtered image, which removes the remaining noise signal and results in a binary image, in which each non-zero pixel cluster represents a potential star. After a contour detection algorithm has extracted each cluster, an exact localization is obtained by computing the center of gravity, based on the intensity values of the cluster (the signal from the stars).

The coordinates from the star catalog are generally given as a pair of the north-aligned azimuth angle φ^N and altitude angle θ_a, so that a star at the northern horizon is located at $\varphi^N = 0°$ and $\theta_a = 0°$. They are converted to

projection rays according to:

$$\mathbf{x''_{cat}} = \begin{bmatrix} \cos(\varphi^N + \pi/2)\cos(\theta_a) \\ -\sin(\varphi^N + \pi/2)\cos(\theta_a) \\ \sin(\theta_a) \end{bmatrix}. \tag{2.12}$$

Finally, the correspondences between the reference stars from the catalog and the extracted stars from the images must be established. Starting with an approximate orientation, represented by the angles ${}^a\alpha, {}^a\beta$ and ${}^a\gamma$, each reference star is mapped to camera coordinates via $\mathbf{R}^\top_{\mathbf{abs}}\ \mathbf{x''_{cat}}$, based on the currently estimated orientation angles, and can be used to find the closest candidate $\mathbf{x''_s}$ from the set of potential stars. The set of established correspondences then serves as input for a least-squares estimation that refines the orientation parameters according to the error

$$\epsilon_i = 1 - ((\mathbf{x''_{cat,i}})^\top\ \hat{\mathbf{R}}_{\mathbf{abs}}\ \mathbf{x''_{s,i}}), \qquad i = 1, ..., n_{stars}. \tag{2.13}$$

The obtained rotation matrix $\hat{\mathbf{R}}_{\mathbf{abs}} = \mathbf{R_z}(\hat{\gamma})\mathbf{R_y}(\hat{\beta})\mathbf{R_x}(\hat{\alpha})$ then minimizes the term in Eq. (2.13) and defines the orientation that fits best to the observed reference data (Figure 2.7b).

3 Stereo Reconstruction with Hemispheric Cameras

This chapter describes the 3-D reconstruction of clouds using pairs of hemispheric sky imagers. The reconstruction is based mainly on a dense stereo analysis, which processes the whole image content rather than a sparse set of image features, such as corners or edges. Stereo reconstruction requires a precisely known relative orientation of the two cameras, as it affects the geometric accuracy of the obtained cloud morphologies. Sec. 3.2 and Sec. 3.5.1 describe its estimation from a set of image point correspondences, extracted from the synchronously recorded stereo images. Stereo matching is introduced in Sec. 3.5, which includes a dense stereo matching algorithm (Sec. 3.5.2) and a high-accuracy multi-view reconstruction technique (Sec. 3.5.3). The latter is assumed to produce more accurate reconstructions in certain situations. Dense matching algorithms often require epipolar rectified stereo images, which are geometrically transformed versions of the original fisheye images. A rectification scheme designed for hemispheric cameras is presented in Sec. 3.3. The theoretical precision of the system and the effect of an asynchronous recording, which could be observed at times during the first experimental phase, is discussed in Sec. 3.6 and Sec. 3.7, respectively.

3.1 Epipolar Geometry

The epipolar geometry defines the geometric relations within a stereo camera setup, as illustrated in Figure 3.1 for two hemispheric cameras. Let $\mathbf{X}_1$ be an object point within $\mathbf{\Omega_1}$, the reference system of camera one, and $\mathbf{X}_2$ the same point in $\mathbf{\Omega_2}$, the reference system of camera two. The relative translation $\mathbf{b}$ between the cameras and their relative orientation $\mathbf{R}$ are given by

$$\mathbf{b} = \mathbf{C_2} - \mathbf{C_1}, \qquad \mathbf{R} = \mathbf{R_1}^\top \mathbf{R_2}. \tag{3.1}$$

The epipolar plane, which is spanned by the baseline vector $\mathbf{b}$ and either of the two projection rays $\mathbf{x}_1''$ and $\mathbf{x}_2''$ intersects each image hemisphere in the *epipolar curves* $\mathbf{l}_1$ and $\mathbf{l}_2$, which denote the respective normal vector of the epipolar plane w.r.t. $\mathbf{\Omega_1}$ and $\mathbf{\Omega_2}$. This geometric relationship leads to the

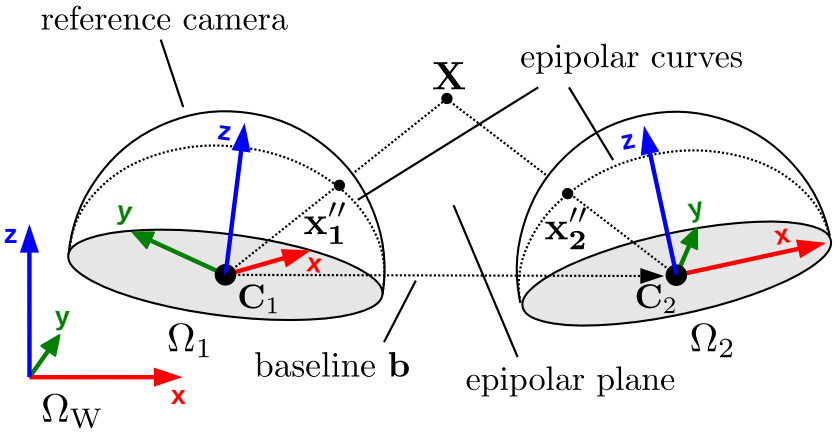


Figure 3.1: Epipolar geometry of two hemispheric cameras located at $\mathbf{C_1}$ and $\mathbf{C_2}$, and an independent orientation within the world coordinate system $\mathbf{\Omega_W}$. The spherical projections $\mathbf{x}_1''$ and $\mathbf{x}_2''$ of a 3-D point $\mathbf{X}$ on the image hemisphere in each camera system and the baseline vector $\mathbf{b}$ span the epipolar plane and can be used to reconstruct $\mathbf{X}$ via triangulation. If only $\mathbf{x}_1''$ or $\mathbf{x}_2''$ is known, its correspondence is restricted to the respective *epipolar curve* on the other image hemisphere. One of the cameras is used as reference for intermediary results from the analysis and reconstruction.

formulation of the *coplanarity constraint*:

$$\begin{aligned}(\mathbf{R_1}\mathbf{X}_1)^\top \left[(\mathbf{C_2}-\mathbf{C_1})\times(\mathbf{R_2}\mathbf{X}_2)\right] = 0 \Leftrightarrow \\ \mathbf{X}_1^\top \quad \mathbf{R_1}^\top \mathbf{S_b} \mathbf{R_2} \quad \mathbf{X}_2 = 0 \Leftrightarrow \\ \mathbf{X}_1^\top \quad \mathbf{E} \quad \mathbf{X}_2 = 0 \end{aligned} \tag{3.2}$$

The matrix $\mathbf{E}$ is called the *essential matrix* and encapsulates information about the relative orientation between two cameras. It is a homogeneous entity and is thus only defined up to scale; each multiple also satisfies Eq. (3.2). It can be written in terms of the reference camera as the product between the skew-symmetric matrix $\mathbf{S_b}$, related to the baseline vector $\mathbf{b} = [x_b, y_b, z_b]^\top$, and the relative rotation matrix $\mathbf{R}$ of Eq. (3.1):

$$\mathbf{E} = \mathbf{S_b}\mathbf{R}, \qquad \text{where} \qquad \mathbf{S_b} = \begin{bmatrix} 0 & -z_b & y_b \\ z_b & 0 & -x_b \\ -y_b & x_b & 0 \end{bmatrix}. \tag{3.3}$$

Given two projection rays represented by the 3-D unit vectors $\mathbf{x}_1''$ and $\mathbf{x}_2''$, the following relations are equivalent:

$$\mathbf{x}_1''^\top \; (k\,\mathbf{E}) \; \mathbf{x}_2'' = 0, \qquad \forall k \in \mathbb{R}, \tag{3.4}$$

$$\mathbf{l}_1 = \mathbf{E}\,\mathbf{x}_2'', \tag{3.5}$$

$$\mathbf{l}_2 = \mathbf{E}^\top \mathbf{x}_1''. \tag{3.6}$$

The coplanarity constraint states that $\mathbf{x}_2''$ is restricted to lie on the epipolar curve defined by the normal vector $\mathbf{l}_2 = \mathbf{E}^\top \mathbf{x}_1''$, and correspondingly for the $\mathbf{x}_1''$ and $\mathbf{l}_1$. The complexity of the correspondence analysis can thus be reduced from two dimensions to one, i.e. along the epipolar curve, which is useful in the context of epipolar image rectification (Sec. 3.3) and stereo matching (Sec. 3.5). Eq. (3.2) can be used to compute the essential matrix based on image point correspondences: Uncertainties in the localization and matching of image points usually lead to deviations from the coplanarity constraint, which can be used as an optimization criteria to compute the parameters of the relative orientation.

Due to the vector product in Eq. (3.3), the essential matrix is singular and has a rank of 2; it has two identical singular values σ_1 and σ_2.

Following Hartley (2001), an essential matrix can be decomposed to yield the rotation and the baseline vector. Given the matrices

$$\mathbf{Z} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Eq. (3.3) can be expanded to

$$\mathbf{E} = \mathbf{SR} = \mathbf{U}\,\mathbf{Diag}([1,1,0])\,\mathbf{V}^\top = \mathbf{UZWV}^\top = \mathbf{UZU}^\top\mathbf{UWV}^\top.$$

There are two solutions of $\mathbf{S_b}$ and $\mathbf{R}$ each:

$$\begin{aligned} \mathbf{S_1} &= \mathbf{UZU}^\top, & \mathbf{R}_1 &= \mathbf{UWV}^\top, \\ \mathbf{S_2} &= \mathbf{UZ}^\top\mathbf{U}^\top, & \mathbf{R}_2 &= \mathbf{UW}^\top\mathbf{V}^\top, \end{aligned}$$

which yields four valid decompositions of $\mathbf{E}$, all of which yield potential solutions for the relative orientation. The solutions for $\mathbf{R}$ are already given so that only the baseline remains to be extracted from $\mathbf{S_1}$ and $\mathbf{S_1}$ respectively. Since $\mathbf{b}$ lies in the null-space of $\mathbf{S}$ and $\mathrm{rank}(\mathbf{S}) = 2$, $\mathbf{b}$ must be the last column of $\mathbf{U}$.

The four solutions for $\mathbf{b}$ and $\mathbf{R}$ are thus

$$\begin{aligned} \mathbf{b}_1 &= \mathbf{u_3}, & \mathbf{R}_1 &= \mathbf{R}, \\ \mathbf{b}_2 &= -\mathbf{u_3}, & \mathbf{R}_2 &= \mathbf{R}^\top. \end{aligned} \tag{3.7}$$

The correct solution is the one, in which the projection rays triangulate in front of both cameras.

A 3-D reconstruction based on the relative orientation alone is only defined within the coordinate system of the reference camera. A meaningful metric reconstruction in world coordinates requires additional knowledge about the absolute orientation $\mathbf{R}_{abs}$ of the reference camera and the length of the real baseline, which can be obtained for example from a satellite navigation system, as described in section Sec. 2.4.5.

Algorithm 2 Computation of the Relative Orientation

1. Collect a number of interest point correspondences (e.g. SIFT features), as described in Sec. 3.5.1.
2. Robust estimation of the essential matrix via RANSAC: Determine the minimum number of iterations N_{iter} for a confidence p and an assumed outlier rate q, according to Eq. (3.9). In each iteration i,
   a) select a random set of at least 8 correspondences and compute a linear estimate ${}^{\mathbf{a}}\mathbf{E}_i$ of the essential matrix, as described in Sec. 3.2.1 and
   b) determine the inlier set of correspondences

   $$\{(\mathbf{x}_1'', \mathbf{x}_2'')\}_{inlier}$$

   where each correspondence satisfies

   $$\mathbf{x}_1''^{\top}\ {}^{a}\mathbf{E}\ \mathbf{x}_2'' < K$$

   and K may be defined according to Eq. (3.8).
   c) Save the current estimate ${}^{a}\mathbf{E}$ and the set of inliers if it contained more elements than any previous set.
3. Use the set of inliers and the corresponding essential matrix ${}^{a}\mathbf{E}$ to compute an optimal estimate of the relative orientation, as described in Sec. 3.2.2.

3.2 Relative Orientation Estimation

The computation of an accurate relative orientation requires an initial estimate of the rotation matrix ${}^{a}\mathbf{R}$ and baseline vector ${}^{a}\mathbf{b}$, and a set of corresponding projection rays that can be used to refine the parameters in terms of the coplanarity constraint (Eq. (3.2)). A feature detection and matching (Sec. 3.5.1) usually provides a set of a few hundred up to a thousand corresponding projection rays, which, however, contains a significant percentage (10% to 20%) of outliers w.r.t. the coplanarity constraint. An approximate solution to the relative orientation and a supporting set of corresponding projection rays is obtained from a RANSAC robust parameter estimation (Chum et al., 2004), which iteratively selects a random set of 8 correspondences, computes the essential matrix (Sec. 3.2.1) and then determines from all correspondences the set that satisfies the coplanarity constraint up to a certain threshold. This set is called the *inlier set*. The one relative orientation that corresponds to

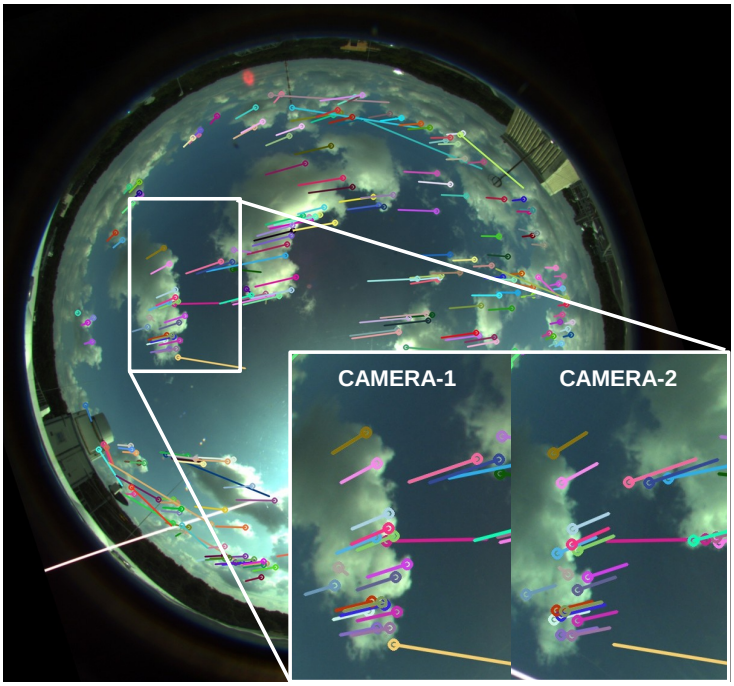


Figure 3.2: Detected interest points can be matched across the stereo images and are marked by the same color. At least 5 correspondences are needed to compute the relative orientation between the cameras and each also provides one 3-D point.

the largest inlier set then defines the approximate solution of ${}^{a}\mathbf{R}$ and ${}^{a}\mathbf{b}$. Following (Förstner and Wrobel, 2016, p. 599), the threshold for the coplanarity condition

$$K = k\,\sigma_x \sqrt{\mathbf{x}_2''^{\top}\,\mathbf{E}^{\top}\mathbf{E}\,\mathbf{x}_2'' + \mathbf{x}_1''^{\top}\,\mathbf{E}\mathbf{E}^{\top}\,\mathbf{x}_1''}, \qquad k > 0, \tag{3.8}$$

reflects the impact of the average angular uncertainty σ_x of the projection rays on the coplanarity constraint. The number of iterations N_{iter} required to select 8 correct inliers depends on the assumed outlier ratio q and the desired confidence p:

$$N_{iter} = \left\lceil \frac{\log(1-p)}{\log(1-(1-q)^8)} \right\rceil. \tag{3.9}$$

The relative orientation is then further refined in an iterative non-linear adjustment, minimizing the residuals of the coplanarity constraint w.r.t. rotation and baseline (Sec. 3.2.2).

3.2.1 Direct Solution of the Essential Matrix

This section describes the direct computation of the essential matrix from at least 8 corresponding projection rays using an adaptation of the original

8-Point-algorithm (Hartley, 2001; Longuet-Higgins, 1981). Given a set of corresponding projection rays derived from matched interest points $\{(\mathbf{x}''_{1i}, \mathbf{x}''_{2i}) \,|\, i \in 1, ..., n\}$, the coplanarity constraint leads to the homogeneous equation system

$$\begin{aligned} \mathbf{x}''^{\top}_{10}\,\mathbf{E}\,\mathbf{x}''_{20} &= 0 \\ &\vdots \\ \mathbf{x}''^{\top}_{1n}\,\mathbf{E}\,\mathbf{x}''_{2n} &= 0 \end{aligned} \qquad \text{with} \qquad \mathbf{E} = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}. \tag{3.10}$$

Each equation can then be rearranged to form a product between the vectors $\mathbf{a}$ and $\mathbf{e}$:

$$\begin{aligned} \mathbf{a}_0\,\mathbf{e} &= 0 \\ &\vdots \\ \mathbf{a}_n\,\mathbf{e} &= 0 \end{aligned} \qquad \Leftrightarrow \qquad \mathbf{A}\,\mathbf{e} = \mathbf{0}, \tag{3.11}$$

where $\mathbf{a}_i$ and $\mathbf{e}$ are defined as

$$\mathbf{a}_i = \begin{bmatrix} x''_{1i}\,x''_{2i}, & x''_{1i}\,y''_{2i}, & x''_{1i}\,z''_{2i}, & y''_{1i}\,x''_{2i}, & y''_{1i}\,y''_{2i}, & y''_{1i}\,z''_{2i}, & z''_{1i}\,x''_{2i}, & z''_{1i}\,y''_{2i}, & z''_{1i}\,z''_{2i} \end{bmatrix},$$

$$\mathbf{e}^{\top} = \begin{bmatrix} e_{11}, & e_{12}, & e_{13}, & e_{21}, & e_{22}, & e_{23}, & e_{31}, & e_{32} & e_{33} \end{bmatrix}.$$

The matrix $\mathbf{A}$ consists of the vertically stacked row vectors $\mathbf{a}_i$ and the parameter vector $\mathbf{e}$ is the essential matrix row-wise reshaped as a 9x1 vector. Eq. (3.11) can be solved via Singular Value Decomposition and $\mathbf{e}$ is found to be the column vector of $\mathbf{V}$ corresponding to the smallest singular value.

The obtained solution yields a matrix $\mathbf{E}'$, which may have full rank due to measurement and localization errors that cause deviations from the coplanarity constraint. The approximate essential matrix ${}^{\mathbf{a}}\mathbf{E}$, which is closest to $\mathbf{E}'$, can be obtained via SVD of $\mathbf{E}'$, where each of the singular values might be non-zero:

$$\mathbf{E}' = \mathbf{U}'\,\mathbf{Diag}([\sigma'_1, \sigma'_2, \sigma'_3])\,\mathbf{V}'^{\top}.$$

The essential matrix ${}^{\mathbf{a}}\mathbf{E}$, which minimizes $\|\mathbf{E}' - {}^{\mathbf{a}}\mathbf{E}\|_F$ with respect to the Frobenius [1] matrix norm, is the one with σ'_3 set to zero as it is by definition the smallest singular value. Since an essential matrix is defined up to a constant factor, the singular values σ_1 and σ_2 may be set to 1. The final matrix reads as

$${}^{\mathbf{a}}\mathbf{E} = \mathbf{U}'\,\mathbf{Diag}([1, 1, 0])\,\mathbf{V}'^{\top}. \tag{3.12}$$

3.2.2 Refinement of the Relative Orientation

The solution presented in Sec. 3.2.1 solves for the entries of the essential matrix directly; each entry contributes linear to the cost function, i.e. the coplanarity

[1] $\|A\|_F = \sqrt{\sum_i \sum_j a_{ij}^2}$, where a_{ij} is the matrix entry at row i and column j.

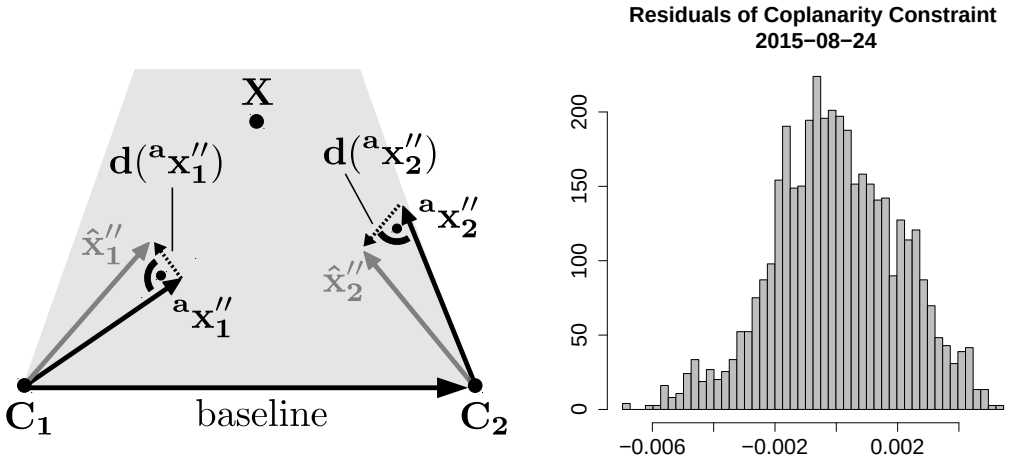


Figure 3.3: **Left**: Refining the relative orientation minimizes the deviations of corresponding projection rays, ${}^{a}\mathbf{x}_1''$ and ${}^{a}\mathbf{x}_2''$, from the coplanarity constraint. The corrections $\mathrm{d}({}^{a}\mathbf{x}_1'')$ and $\mathrm{d}({}^{a}\mathbf{x}_2'')$ quantify the deviations of the observations from the estimated true observations, $\hat{\mathbf{x}}_1''$ and $\hat{\mathbf{x}}_2''$. **Right**: Residuals of the coplanarity constraint after parameter estimation for 24 August 2015.

constraint. The relative orientation is well defined by rotation and baseline vector and will thus be optimized w.r.t. these entities. Since the coplanarity constraint is non-linear in the relative rotation of the camera system and the observed projection rays, a least-squares estimation is computed iteratively and requires a linearization of the cost function.

Given two observed projection rays ${}^{a}\mathbf{x}_1''$ and ${}^{a}\mathbf{x}_2''$ as well as an approximate relative rotation matrix ${}^{a}\mathbf{R}$ and baseline vector ${}^{\mathbf{a}}\mathbf{b}$, a linear Gauss-Helmert model is realized by an expansion of the coplanarity constraint:

$$\begin{aligned}\hat{\mathbf{x}}_1''^{\top}\,\hat{\mathbf{E}}\,\hat{\mathbf{x}}_2'' \;=\; \hat{\mathbf{x}}_1''^{\top}\,\hat{\mathbf{S}}_{\mathbf{b}}\,\hat{\mathbf{R}}\,\hat{\mathbf{x}}_2'' \;=\; \overbrace{{}^{a}\mathbf{x}_1''^{\top}\,{}^{a}\mathbf{S}_{\mathbf{b}}\,{}^{a}\mathbf{R}\,{}^{a}\mathbf{x}_2''}^{approximation} &+ \overbrace{\mathrm{d}({}^{a}\mathbf{x}_1''^{\top})\,{}^{a}\mathbf{S}_{\mathbf{b}}\,{}^{a}\mathbf{R}\,{}^{a}\mathbf{x}_2''}^{corrections} \\ &+ {}^{a}\mathbf{x}_1''^{\top}\,\mathrm{d}({}^{a}\mathbf{S}_{\mathbf{b}})\,{}^{a}\mathbf{R}\,{}^{a}\mathbf{x}_2'' \\ &+ {}^{a}\mathbf{x}_1''^{\top}\,{}^{a}\mathbf{S}_{\mathbf{b}}\,\mathrm{d}({}^{a}\mathbf{R})\,{}^{a}\mathbf{x}_2'' \\ &+ {}^{a}\mathbf{x}_1''^{\top}\,{}^{a}\mathbf{S}_{\mathbf{b}}\,{}^{a}\mathbf{R}\,\mathrm{d}({}^{a}\mathbf{x}_2'')\end{aligned} \quad . \tag{3.13}$$

Here, $\mathrm{d}({}^{a}\mathbf{x}_1'')$ and $\mathrm{d}({}^{a}\mathbf{x}_2'')$ are the corrections to the observed projection rays, while $\mathrm{d}({}^{a}\mathbf{S}_{\mathbf{b}})$ and $\mathrm{d}({}^{a}\mathbf{R})$ are the corrections to the relative orientation terms.

Corrections to unit vectors, such as the projection rays or the baseline, only have two degrees of freedom as they are restricted to the unit sphere. Following Förstner (2011), a suitable parametrization restricts such corrections to the

Date	Cameras	α	β	γ	Δu^r	Δv^r
2015-08-24	JOYCE	0.007°	0.015°	0.013°	0.046°	0.068°
2014-08-05	JOYCE	0.017°	0.011°	0.015°	0.055°	0.076°
2015-08-20	RWE	0.009°	0.019°	0.019°	0.062°	0.056°

Table 3.1: Estimated directional precision of relative rotation angles α, β and γ, and the baseline vector via its reduced coordinates Δu^r and Δv^r for selected days and stereo configurations (see Sec. 4.1.1).

local 2-D tangent space on the unit sphere, i.e. the null-space of the respective unit vector (Figure 3.3). A correction of the unit vector $\mathbf{x}''$ can then be written in form of a parameter update by $[\Delta u^r, \Delta v^r]^\top$, according to

$$\mathrm{d}(\mathbf{x}'') = \mathbf{null}(\mathbf{x}'') \begin{bmatrix} \Delta u^r \\ \Delta v^r \end{bmatrix}, \quad \text{where} \quad \mathbf{null}(\mathbf{x}'') = [\mathbf{u^r}, \mathbf{v^r}] = \begin{bmatrix} x_{u^r} & x_{v^r} \\ y_{u^r} & y_{v^r} \\ z_{u^r} & z_{v^r} \end{bmatrix},$$

followed by a subsequent spherical normalization. Using this parametrization makes sure that the corresponding covariance matrix related to the directional uncertainty of the projection ray is regular.

Each iteration includes an update of the relative orientation parameters, the three rotation angles and two reduced coordinates to correct the baseline vector, and an update of the projection rays via its four reduced coordinates (Förstner, 2011; Förstner and Wrobel, 2016, pp. 586-588).

3.3 Epipolar Image Rectification for Hemispheric Cameras

Epipolar image rectification transforms two stereo images of arbitrary oriented cameras in such a way that the projection of an epipolar plane is aligned to the image rows. Two corresponding image points in the rectified images then differ only by a horizontal shift within the same image row, which largely simplifies subsequent stereo analysis. In case of hemispheric cameras, however, epipolar planes are mapped to epipolar curves rather than epipolar lines as for a pinhole camera. Hence, a rectification based on a projective image transformation (e.g. Loop and Zhang, 1999) alone does not directly lead to a correspondence between epipolar planes and image rows. The rectification of the fisheye stereo images is done following Abraham and Förstner (2005), which allows to use the whole image content of hemispheric cameras for stereo analysis.

The image transformation leads to a *canonical stereo configuration* or a *stereo normal case*, in which the two cameras have equal intrinsic parameters, equal orientation and are separated only along the x-direction of the reference

camera. In terms of the relative orientation, this yields

$$\mathbf{R} = \mathbf{I}, \qquad \mathbf{b} = [b, 0, 0]^\top, \qquad \mathbf{K}_1 = \mathbf{K}_2, \tag{3.14}$$

where b denotes the distance between the cameras. Since the real cameras may have arbitrary spatial orientation and intrinsic parameters, the rectification defines for each real camera a virtual camera that has the desired properties outlined in 3.14.

The rectification transform and the virtual stereo setup are illustrated in (Figure 3.4): An object point $\mathbf{X_V}$ observed by two virtual stereo cameras can then be described by the angle β, which determines the respective epipolar plane, and the two angles ψ_1 and ψ_2 within the epipolar plane representing the two projection rays. Based on this geometry a rectification scheme can be constructed that covers the whole 3-D space by mapping the angle β to the image rows and the angles ψ_1 and ψ_2 to the columns of the respective image, which guarantees that corresponding image points have the same image row index:

$$\mathbf{x'_V} = \begin{bmatrix} \psi \\ \beta \end{bmatrix} = \begin{bmatrix} \text{atan2}(X_V, \sqrt{Y_V^2 + Z_V^2}) \\ \text{atan2}(Y_V, Z_V) \end{bmatrix} \quad \text{with} \quad \mathbf{X_V} = [X_V, Y_V, Z_V]^\top.$$

The mapping from a projection ray $\mathbf{x}''$ to its correspondence $\mathbf{x}''_V$ in the virtual reference system is done by a sequence of rotations. First, both cameras are aligned so that they have equal orientation (step 1 in Figure 3.4), which is done by correcting for the relative rotation of the second camera. In a second step, the rotation $\mathbf{R_V}$ aligns both reference systems so that the respective x-axes and the baseline vector are collinear (step 2 in Figure 3.4). The first column of $\mathbf{R_V}$ is the normalized baseline vector. The other two columns can be chosen freely as long as they are orthogonal and $\mathbf{R_V}$ is a valid rotation matrix. This implies that the epipolar rectification is determined up to a rotation about the baseline. It is convenient to chose a virtual y-axis that lies within the original x-y-plane of the reference camera, which then also determines the virtual z-axis. The rotation $\mathbf{R_V}$ is then given by

$$\mathbf{R_V} = [\mathbf{e_1}, \mathbf{e_2}, \mathbf{e_3}], \quad \text{where} \quad \begin{aligned} \mathbf{e_1} &= \mathbf{b} \cdot \|\mathbf{b}\|^{-1}, \\ \mathbf{e_2} &= [-y_b, x_b, 0]^\top \cdot \|[-y_b, x_b, 0]\|^{-1}, \\ \mathbf{e_3} &= \mathbf{e_1} \times \mathbf{e_2}, \end{aligned}$$

which also yields the mapping from $\mathbf{x}''$ to $\mathbf{x}''_V$:

$$\begin{aligned} \mathbf{x}''_{V,1} &= \mathbf{R_V}^\top \cdot \mathbf{I}\, \mathbf{x}''_1, \\ \mathbf{x}''_{V,2} &= \mathbf{R_V}^\top \cdot \mathbf{R}\, \mathbf{x}''_2. \end{aligned} \tag{3.15}$$

Based on the angles β and ψ related to $\mathbf{x}''_V$, the corresponding virtual image coordinates $[u_V, v_V]^\top$ are given by

$$\begin{aligned} u_V &= c\,\psi + u_{0,V}, \\ v_V &= c\,\beta + v_{0,V}, \end{aligned} \quad \text{with} \quad \begin{aligned} u_{0,V} &= c\,(\psi_{max} - \psi_{min})/2, \\ v_{0,V} &= c\,(\beta_{max} - \beta_{min})/2, \end{aligned} \tag{3.16}$$

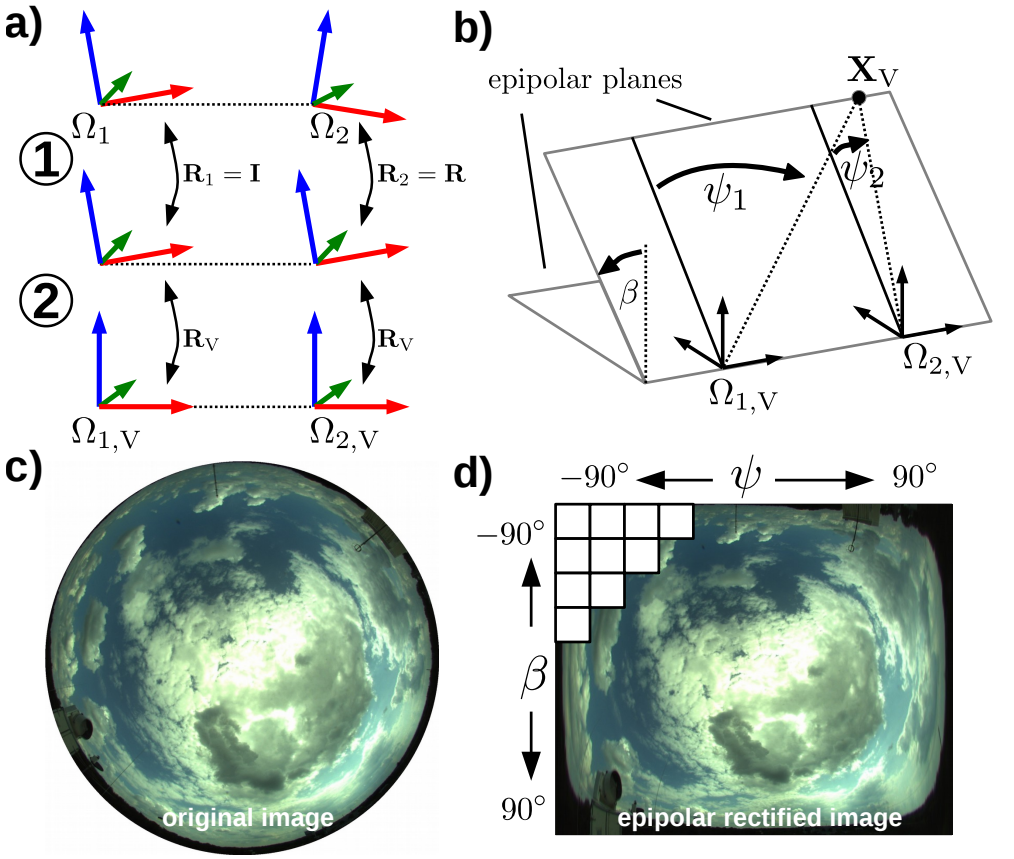


Figure 3.4: Epipolar rectification for hemispheric cameras: (a) The two-step rotational mapping between real and virtual cameras results in (b) a canonical camera setup of virtual cameras during rectification. A fisheye image (c) is then transformed so that the angles β and ψ correspond to (d) the rows and columns respectively.

where $[u_{0,V}, v_{0,V}]^\top$ denotes the coordinates of the virtual principal point in pixel coordinates, while c is the camera constant. The view field of the virtual cameras is defined by β_{min}/β_{max} and ψ_{min}/β_{max}, respectively.

The backward transformation from the virtual image point $\mathbf{x}_\mathrm{V}$ to the virtual projection ray $\mathbf{x}''_V$ is obtained from Eq. (3.16) as

$$\mathbf{x}''_V = \begin{bmatrix} x''_V \\ y''_V \\ z''_V \end{bmatrix} = \begin{bmatrix} \sin(\psi) \\ \cos(\psi)\sin(\beta) \\ \cos(\psi)\cos(\beta) \end{bmatrix}, \quad \text{where} \quad \begin{aligned} \psi &= (u_V - u_{0,V})/c, \\ \beta &= (v_V - v_{0,V})/c. \end{aligned} \tag{3.17}$$

Accordingly, the projection ray within the (real) camera reference system $\mathbf{\Omega_1}$ is given by

$$\mathbf{x}''_1 = \mathbf{R_V}\,\mathbf{x}''_{V,1}, \tag{3.18}$$

Note that the rectification produces strong distortions at the left and right margins of the images near the epipoles, which leads to a diminished precision

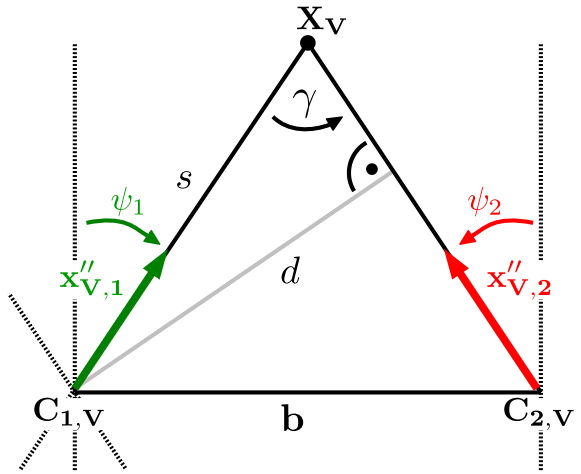


Figure 3.5: Triangulation of point $\mathbf{X_V}$ within the epipolar plane in a canonical stereo setup exploits the coplanarity of the two projection rays $\mathbf{x}''_{V,1}$ and $\mathbf{x}''_{V,2}$. The edge d creates two right triangles, which allows to derive the depth value s via ψ_1 and ψ_2.

of a stereo analysis in those regions and subsequently the reconstruction (Figure 3.8). Towards the image center, however, the distortion becomes smaller and vanishes for viewing directions perpendicular to the baseline ($\psi \to 0$).

3.4 Triangulation

The computation of the 3-D object point $\mathbf{X}$ from two corresponding projection rays $\mathbf{x}''_1$ and $\mathbf{x}''_2$ is done via triangulation (Figure 3.5). One solution makes use of the fact that in this study stereo analysis is based of epipolar image rectification, which guarantees that two corresponding projection rays are coplanar, i.e. lie within the same epipolar plane. Hence, $\mathbf{X_V}$ can be computed directly using the following relations derived from Figure 3.5:

$$\begin{aligned} s \cdot \cos(\psi_2) = d &= b \cdot \sin(\psi_1 - \psi_2), \\ \mathbf{X_V} &= s \cdot \mathbf{x}''_{V,1}. \end{aligned}$$

Including the absolute orientation of the reference camera, represented by the rotation matrix $\mathbf{R_{abs}}$, the object point $\mathbf{X}$ in world coordinates is then given by

$$\mathbf{X} = \mathbf{R_{abs}}\,\mathbf{R_V}\,\mathbf{X_V} \qquad \text{with} \qquad \mathbf{X_V} = b \cdot \left(\frac{\cos(\psi_2)}{\sin(\psi_1 - \psi_2)} \right) \mathbf{x}''_{V,1}. \tag{3.19}$$

If the two projection rays are not coplanar, one can obtain a linear solution via Direct Linear Transformation (DLT) as follows. Following (Förstner and Wrobel, 2016, p. 602), an unknown object point $\tilde{\mathbf{X}}$ in homogeneous form and

the corresponding projection rays $\mathbf{x}_1''$ and $\mathbf{x}_2''$ should satisfy

$$\begin{aligned} \lambda_1 \mathbf{x}_1'' &= \mathbf{P}_1 \, \tilde{\mathbf{X}}, \\ \lambda_2 \mathbf{x}_2'' &= \mathbf{P}_2 \, \tilde{\mathbf{X}}, \end{aligned} \qquad \text{for} \qquad \lambda_1, \lambda_2 \neq 0,$$

where $\mathbf{P}_1$ and $\mathbf{P}_2$ are the respective (calibrated) camera matrices with $\mathbf{K} = \mathbf{I}$ (compare Eq. (2.10)). Applying the vector product leads to the equation system

$$\begin{bmatrix} \mathbf{x}_1'' \times \mathbf{P}_1 \\ \mathbf{x}_2'' \times \mathbf{P}_2 \end{bmatrix} \tilde{\mathbf{X}} = 0 \quad \Leftrightarrow \quad \underbrace{\begin{bmatrix} \mathbf{S}(\mathbf{x}_1'') \, \mathbf{P}_1 \\ \mathbf{S}(\mathbf{x}_2'') \, \mathbf{P}_2 \end{bmatrix}}_{\mathbf{A}} \tilde{\mathbf{X}} = 0,$$

where $\mathbf{S}(\mathbf{x}'')$ denotes the skew symmetric matrix related to $\mathbf{x}''$ (see Eq. (3.3)). A Singular Value Decomposition then yields $\tilde{\mathbf{X}}$ as the last column of the matrix $\mathbf{V}$.

3.5 Stereo Analysis

3-D-stereo reconstruction via triangulation requires precise estimations of the projection rays $\mathbf{x}_1''$ and $\mathbf{x}_2''$, which in turn are determined by corresponding image points $\mathbf{x_1}$ and $\mathbf{x_2}$. A stereo matching algorithm automatically establishes such correspondences across two images by means of image pixels, patches or extracted geometric features.

The discrete nature of digital images implies that an object is only resolved at a limited spatial resolution. The projection by two cameras at different positions leads to variations in the image signal regarding intensity (surface reflectance, scattering) and geometric shape (projection of a 3-D object onto a 2-D image), which introduce uncertainties to the stereo matching and thus affect the directional precision of $\mathbf{x}_1''$ and $\mathbf{x}_2''$. In general, the precision can be expected highest if the object surface is approximately planar and aligned parallel to the camera sensor.

Individual image points (pixels) yield only a single color or intensity information, which is typically not sufficient to establish reliable correspondences. Stereo analysis often matches image patches or detectable geometric features, such as lines or regions, which contain a much higher amount of information and can be identified in both images. The matching of complex geometric entities (e.g. the silhouette of a cloud), however, can be challenging, because the object or parts of it may significantly change its appearance in the images due to occlusion or perspective change. A convenient approach is to use small image patches or interest points (Sec. 3.5.1), which can often be found in large numbers and matched with sufficient reliability.

Stereo matching relies on the visual appearance of an object in the images. Changes in illumination (whether by surface reflection, scattering or camera

settings), perspective distortions or specular reflections can lead to incorrect or flawed correspondences. For the stereo reconstruction of clouds, an obvious challenge in many situations is the lack of contrast, which leads to a missing or uncertain reconstruction in those areas. Thus, the base of a cumulus cloud is generally reconstructed with a higher geometric uncertainty than its well-structured convective top. Further sources of uncertainty are the discretization of the intensity signal, electronic noise and a limited computational precision.

A stereo algorithm determines a set of image point correspondences based on the similarity of the local image signal. Let $\mathbf{x}_{1,\mathrm{V}}$ and $\mathbf{x}_{2,\mathrm{V}}$ be two corresponding image points determined by a stereo algorithm from epipolar rectified images (see Sec. 3.3). Since the epipolar lines are aligned to the image rows, both image points differ only by their u-coordinate. The *disparity* is defined as the horizontal difference between the two image points in pixel coordinates

$$d = u_{1,V} - u_{2,V}.$$

A *dense* stereo algorithm seeks a valid disparity value for each pixel in the image, which results in a *disparity map* $D(u_{1,V}, v_{1,V})$ containing for each image point its corresponding disparity value.

According to Scharstein et al. (2001), most dense stereo matching algorithms follow the steps (a) matching cost computation, (b) matching cost aggregation, (c) disparity computation and (d) disparity refinement. Matching cost computation defines a cost function to quantify the similarity between individual pixels, whereas the cost aggregation step integrates the matching cost over a local support region or template window. The disparity computation step may implement further constraints regarding spatial smoothness, visibility or order to avoid mismatches or provide reasonable disparity estimates in low-contrast regions. In a last step, the obtained disparity values are refined to sub-pixel precision.

Depending on the degree an algorithm assesses the steps (b) and (c), it may be classified as a local or a global method. Local methods rely on the matching cost aggregation alone and implement a winner-takes-it-all strategy, which allows an efficient computation, but also makes it less robust against ambiguities resulting from low-contrast regions or repeating patterns. Global methods emphasize the disparity computation step and often minimize a global energy functional $E = E_{data} + E_{smooth}$, which combines data consistency (steps (a) and (b)) with a spatial smoothness constraint that demands similar disparities for neighboring image points. This leads to a propagation of disparity information across the image and a generally more robust and consistent reconstruction. Many global methods require epipolar rectified images in order to limit the high computational complexity resulting from the global disparity computation. Epipolar image rectification, however, can lead to image deformations and may affect the matching cost terms in those regions. Global optimization techniques used to compute a disparity map range from Markov Random Fields (Felzenszwalb and Huttenlocher, 2006) to graph-based methods such as

Min Cut/Max Flow (Boykov et al., 2001), Dynamic Programming (Birchfield and Tomasi, 1998) or the variational methods (Slesareva et al., 2005).

The following sections introduce the stereo matching and reconstruction techniques applied in this thesis.

3.5.1 Feature Matching

Feature-based stereo analysis relies on the detection and matching of well-defined geometric image features, such as lines (Bay et al., 2005), regions (Matas et al., 2004) or interest points (Tuytelaars and Mikolajczyk, 2007). The latter are also known as *keypoints*. This section focuses on the detection, localization and matching of interest points using the Scale Invariant Feature Transform (SIFT, Lowe, 2004), which is widely used for 3-D reconstruction, camera pose estimation or image registration. Like all interest point operators, SIFT-features mark well-defined image areas with low entropy (high information) that show strong image intensity gradients, such as corners, junctions or blobs.

The SIFT operator conducts *detection* and *description* of interest points. Detection relies on edge localization via convolution with the *Difference of Gaussians* (DoG) filter, which approximates the *Laplacian of Gaussians* (LoG) edge detector. The image is convoluted with a Gaussian filter (G) twice: once using a kernel size of σ_1 and another time with a larger kernel size σ_2. Subtracting the two filtered images yields the Difference of Gaussians (DoG) image:

$$\mathrm{DoG}(u, v, \sigma_1, \sigma_2) = \mathrm{G}(u, v, \sigma_2) - \mathrm{G}(u, v, \sigma_1) \quad \text{with} \quad \sigma_1 > \sigma_2.$$

Here, $\mathrm{G}(u, v, \sigma)$ denotes the Gaussian filter of kernel size σ applied at the image point $\mathbf{x} = [u, v]^\top$. The DoG filter is applied several times at each image point with increasing values for σ_1 and σ_2, resulting in a set of DoG-layers (Figure 3.6), also known as *scale space*. The detected local extrema in each DoG-layer are potential interest points, but may be declined due to a low absolute response strength. Unstable interest points that cannot be reliably located throughout scale space are avoided by taking into account a 3x3 neighborhood in the DoG-layer below and above the current layer, which must not contain an interest point with a smaller (larger) response if it represents a local minimum (maximum). The interest point localization is then refined to sub-pixel accuracy by fitting a three-dimensional quadratic function through the local 3x3x3 DoG-neighborhood (Figure 3.6).

Once the detection step is finished and a set of interest points is collected, each must be assigned a *descriptor* that can be used to identify the same interest point in other images. The SIFT feature descriptor consists of a 128 (or 256)-bin histogram containing the gradient magnitudes around the detected interest point. Based on the radial distance to the interest point, the gradients are weighted with a Gaussian kernel and are subsequently normalized to

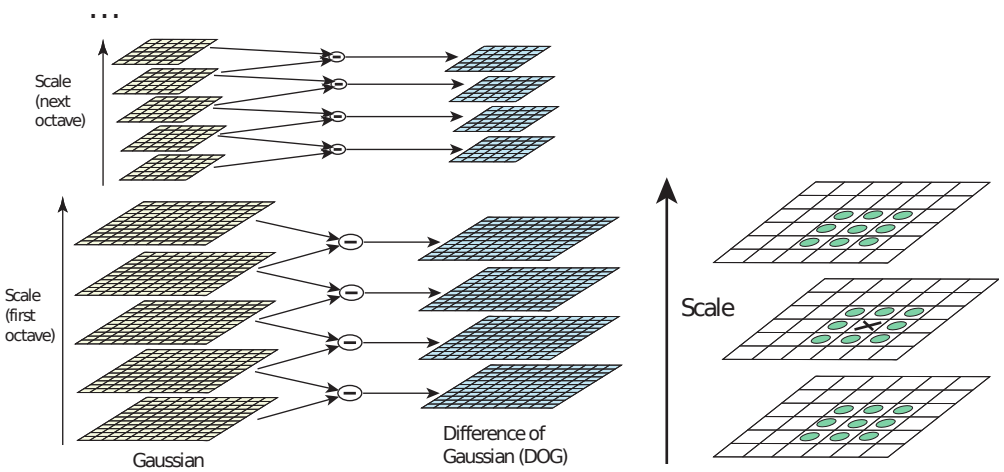


Figure 3.6: Left: Scale space analysis of an image via Gaussian Image Pyramid. **Right**: Interest point localization of the SIFT-operator via scale space extrema detection. (adapted from Lowe, 2004).

provide some tolerance to variations in illumination. The histogram is then aligned to the dominant gradient direction, which allows the to identify the interest point in case of rotated images. For more detailed information of the SIFT-descriptor, see Lowe (2004).

Once interest points and their descriptors are computed, they can be matched across the different images. Figure 3.2 shows a matching result used to determine the essential matrix via interest point correspondences. At this stage the coplanarity constraint (see Sec. 3.1) states that corresponding interest points are restricted to lie on the respective epipolar curve. This can be used to filter out interest point candidates that violate this constraint and thus improve the matching. Image sampling, measurement errors and image noise affect the localization of interest points, which makes it necessary to relax the coplanarity constraint and tolerate deviations to some degree, e.g. 1 pixel. Besides the coplanarity constraint, a maximum valid disparity range can reduce the number of invalid correspondences as it formulates as a maximum enclosing angle between two projection rays, e.g. 15°.

The interest points are assumed to offer sufficient information for a reliable matching. Corresponding interest points are identified by their highest matching score based on the gradient information contained in the descriptors $\mathbf{q_1}$ and $\mathbf{q_2}$ via a similarity measure, such as the sum of squared differences

$$f(\mathbf{q}_1, \mathbf{q}_2) = \sum_i \|\mathbf{q}_1(i) - \mathbf{q}_2(i)\|^2,$$

where i is the descriptor entry index. It is recommended to reject an interest point match if the matching score ratio between the best and the second best

candidate is too small:

$$\frac{f(\mathbf{q}_1, \mathbf{q}_2)}{f(\mathbf{q}_1, \mathbf{q}_3)} < \epsilon_f \qquad \text{with} \quad \epsilon_f \in [0.5, 0.8]$$

In a last step, the correspondences are validated by a consistency check, which requires the whole matching procedure to be repeated for the reversed image pair, i.e. interest points in image two are matched against interest points in image one. Thus, a correspondence between two interest points $\mathbf{x}_1$ and $\mathbf{x}_2$ is valid if $\mathbf{x}_1$ matches $\mathbf{x}_2$ and vice versa.

3.5.2 Semi Global Block Matching (SGBM)

The Semi-Global Block Matching (SGBM) is a modification of the technique introduced by Hirschmuller (2005), and is implemented as part of the open-source computer vision library OpenCV (Bradski, 2000). The technique has become very popular due to its convincing results, comparatively simple implementation and high efficiency. While computational complexity is generally not a crucial aspect in 3-D cloud reconstruction, for some applications, such as short-term surface solar radiation forecast or an operational use at observation sites, the economic aspect should not be neglected.

The algorithm computes a disparity map from a pair of epipolar rectified stereo images (Sec. 3.3) by minimizing a global energy functional. Based on a chosen range of valid disparities, the first step is to compute the matching cost for each pixel in the reference image for all possible disparity values. The algorithm performs a template-based matching cost computation, which combines the sampling-insensitive similarity measure of Birchfield and Tomasi (1998) and an image preprocessing step to compensate small errors in the relative orientation of the camera system (Hirschmuller and Gehrig, 2009). For the matching cost aggregation the sum of absolute differences (SAD) is used.

The disparity computation is done in the second step by imposing a spatial smoothness constraint, which penalizes disparity differences between neighboring pixels along 8 different directions. The penalty is small, P_1, in case the difference is below a certain threshold (e.g. 1 pixel), and higher, P_2, if the difference is larger. The recursive formulation of the updated matching cost $L_r(\mathbf{x}_\mathrm{V}, d)$ for an image point $\mathbf{x}_\mathrm{V}$ in the reference image, a disparity value $d \in [d_{min}, d_{max}]$ and the path vector $\mathbf{r}$, is given by

$$\begin{aligned} L_r(\mathbf{x}_\mathrm{V}, d) \;=\; C(\mathbf{x}_\mathrm{V}, d) + \min\{ & L_r(\mathbf{x}_\mathrm{V} - \mathbf{r}, d), \\ & L_r(\mathbf{x}_\mathrm{V} - \mathbf{r}, d-1) + P_1, \\ & L_r(\mathbf{x}_\mathrm{V} - \mathbf{r}, d+1) + P_1, \\ & L_r(\mathbf{x}_\mathrm{V} - \mathbf{r}, k) + P_2 \} \end{aligned} \tag{3.20}$$

with

$$\mathbf{r} \in \{-1, 0, +1\}^2 \setminus [0, 0]^\top, \qquad k \neq \{d-1, d, d+1\}, \qquad P_1 < P_2.$$

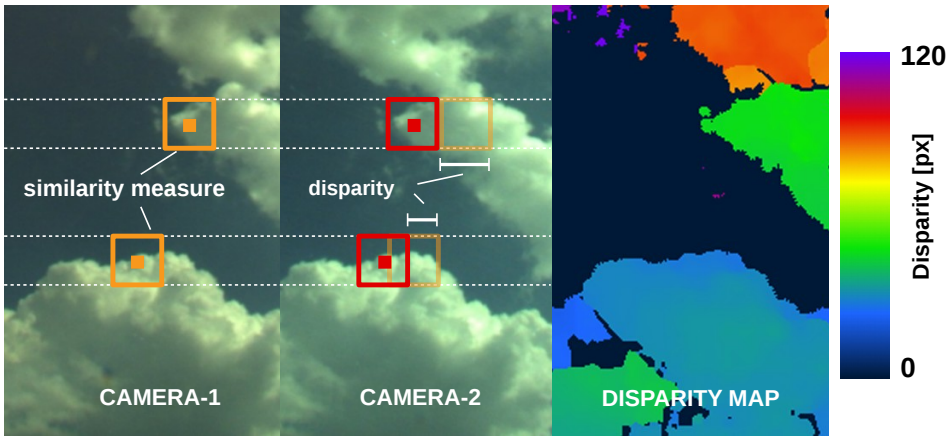


Figure 3.7: Dense stereo matching with SGBM using epipolar rectified stereo images. **Left/Center**: Correspondence information is computed for all pixels if possible via template-based similarity measure. **Right**: Each disparity value $d > 0$ then allows to reconstruct one 3-D point, resulting in a dense 3-D point cloud.

Summed up over all directions $\mathbf{r}$, the total energy for $\mathbf{x}_\mathrm{V}$ and d is

$$S(\mathbf{x}_\mathrm{V}, d) = \sum_r L_r(\mathbf{x}_\mathrm{V}, d).$$

and the disparity minimizing the global energy for $\mathbf{x}_\mathrm{V}$ is

$$D(\mathbf{x}_\mathrm{V}) = \arg\min_d \{S(\mathbf{x}_\mathrm{V}, d)\}$$

In a last step, the disparities are refined to sub-pixel accuracy by fitting a parabola at the disparity minimum, i.e. $S(\mathbf{x}_\mathrm{V}, D(\mathbf{x}_\mathrm{V}) - 1), S(\mathbf{x}_\mathrm{V}, D(\mathbf{x}_\mathrm{V}))$ and $S(\mathbf{x}_\mathrm{V}, D(\mathbf{x}_\mathrm{V}) + 1)$. The suitability of the fitted function may vary with the underlying energy functional.

The disparity map is computed twice, once with image 1 as reference and vice versa, to make sure the computed disparities are mutually consistent and differ at most by a given threshold ϵ_d (e.g. 1 pixel):

$$D(\mathbf{x}_\mathrm{V}) = \begin{cases} d & \text{if} \quad |D_{1\to2}(\mathbf{x}_{1,\mathrm{V}}) - D_{2\to1}(\mathbf{x}_{2,\mathrm{V}})| < \epsilon_d \\ \text{invalid} & \text{otherwise} \end{cases}$$

Recall that $d > 0$.

3.5.3 Patch-based Multi-View Stereo (PMVS)

Patch-Based Multi-View Stereo (Furukawa and Ponce, 2010) models an object as a collection of oriented and textured (radiation reflecting) rectangular 3-D-patches in space. A patch that is close to its corresponding surface element of the real object, including its orientation, is assumed to produce a

radiometric and geometrically consistent projection in each observing camera. The algorithm uses both the SIFT and the Harris feature detector to collect a set of initial interest point correspondences and follows an expand-and-filter strategy in order to achieve a denser surface reconstruction. Such an approach is flexible regarding the number of employed cameras and results in a high accuracy, since each patch is optimized w.r.t. its location and orientation to achieve the best matching score across the images. According to the ETH3D[2] multi-view benchmark, PMVS ranks high at accuracy, but relatively low at geometric completeness.

Consider a set of spatially distributed cameras observing the same object from different views. The image I_k of camera k is subdivided by a regular grid, e.g. with a cell size of 32×32 pixel. After a feature detection by the Difference-of-Gaussians (Sec. 3.5.1) and Harris detector (Harris and Stephens, 1988), each such grid cell keeps only a certain number of those features with highest gradient magnitude, which are then matched across the individual images by implementing a maximum deviation from the coplanarity constraint and using normalized cross-correlation as matching cost function. The resulting set of matched interest points is then triangulated to yield the respective centers of the initial set P of patches. Each patch is associated with a reference image, which provides the corresponding texture and initial orientation. After the matching phase, the collected patches are refined w.r.t. patch location and orientation to yield the highest average cross-correlation, $\overline{N}(p)$, across the images:

$$\overline{N}(p) = \frac{1}{|T(p)| - 1} \sum_{I \in T(p), I \neq R(p)} NCC(p, R(p), I), \tag{3.21}$$

where $T(p)$ is the set of images where p is visible and $R(p)$ is the reference image of p.

A subsequent iterative patch expansion and outlier filtering then creates a consistent and robust quasi-dense 3-D-reconstruction. The patch expansion step is similar to the smoothness constraint, which is usually implemented in global stereo matching algorithms and assumes similar disparities for neighboring pixels. New patches are only created if an image grid cell does not have its maximum number of registered (projected) patches and existing ones in the vicinity have a similar patch orientation. Outliers are removed during the filter stage if patches in front or behind the correct surface patch are detected via low correlation and visibility values compared to neighboring patches.

The current implementation works for central-perspective cameras only and cannot be applied to the whole image content of a fisheye image simultaneously. For practical purposes, the application of PMVS is limited to a subregion of the hemispheric stereo images, which is projected onto the local tangent image plane to obey the central-perspective projection.

[2] https://www.eth3d.net

3.5.4 Discussion

In this thesis, stereo analysis relies primarily on SGBM, as it provides a geometrically more complete reconstruction than PMVS; the resulting dense 3-D point cloud is well suited for a further analysis like segmentation of individual clouds. Such stereo algorithms establish image point correspondences, while implicitly assuming a horizontal shift of the image signal. This assumption is approximately valid for planar surface elements with an orientation similar to the viewing direction. In case of slanted or highly curved surfaces, however, the fixed geometry of the template window is often not suited to adequately resolve the depth (disparity) range within the template, which then results in mismatches or uncertainties in the reconstruction. In this application, the base of distant clouds yields such a slanted surface relative to the viewing direction.

PMVS combines stereo matching and 3-D reconstruction by optimizing the object geometry directly in space via oriented 3-D patches. This approach has the potential to achieve a higher accuracy and precision for cloud top areas, which are often not aligned towards the cameras. In this application, the full potential of PMVS is not exploited because large distances between the stereo cameras will generally make parts of a cloud visible only to one stereo camera, i.e. two camera devices. It can be expected, however, that the method produces slightly more accurate results for cloud top areas, which are not aligned towards the cameras.

3.6 Geometric Precision of Hemispheric Stereo

The theoretical precision is derived for a pair of hemispheric cameras w.r.t. of geometric distance and cloud base height, and is illustrated for a hypothetical cloud layer at 1500 m height. The geometric precision of a fixed stereo camera system depends primarily on the precision of the disparity obtained from the stereo matching. While the relative orientation and the intrinsic parameters influence the overall geometric accuracy of the reconstruction, they can be assumed constant; a flawed relative orientation would lead primarily to a persistent distortion of the 3-D reconstruction rather than random disturbances.

The quality of the computed disparity map depends primarily on contrast and structure in the respective image area: Cloud edges provide high contrast and thus a high signal-to-noise-ratio (SNR); a noisy image signal then has only small effects on the recovered object geometry. In image areas with low contrast, however, image noise has a larger influence on the matching result, leading to an uncertainty in the computed disparity estimate. Obviously, the size of the template window - or the patch size of interest points - also has a large influence on the uncertainty of the disparity map. A larger window allows to use a higher amount of information to compute the matching cost and thus increases the robustness to image noise. Following Förstner (1993), a good approximation to the precision of template matching is given by the

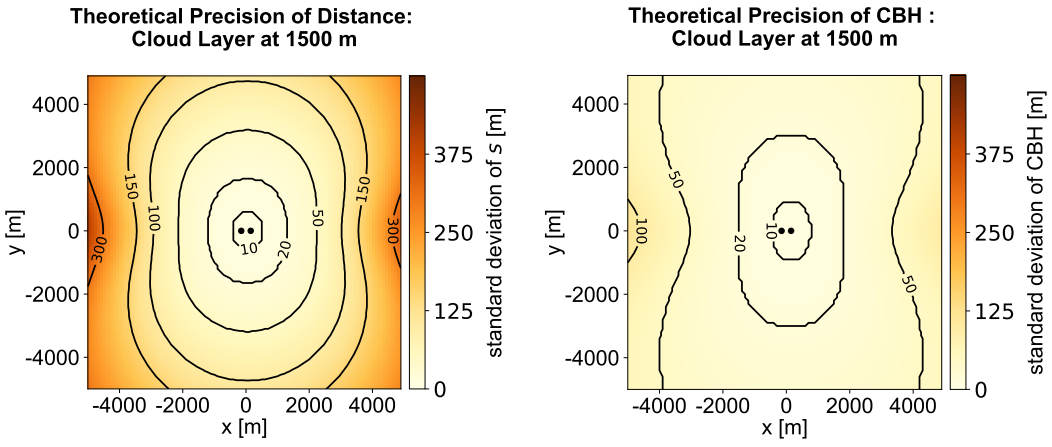


Figure 3.8: Spatial distribution of the theoretical precision for a stereo setup of hemispheric sky imagers within an area of 10 km × 10 km in meters, assuming a baseline of 300 m. Shown is the uncertainty of (a) the geometric distance and (b) cloud base height (CBH) for a hypothetical cloud layer at 1500 m height. The cameras are visualized as two black dots in the center. The directional uncertainty of the reference projection ray $\mathbf{x}''_{V,1}$ and the uncertainty of the angle γ is set to 0.07°.

covariance matrix $\boldsymbol{\Sigma}_{\mathbf{dd}}$

$$\boldsymbol{\Sigma}_{\mathbf{dd}} = \sigma_n^2 \, \mathbf{T}^{-1}, \qquad \text{where} \qquad \mathbf{T} = \begin{bmatrix} \sum_{i\in\Omega} f_x^2 & \sum_{i\in\Omega} f_x \, f_y \\ \sum_{i\in\Omega} f_x \, f_y & \sum_{i\in\Omega} f_y^2 \end{bmatrix}.$$

The matrix $\mathbf{T}$ is often called the *structure tensor* and contains information about the local gradient structure within the template domain $\boldsymbol{\Omega}$, as f_x and f_y denote the horizontal and vertical derivatives of the image signal. Higher gradients allow for a more precise matching. The image noise variance σ_n^2 can be obtained a priori from the sensor characteristics and the local image intensity, or may be derived from the difference of the matched templates.

The precision of a stereo camera system is derived via propagation of uncertainty using a first-order approximation of the triangulation step (Sec. 3.4). For the sake of simplicity, the analysis assumes a canonical stereo camera setup, i.e. $\mathbf{R_1} = \mathbf{R_2} = \mathbf{I}$ and $\mathbf{b} = [b, 0, 0]^\top$, where b denotes the distance between the cameras. It is assumed that the projection rays are subject to a directional uncertainty within the epipolar plane only.

Given a pair of projection rays $\mathbf{x}''_{V,1}$ and $\mathbf{x}''_{V,2}$, and corresponding angles ψ_1 and ψ_2 (Figure 3.5). One way to model the random perturbation is to interpret ψ_1 and ψ_2 as two independent random variables. This is justified in case the corresponding image points are localized independently. Another possibility is

that the angles ψ_1 and γ are both independent random variables, which better represents the process of template matching, as done by many dense stereo algorithms like SGBM, and generally leads to a smaller geometric uncertainty. This approach is pursued in the following analysis.

According to Eq. (3.19), a 2-D object point, ${}^{\text{2D}}\mathbf{X_V}$, within an epipolar plane can be rewritten in terms of the angles ψ_1 and γ as

$${}^{\text{2D}}\mathbf{X_V} = \underbrace{b\left(\frac{\cos(\psi_1 - \gamma)}{\sin(\gamma)}\right)}_{s} \begin{bmatrix}\sin(\psi_1)\\ \cos(\psi_1).\end{bmatrix}$$

The variance of the geometric distance s along the reference ray is then given as

$$\sigma_s^2 = \mathbf{J_s}\,\mathbf{Diag}([\sigma_{\psi_1}^2, \sigma_\gamma^2])\,\mathbf{J_s}^\top,$$

where the Jacobian

$$\mathbf{J_s} = \begin{bmatrix}\frac{\partial s}{\partial \psi_1} & \frac{\partial s}{\partial \gamma}\end{bmatrix}$$

is best obtained via numerical differentiation due to a bad numeric conditioning of the analytic solution (Schneider et al., 2016). The variance of cloud base height (CBH) depends directly on the zenith angle θ, which yields

$$\sigma_{\text{cbh}}^2 = \cos^2(\theta)\,\sigma_s^2.$$

Based on these formulations, one can investigate the spatial variation of the theoretical precision of distance (Figure 3.8a) and cloud base height (Figure 3.8b) for a pair of sky imagers. The virtual setup includes a hypothetical cloud layer at 1500 m height, which is observed by two sky imagers in a canonical stereo configuration and a baseline length of 300 m. The directional uncertainty of ψ_1 and γ is set to $\sigma_{\psi_1} = \sigma_\gamma = 0.07°$. The increasing uncertainty along the baseline is clearly recognizable, as is also the elliptic shape of the spatial distribution of the geometric error. This is particularly important for low clouds like shallow cumulus, which often have a cloud base height between 1000 m and 1500 m. In case of a mid- to high-altitude cloud, however, the error distribution becomes increasingly isotropic.

A single camera pair thus cannot monitor clouds with the same quality in all directions; a morphological analysis should focus on areas with lowest uncertainty, i.e. perpendicular to the baseline. Alternatively, a stereo camera pair could be extended to a 3-camera setup, for example aligned in a right-angled triangle. The additional camera would allow to reduce the geometric uncertainty and to improve stereo matching and triangulation (trifocal geometry). The uncertainty of cloud base height, i.e. the vertical component of the geometric uncertainty, depends also on the zenith angle of the respective projection ray (Figure 3.8b) and shows a slower increase with distance to the cameras. Hence, cloud height estimations may still be reliable even at larger distances to the cameras.

3.7 Effect of Asynchronous Recording

3-D stereo reconstruction of dynamic objects is sensitive to asynchronously recorded images. Given a fixed camera setup, a time difference between the individual recordings can lead to errors during stereo analysis that result in distortions or missing data in the 3-D reconstruction. The quality of the effect, however, varies with the motion direction of the recorded object relative to the camera setup.

If the motion direction is approximately parallel to the baseline, a small time shift between both recordings results in a positive or negative disparity bias (depending on the sign of motion) caused by the additional cloud motion between the recordings. Without additional information from other observations, it is impossible to judge whether the computed disparity map results from a lower (or higher) cloud layer base height or is caused by asynchronously recording cameras.

The bias of the 3-D model is directly proportional to the ratio of the translation between both recordings and the baseline length:

$$\mathrm{d}z_{cbh} = \frac{\Delta t \cdot v}{b}\, z_{cbh}.$$

In case the layers motion direction is perpendicular to the baseline, the resulting projection rays do not lie within the same epipolar plane and the triangulated point would lie between the two epipolar planes at a similar height. The effect on a disparity map, computed via epipolar rectified images, is an increased variability of the 3-D reconstruction without showing a clear vertical bias.

4 Evaluation of 3-D Cloud Reconstructions

This chapter evaluates the 3-D cloud geometries produced by the dense stereo reconstruction, as described in Chapter 3, with observational data from a micro-pulse lidar-ceilometer and a 35-GHz cloud radar situated at the Jülich Observatory for Cloud Evolution (JOYCE; Löhnert et al., 2015) within the Forschungszentrum Jülich GmBH.

Sec. 4.1 introduces the experimental setup, including a technical description of the instruments and the spatial configurations of the stereo camera system. The accuracy of the method is then evaluated for cases of cloud layers and cumulus clouds. The planar geometry of mid-altitude cloud layers is used to investigate the general accuracy and reliability of the reconstruction - especially at off-zenith regions (Sec. 4.2). Reconstructed morphologies of cumulus clouds are evaluated in Sec. 4.3 and compared to the 2-D reflectivity signal obtained from a cloud radar scan in Sec. 4.4. The efficient backscatter of clouds at wavelengths in the visible spectrum allows to confront the reconstructions with time-series of cloud base height retrievals from a lidar-ceilometer and investigate the accuracy of the system in various atmospheric situations (Sec. 4.5).

4.1 Experimental Setup

4.1.1 Stereo Camera Setup at Jülich

When designing a network of stereo cameras for monitoring the local cloud field and cloud morphology, choosing a spatial camera setup with a rather short baseline (e.g. 300 to 500 m) compared to previous studies (e.g. Seiz, 2003; Öktem et al., 2014), can be favorable. While a shorter baseline yields larger geometric uncertainties with increasing distance to the cameras, the small size and spacing of cumulus clouds makes a large baseline (e.g. 1 km) inappropriate because the amount of mutually visible cloud areas diminishes rapidly, impeding a dense stereo reconstruction.

The experimental camera setup in this study included up to two pairs of stereo cameras, i.e. four individual sky imagers, separated by several kilometers with the aim to reconstruct clouds from different perspectives. The sky imagers were situated within and around the Forschungszentrum Jülich GmBH (FZJ) in the state of North-Rhine-Westphalia, Germany. The FZJ is

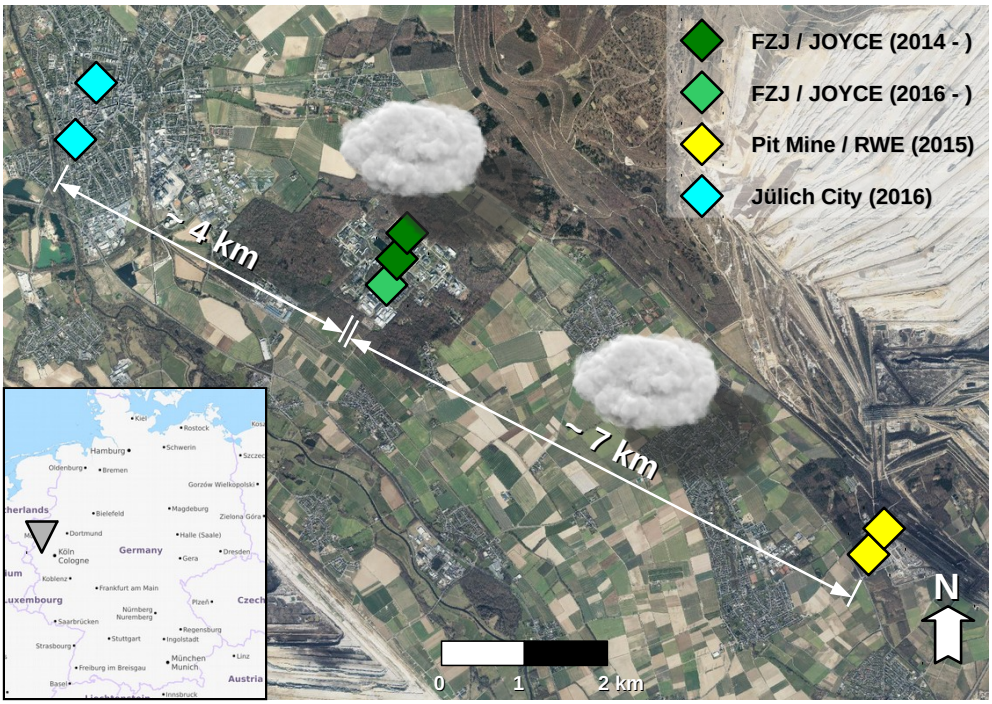


Figure 4.1: The experimental stereo camera setup consists of four hemispheric sky cameras grouped in two pairs. While two cameras were always situated within the the Forschungszentrum Jülich GmBH near the Jülich Laboratory for Cloud Evolution (JOYCE), the other camera pair was installed for a limited time period near the open pit mine (RWE) in 2015 and within the city center of Jülich since 2016 (JUL). Large map: *Land NRW (2017) Datenlizenz Deutschland - Namensnennung -Version 2.0 (www.govdata.de/dl-de/by-2-0).* Small map: ©*OpenStreetMap-Mitwirkende - www.openstreetmap.org/copyright.*

located between the cities Aachen in the West, Cologne in the East and 4 kilometers from the city of Jülich in the Northwest. While the natural land surface is plain, the local vicinity has a heterogeneous character with adjacent open pit mines to the West and East, an artificial hill (Sophienhöhe), a coal power plant *Weisweiler* and varying land use of forests and acres. Figure 4.1 gives an overview of the local geography and the spatial setup of the cameras. Three different stereo camera configurations were tested during the experimental phase, each configuration consisting of two cameras. In the following, these stereo camera setups are abbreviated as JOYCE, RWE, JUL. The stereo configuration JOYCE within the research center was operational from May 2014 to September 2015. The reference camera was situated directly at the observation site; the second camera was installed on the roof of a building approximately 300 m away. In October 2015, the second camera was re-located to another building, which increased the baseline to approximately 700 m. In 2015, two additional cameras - stereo configuration RWE - were installed

within the administrative area of the pit mine *Hambach* - property of the RWE group - located approximately 7 km distant from JOYCE, and was operational from June to early September 2015. The additional camera setup allowed to reconstruct a geometrically more complete boundary of passing clouds from two perspectives. Since a regular internet connection was not available, the cameras were controlled via a mobile connection. While placement options for the cameras were limited, the area of highest reconstruction precision could be aligned towards JOYCE. A baseline of 300 m was sufficient for an analysis of passing clouds between JOYCE and RWE, but resulted also in a rather low geometric precision for clouds very close to JOYCE. A third camera setup, JUL, which was located in the city center of Jülich, was tested in 2016 and included the cameras of the former RWE camera pair. In this configuration, both stereo configurations, JOYCE and JUL, had a baseline of more than 700 m. The shorter distance between JOYCE and JUL of about 4 km principally allowed to compare convective top reconstructions of passing clouds to the corresponding reflectivity profile from the cloud radar at JOYCE, but due to technical problems the scanning capability of the radar was not available during most of 2016. During this period, only the vertical profile scan (soda straw) was operational, which provides only a 1-D reflectivity profile at zenith.

The baseline of each stereo camera setup was aligned in Southwest - Northeast direction with the highest geometric precision in a perpendicular direction (Northwest - Southeast), i.e. along the line connecting the stereo camera locations (JOYCE - JUL - RWE). The baseline of the camera setup in the research center is to a large degree determined by the limited localization options, since few buildings are high enough to offer an open view that is unobstructed by trees or buildings, or have power-supply and network access without violating security standards. Based on the camera setup JOYCE, any further suitable location should generally satisfy a sufficiently large distance to JOYCE (e.g. 4 to 7 km), which allows to reconstruct clouds from different view perspectives. The additional camera pair should be oriented towards JOYCE to create overlapping areas of highest precision and generate synergy between the cameras. During the experiments, the maximum precision areas were oriented Northwest to Southeast and were suited to reconstruct passing clouds from two sides, moving in Southwest - Northeast direction. Clouds moving in Northwest - Southeast direction can be tracked for a long time and permit observation of cloud evolution and life-cycle.

4.1.2 Instruments for Validation

The JOYCE observation site hosts many instruments aimed at studying the evolution of boundary layer clouds, radiation closure and precipitation. Suitable information about the geometric aspects of passing clouds is provided by a lidar-celometer and a scanning cloud radar, which are described in more detail in the following sections.

4.1.2.1 Cloud Radar JOYRAD-35

The JOYRAD-35 (Figure 4.2a) is a polarimetric Doppler radar that measures the reflectivity at 35.5 GHz. Its capacity to emit and receive vertically and horizontally polarized radiation including their Doppler shift allows to analyze cloud macro-physical properties, such as cloud vertical structure and thickness, and retrieve micro-physical properties, such as cloud particle type, phase and velocity. Its sensing capability relies on a pencil beam with a measuring range between 150 m and 15 km, a range resolution of 30 m and a beam width of 0.6°, which corresponds to 52 m at 5 km distance.

The capability of the radar to sense clouds depends on its sensitivity to detect the backscattered radiation by a distribution of cloud droplets located at a distance d. For typical droplet sizes of 6 to 40 microns in diameter, Rayleigh-scattering theory then states that the received power P_r of the backscattered signal is

$$P_r \propto \frac{\pi^5}{\lambda^4} \left| \frac{m^2 - 1}{m^2 + 2} \right|^2 \frac{Z}{d^2},$$

where λ the wavelength of the radar beam and m the complex refractive index of water (or ice). The proportionality of P_r to λ^{-4} indicates that at a smaller wavelength, the amount of backscattered radiation increases, but also the beam extinction increases. The *reflectivity factor* Z is a property of the target, which, in this case, refers to a population of cloud droplets, and depends on the number density distribution $n(D)$ of cloud droplets. It is defined as

$$Z = \int_0^\infty n(D) D^6 \, \mathrm{d}D,$$

where D is the diameter of the droplet or particle. Due to the large range of possible reflectivity values, Z is usually given in logarithmic scale:

$$Z[dBZ] = 10 \log_{10}(Z).$$

Accordingly, a single large cloud droplet can have the same reflectivity factor as millions of smaller droplets. The emitted signal by the radar and the backscattered signal by the cloud droplets both decrease with distance $\propto 1/d^2$, resulting in a combined decrease $\propto 1/d^4$. The volume of the beam cell, however, increases increases $\propto d^2$, leading to a net decrease of the signal $\propto 1/d^2$. The radar is thus often not sensitive enough to detect distant parts of a cloud that contain very small droplets. These conditions are usually met at the edges of a cloud and near the cloud base, where entraining sub-cloud air just produces a narrow and small-size droplet spectrum. This effect is more significant for continental than marine cumulus. Typical cloud droplet sizes of marine cumulus reach sizes of 10 to 40 μm in diameter. Continental cumulus consist of smaller droplets due to the higher availability of cloud condensation nuclei

(a) **(b)**

Figure 4.2: Instruments for validation of 3-D stereo reconstructions. **Left**: Polarimetric Doppler-radar JOYRAD-35 produced by METEK. **Right**: Lidar-ceilometer CT25K by Vaisala. (Images adapted from METEK instrument data sheet (`http://metek.de/`) (left) and `http://gop.meteo.uni-koeln.de/ag_crewell` (right))

(CCN) and reach only sizes of 5 to 20 μm in diameter (Pruppacher and Klett, 1996, p. 18). At the cloud base, where droplets just reached their activation radii, the diameter can be even smaller.

The radar operates either in a 1-D vertical scan (*Soda Straw*) or 2-D cross-wind range-height indicator scan mode (CW-RHI), which is scheduled for the first 15 minutes of each hour and is aligned perpendicular to the dominant wind direction as derived from the nearby wind-lidar. The scan aims at sensing the 3-D cloud morphology of passing clouds by a sequence of 2-D reflectivity profiles (Kollias et al., 2014), and is used in this thesis to evaluate the stereo reconstructions.

An additional scanning sequence was operational irregularly in 2014 and 2015, which was designed to observe clouds within the regions of maximum geometric precision of the stereo camera system rather than in cross-wind direction as in CW-RHI mode. Since only a few scans per hour could be conducted, the yield of those scanning modes was not relevant for the results of this thesis.

4.1.2.2 Lidar-Ceilometer CT25K

The CT25K lidar-ceilometer (Figure 4.2b) emits laser pulses at a wavelength of 905 nm (near infra-red) and measures the backscattered signal by aerosols and cloud droplets. To increase the signal-to-noise-ratio, each pulse has a length of 100 ns, which allows to average 65.536 backscatter profiles over a period of 11.7 seconds. The laser beam has an opening angle of 0.037°, which corresponds to a diameter of 2.25 m at 3000 m height. The observation range reaches from 30 m up to 7500 m with a range resolution of about 30 m. Output is produced every 15 seconds, including data processing and transfer. An analysis of the height-resolved backscatter profile then allows to estimate cloud base height (CBH), cloud vertical dimension for optically thin clouds and boundary layer height.

Following Weitkamp (2005), the received power $P(r, \lambda)$ of the backscattered light at distance R and wavelength λ is given as

$$P(R, \lambda) = P_0 \frac{c\tau}{2} A \eta \frac{O(R)}{R^2} \beta(R, \lambda) \exp\left[-2 \int_0^R \alpha(r, \lambda)\, \mathrm{d}r\right], \qquad (4.1)$$

where P_0 is the power and τ the duration of a single emitted laser pulse, c the speed of light, A the area of the receiver optics, η is the device-specific optical efficiency and $O(r)$ the overlap ratio between the solid angle of the laser beam and the telescope area A. The exponential term determines the transmission of light according to the law of *Beer–Lambert–Bouguer*. The backscatter coefficient β, which specifies the amount of radiation that is scattered in backward direction, depends to the droplet size distribution within a volume and the laser light wavelength λ.

A retrieval algorithm analyzes the backscatter signal and detects sharp peaks that indicate a change in particle or droplet concentration, which reveals the edge of a cloud and thus CBH. Depending on the algorithm and the specifications of the lidar, the retrieved CBH can vary up to 100 m (Martucci et al., 2010).

In order to reduce backscatter signal induced by falling rain drops, the instrument is tilted by 2.5° (before 2015 the tilt was 7°), which permits estimation of CBH under rainy conditions. This strategy is, however, only effective for light rain, whereas strong rain events produce a backscatter which is dominated by the integrative effect of the rain drops. One has to consider this tilt as well as the absolute orientation of the lidar when comparing CBH observations to those derived by the stereo camera system, since the instruments might not point to exactly the same spot in the sky. This uncertainty, however, should not severely affect a comparison; larger disagreements can be expected for shallow cumulus clouds due to a higher small-scale variability. The exact orientation of the lidar beam was not estimated, but could be determined, for example, from the cross-correlation of time-height profiles of CBH from both the lidar-ceilometer and the stereo reconstruction.

4.1.3 Evaluation Methodology

A convenient way to validate the stereo reconstructions includes a comparison of zenith CBH from a lidar-ceilometer to the corresponding height values extracted from the reconstructions, as done for example in Seiz et al. (2002), Öktem et al. (2014) and Allmen and Kegelmeyer (1996). In this chapter, the reconstructions are also compared to the reflectivity profiles obtained from a scanning cloud radar. The methodology for a comparison is described in the following.

Due to the central-projective geometry of the cameras one should only compare observations, which are approximately collinear to the viewing direction of the reference camera in the respective stereo setup. A comparison of two corresponding object points, related to the reflectivity profile and the reconstruction respectively, is thus only valid if both are projectively equivalent, i.e. project approximately to the same image point. The 3-D geometric difference is expected to depend primarily on the stereo analysis (parallax/disparity), which may cause an uncertainty in the estimated distance along the respective projection ray (compare Sec. 3.6).

The (relative) *geometric difference* ϵ_{proj} (ϵ_{rel}) between a 3-D object point $\mathbf{X}$ from the stereo reconstruction and a reference point $\mathbf{X_{ref}}$, e.g. derived from the lidar-ceilometer or the cloud radar, is defined as:

$$\begin{aligned} \epsilon_{proj} &= \|\mathbf{X}\| - \|\mathbf{X_{ref}}\|, \\ \epsilon_{rel} &= \frac{\|\mathbf{X}\| - \|\mathbf{X_{ref}}\|}{\|\mathbf{X_{ref}}\|} \qquad \text{with} \qquad \angle(\mathbf{X}, \mathbf{X_{ref}}) < \epsilon_\gamma. \end{aligned} \tag{4.2}$$

The angle ϵ_γ defines the maximum valid angle between the projection ray of $\mathbf{X}$ and $\mathbf{X_{ref}}$. In the experiments, ϵ_γ is set to a value of 0.1°, which corresponds approximately to an error of one pixel in the image.

To obtain a meaningful error statistic and remove the influence of gross outliers, a comparison is done after the removal of values outside the 1.5-fold interquartile range above the upper and below the lower quartile of the dataset.

4.1.3.1 Evaluation via Zenith-CBH Time-Series of Lidar-Ceilometer

CBH is derived from the reconstruction by averaging the height of all points within a vertically aligned cone of 5° diameter. Such a large sampling area, compared to the beam diameter of the lidar, increases the chance of a valid CBH estimation, since the stereo analysis might not produce a valid disparity for all cloudy pixels in the reference image. Averaging over a range of CBH values, however, can also lead to disagreements with the corresponding value from the lidar-ceilometer. The effect is largest for clouds with a heterogeneous cloud base morphology or at the edges of a cloud where different parts of the cloud become visible. Also, the lidar-ceilometer was tilted by 7° (later 2.5°) and the absolute orientation of the laser beam is not exactly known.

For the statistical evaluation a temporal sampling of 60 seconds was chosen within a six hour time span from 10:00 UTC to 16:00 UTC. Since the sampling times of both instruments were not synchronized - the stereo system was not technically integrated into JOYCE - both the nominal time stamp and the real sampling time can differ by some seconds. The maximum nominal sampling time difference between two corresponding observations is chosen as 7 seconds, with the consequence that some observations do not contribute to the statistics. Due to the different orientation and missing synchronization both methods may not always sample the same part of a cloud, which can result in a higher variability, but also to some gross outliers that affect the comparison.

4.1.3.2 Evaluation via Georeferenced CBH-Time-Series

A non-scanning lidar-ceilometer offers CBH observations only for the zenith area; a validation of the cloud morphology in off-zenith regions is generally not possible. Mid- to high-altitude stratiform clouds and fast moving stratocumulus decks, however, often evolve over time-scales long enough to assume an approximately constant cloud base geometry for periods of 10 to 30 minutes. This allows to simulate a scanning capability of the lidar-ceilometer by comparing a single cross-section of the stereo reconstruction, taken in direction of cloud motion, with a time-series of CBH observations from the lidar-ceilometer. Such a comparison requires to convert the time-height observations by the lidar-ceilometer to a sequence of 3-D points within the world reference system, which can then be used to compute the 3-D geometric difference according to Eq. (4.2).

Let φ be the motion direction of a cloud layer from where it approaches, and v its velocity, for example derived from wind lidar observations or estimated cloud motion vectors (Chapter 5). The cross-section of the reconstruction represents the geometric state of the layer at time t_0 with the horizontal distance s to the reference camera ranging between s_{min} and s_{max} (e.g. $\pm$5000 m). The corresponding CBH time series by the lidar-ceilometer is determined by the time interval $t_{min} < t < t_{max}$ and converted to the common world reference system via

$$\begin{aligned} X_{ceilo}(t,\varphi,v) &= \cos(\varphi)(t_0 - t)v, \\ Y_{ceilo}(t,\varphi,v) &= \sin(\varphi)(t_0 - t)v, \\ Z_{ceilo}(t,\varphi,v) &= \mathrm{CBH}(t), \end{aligned} \quad \text{where} \quad \begin{aligned} t_{min} &= t_0 - \frac{s_{max} - s_{min}}{2v}, \\ t_{max} &= t_0 + \frac{s_{max} - s_{min}}{2v}. \end{aligned}$$

4.1.3.3 Evaluation via Cloud Radar Cross-Section Scans

The cloud radar stores the north-aligned azimuth angle φ^N, elevation angle θ_{elv} and the radial distance R of each reflectivity value. A comparison with the stereo reconstruction thus requires the conversion from spherical to Euclidean

(world) coordinates, according to

$$
\begin{aligned}
X_{rad}(\varphi^N, \theta_{elv}, r) &= R\ \cos(-\varphi^N + \frac{3}{2}\pi)\cos\theta_{elv}, \\
Y_{rad}(\varphi^N, \theta_{elv}, r) &= R\ \sin(-\varphi^N + \frac{3}{2}\pi)\cos\theta_{elv}, \\
Z_{rad}(\varphi^N, \theta_{elv}, r) &= R\ \sin\theta_{elv}.
\end{aligned}
$$

Note that the azimuth angle φ^N describes an orientation/rotation w.r.t. North and in reverse direction of rotation compared to the coordinate system of the stereo cameras.

Since a single radar beam can provide multiple reflectivity values that are almost collinear to a 3-D point of the stereo reconstruction, the computation of the geometric difference according to Eq. (4.2) includes only those reflectivity signals with the smallest radial distance R, which are assumed to represent the cloud boundary.

4.1.4 Evaluation Data

Image data was collected from different stereo configurations over the years 2014, 2015 and 2016 (Sec. 4.1.1). The dataset is selected with the purpose to address a specific evaluation problem (e.g. consistency of 3-D cloud morphology, accuracy via CBH observations) or atmospheric situation (e.g. multiple cloud layers). In general, a comparison between the stereo reconstruction and a lidar-ceilometer can be done almost permanently (e.g. via zenith-CBH statistics), a detailed investigation of cumulus morphology using the scanning capability of the cloud radar is more difficult due to the limited spatial and temporal coverage of a scanning sequence and the degrading sensitivity with distance.

Date	Cloud Type	Description	Focus
2014-07-24	Cu,Cb	Unstable, convection	Cu morphology
2014-08-05	Ac,Cu	Inversion, sparse Cu	Cloud layer
2014-08-11	Cb,Ac,As,Cu	Cyclone flow, showers	Cloud layer/radar
2015-07-29	Cb,Cu,Sc	Cyclone flow, showers	CBH variability
2015-08-19	Cc,Ac,Sc	Mixed clouds	Multi-layer clouds
2015-08-24	Ac,As,Cu	Mixed clouds	Cu sensed by radar

Table 4.1: Overview of experimental datasets obtained from stereo camera JOYCE, listed with dominant cloud type, atmospheric condition and evaluation focus.

4.2 Evaluation of Cloud Layer Reconstructions

Cloud layers allow to investigate both the validity of the estimated camera orientation parameters and the results from the stereo reconstruction. An error in the orientation parameters or a computer system time offset should produce a detectable systematic distortion or offset in the reconstruction, whereas a flawed stereo analysis can result in a random variability, a missing reconstruction or inconsistent artifacts of the cloud geometry. In this analysis, cloud reconstructions are limited to an area of 10×10 km^2 due to the increasing geometric uncertainty with distance to the cameras.

Figure 4.3 shows an analysis of an altocumulus cloud layer from 11 August 2014 at 14:14:00 UTC, which was captured by a cross-section scan of the cloud radar. The average geometric difference between the stereo reconstruction and the lower cloud boundary derived from the radar reflectivity signal amounts to 61 m with a standard deviation of 32 m. The estimated CBH is 2897 m by the lidar-ceilometer and 2881 m by the cameras.

The horizontal extent of the cloud as derived by the stereo cameras and the cloud radar differs significantly. While the reconstruction occupies a distance range from -2000 m to 5000 m, the radar signal suggests a cloud within -1000 m and 3000 m. A likely explanation in this case is the decreasing sensitivity of the radar with increasing distance (between 3000 m and 5000 m/-2000 m and -1000 m) and smaller droplet size due to evaporation at the edge of the cloud. One should, however, take into consideration that a flawed cloud mask (Sec. 2.3.1) can lead to a larger (smaller) geometric extent of the cloud reconstruction if the threshold is chosen too large (low).

Figure 4.3c shows a comparison between a time-series of CBH values by the lidar-ceilometer and a reconstruction cross-section along the layers direction of motion, according to the methodology described in Sec. 4.1.3.2. It reveals a small, but steady increase of the layers base height from 2800 m to 3000 m, which is also detected by the stereo analysis. Within a horizontal distance from 1000 m to 3000 m the reconstruction shows a higher variability, resulting in a difference of up to 400 m to the CBH from the lidar. The variability is most likely caused by the strong forward-scattering of direct solar radiation, causing an over-saturation of the sensor elements in the respective image region and a flawed disparity estimation during the stereo analysis. Ignoring these outliers, the stereo cameras match the lidar-ceilometer CBH on average up to 9 meters with a standard deviation of 59 m (2.04% of ceilometer CBH).

Stereo analysis can detect multiple cloud layers in the stereo reconstruction as long as they are partly visible in both cameras and offer a minimal amount of contrast. A case study from 19 August 2015 at 13:50 UTC shows a two-layer situation with a cirrus/cirrocumulus layer between 6 and 8 km height and a translucent stratocumulus layer at about 2.3 km height, which is barely visible as it was evaporating near zenith (Figure 4.4). Both layers were recognized by the stereo analysis, yielding clearly distinguishable reconstructions. Com-

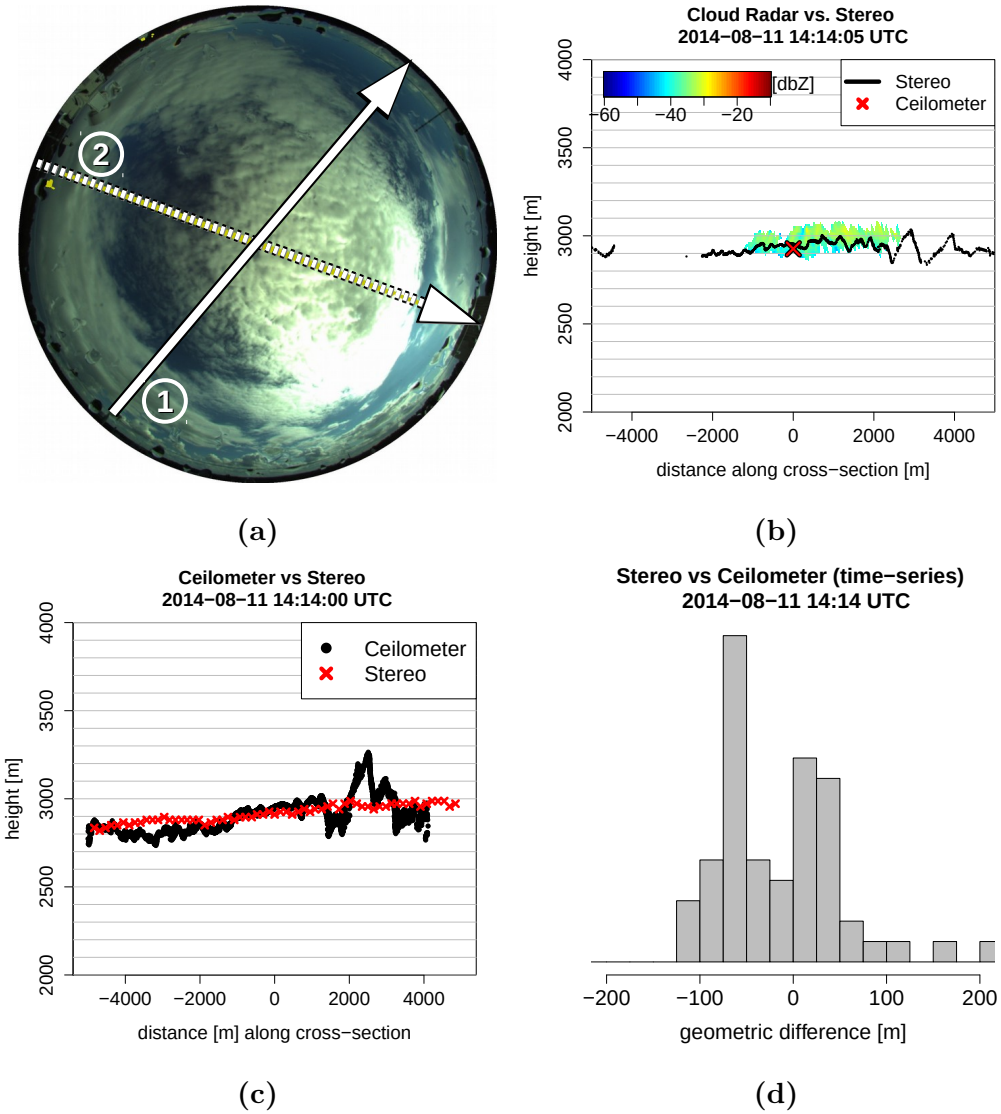


Figure 4.3: Comparison of a stereo reconstruction to the reflectivity profile from a cloud radar and validation with CBH observations by a lidar-ceilometer at 11 August 2014 14:14 UTC. The original fisheye image from the reference camera at JOYCE is shown in (a) with marked directions of two cross-sections (arrow base: negative distance, arrow top: positive distance) with (b) corresponding to (1) and (c) corresponding to (2). The cross-section of the reconstruction is (b) compared to the reflectivity profile by the cloud radar and (c) to the georeferenced time series of CBH observations by the lidar-ceilometer.

pared to the zenith CBH values by the lidar-ceilometer, the reconstruction shows differences between -100 m and -200 meters for the stratocumulus layer. The reconstruction resembles, however, the lower boundary of the backscatter profile up to about 50 m, which can be explained by the CBH retrieval algo-

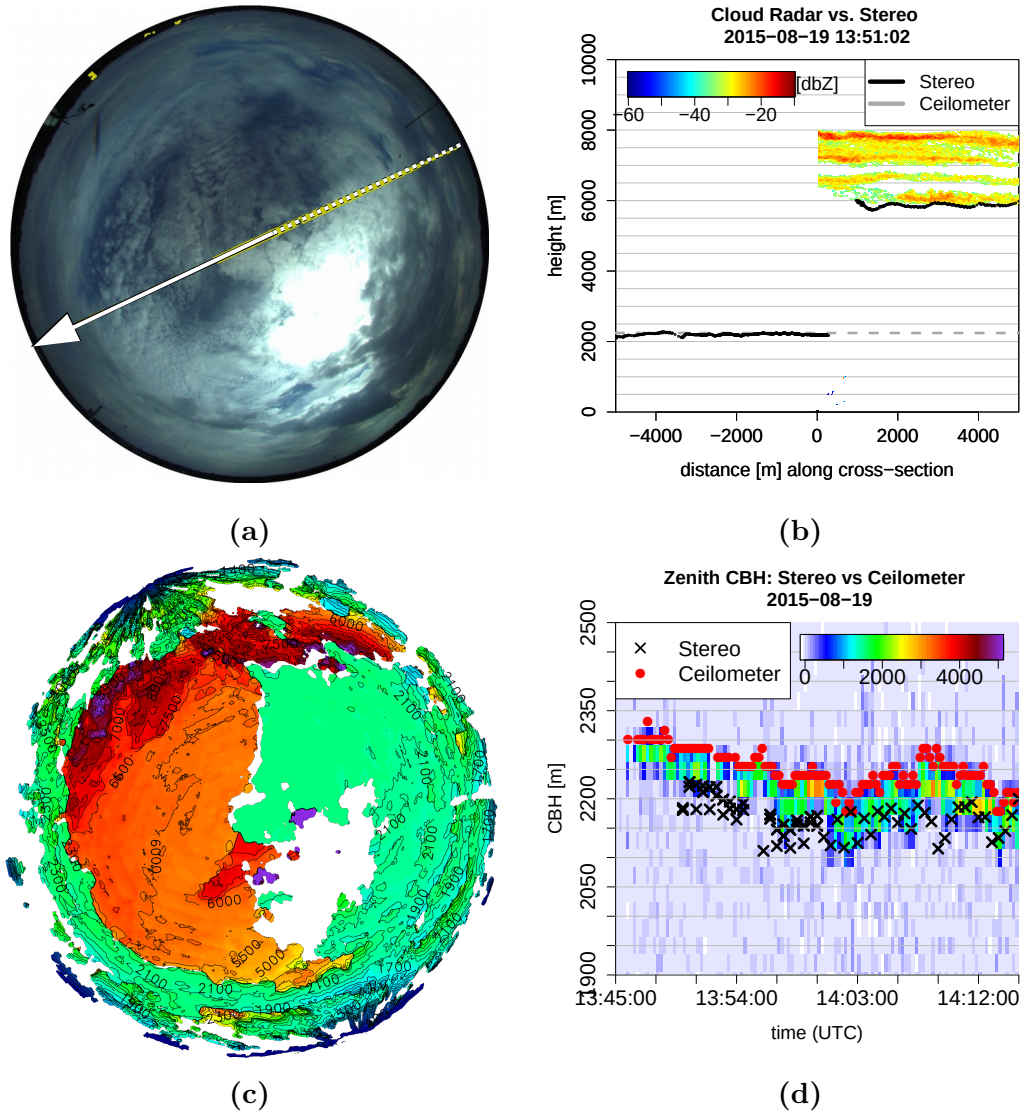


Figure 4.4: Evaluation of multi-layer cloud reconstructions from 19 August 2015 with (a) the original fisheye image and (c) a corresponding cloud height map, which indicates a higher cirrus/cirrocumulus layer at 6 km height and a lower translucent stratocumulus at 2.3 km height. Also shown is (b) a comparison with the reflectivity profile from the cloud radar and (d) a comparison with the cloud base height (CBH) and backscatter profile (in $(srad\,m)^{-1}$) from the lidar-ceilometer in. Note that in this case the radar scan was scheduled to cover an elevation range of only 75°.

rithm that tends to locate the CBH above the strongest increase in backscatter. The reconstruction of the cirrocumulus cloud yields a standard deviation of 53 m with a mean difference to the reflectivity-derived cloud layer boundary of 76 m.

4.3 Cumulus Morphology Retrieval with SGBM and PMVS

The following analysis evaluates the 3-D reconstruction of cumulus clouds and the potential to capture the irregular morphology. In most cases, the scans of the cloud radar did not cover the investigated cumulus clouds due to the limited number of scans per hour (15 minutes) and a fixed scanning direction. An inter-comparison of the SGBM and PMVS stereo reconstruction methods (compare Sec. 3.5.2 and Sec. 3.5.3) is provided to investigate the performance for such clouds.

The shape reconstruction of cumulus clouds was investigated in detail at 24 July 2014. The atmospheric condition could be characterized as unstable, with active convection and significant vertical wind shear. Under these conditions, strong updrafts caused highly tilted convective turrets.

A developing cumulus cloud at 11:32 UTC was analyzed that was reaching its terminal height of approximately 3700 m (Figure 4.5). The vertical wind shear and the convective updraft led to an increasingly concave shape of the cloud. The cross-section through the 3-D reconstruction (yellow line in Figure 4.5) suggests that the principal geometry was recovered correctly, since cloud top (1), the smaller convective tower (3) and the cloud base (2) are recognizable. Judging from the temporally closest and lowest CBH values from the lidar-ceilometer, which indicate a CBH of 1572 m (at 11:22 UTC), the base is overestimated by up to 100 meters at its central region at 4000 m distance.

SGBM and PMVS were both applied to the stereo images in full resolution and a template size of 11 pixel. The reconstructions differed by less than 100 m and mostly below 50 m (Figure 4.5f). Since PMVS relies on the detection of interest points, it is unable to model the cloud base adequately, which lacks any distinctive image areas due to a low contrast.

While cloud base (2) and the shadowed section at (1), below the convective turret were reproduced, such low-contrast areas are likely not sufficiently pronounced; regularization instead of template matching then dominates the disparity estimation during stereo analysis. The geometric uncertainty is well below 25 m at well-structured parts and reaches values of 75 m and more at the texture-less area near (1) Figure 4.5c). The highest uncertainty can be found near the cloud base where it reaches typical standard deviations between 50 and 150 m. Such areas often have insufficient contrast for a valid stereo match, resulting in a missing reconstruction. The secondary convective tower at (3) is also recognizable in the cross-section, reaching a height of about 2700 m. When using a fixed-size template window for the stereo matching, such geometries tend to be smoothed and over-sized, since the small-scale curvature cannot be resolved. A smaller template window could be used without sacrificing too much geometric precision, while capturing those higher frequency geometries in more detail.

The reconstruction of clouds that appear near the left and right border of the

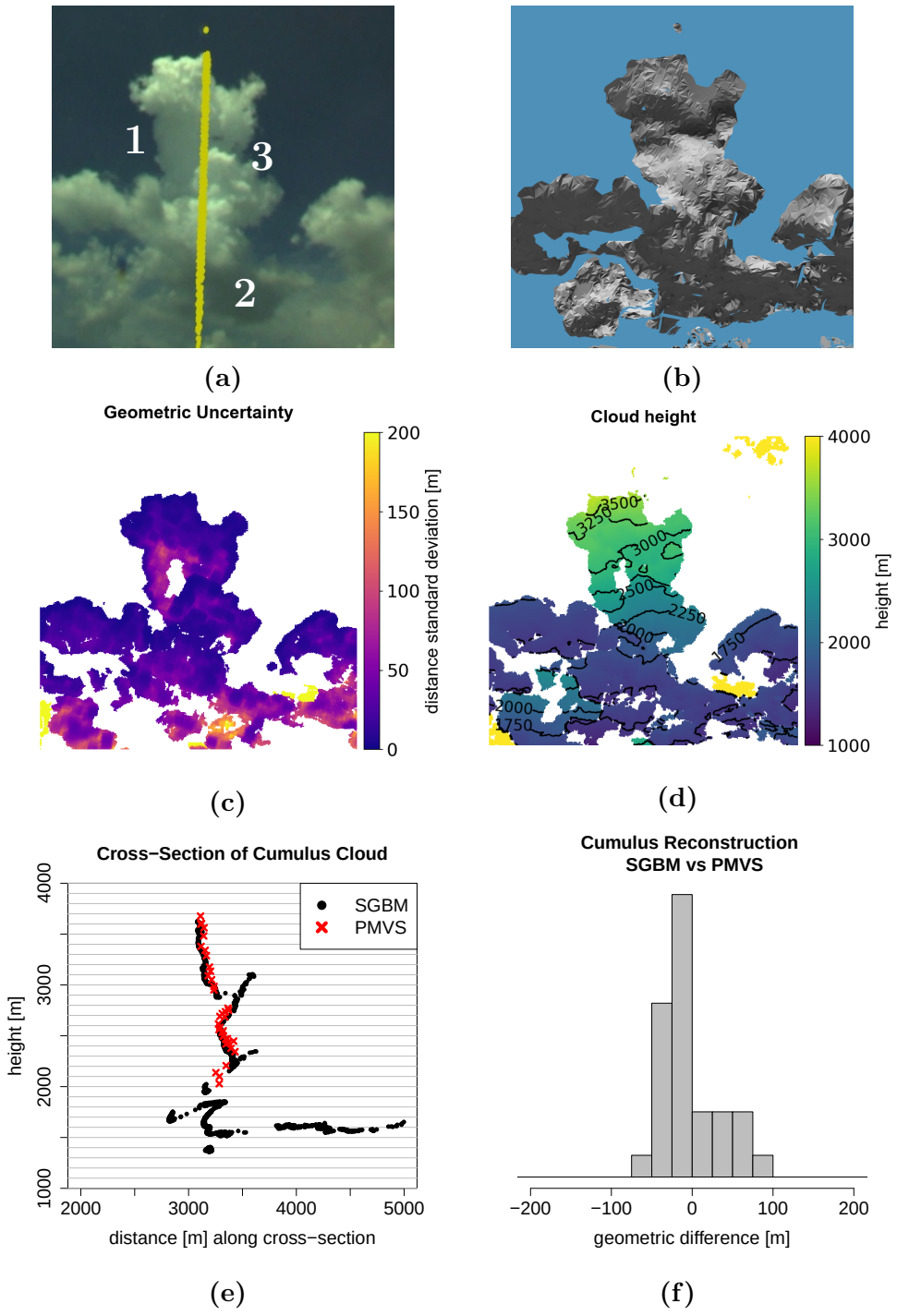
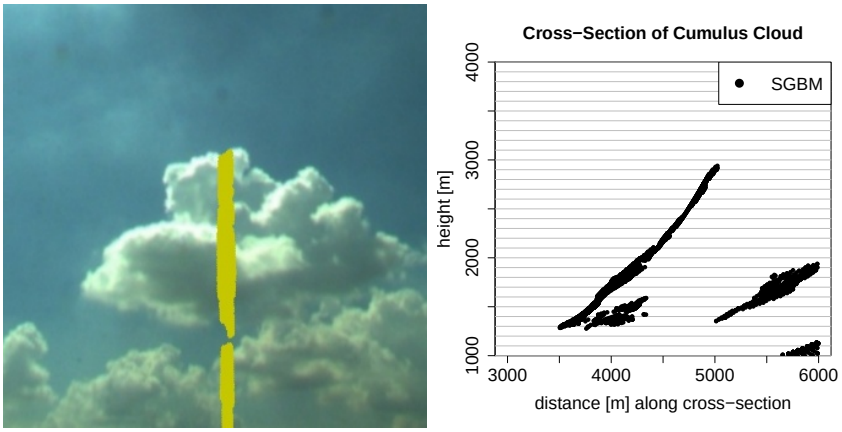


Figure 4.5: Analysis of cumulus morphology and reconstruction precision from 24 July 2014 as produced by SGBM and PMVS in full image resolution. Shown is (a) a close-up view from the original fisheye image with marked cross-section and three areas of interest, (b) a 3-D cloud boundary mesh derived from the point cloud and visualized in *Blender*, (c) uncertainty of geometric distance to reference camera, (d) height-contour plot, (e) cross-section of reconstructions and (f) geometric distance between cloud reconstructions as obtained from SGBM and PMVS. Note, that (b) shows the reconstruction obtained from half image resolution.



(a) **(b)**

Figure 4.6: Stereo reconstruction of cumulus cloud located in a region of bad geometric conditioning, collinear to the baseline. **Left**: Detail view of the cloud and marked location of cross-section (yellow line). **Right**: Distance-height plot of cross-section.

epipolar rectified stereo images are subject to significant geometric uncertainty due to the small enclosing angle between the projection rays (Sec. 3.6). Also, the image distortion in those areas leads to additional deformation of the reconstruction (Figure 4.6). While the effect on cloud base reconstruction can be profound, the geometric conditioning for cloud top height estimation is less severe, as the vertical extent and the high contrast in these areas results in both a larger enclosing angle between the projection rays and a reduced directional uncertainty.

For an estimation of cloud cover or perimeter of individual clouds, the stereo matching can be applied to the images in a lower resolution to achieve a higher degree of geometric completeness, minimizing the ratio of invalid reconstructions due to low contrast. This comes, however, at the cost of a reduced geometric detail.

4.4 Comparison of Cumulus Morphology Retrievals to Cloud Radar Reflectivity Profiles

In this section, the stereo reconstructions are compared to cross-wind RHI scans conducted by the cloud radar, as described in Sec. 4.1.3.3. CBH observations from the lidar-ceilometer are used to complement the comparison for clouds that pass the cameras at zenith.

A site-crossing cumulus cloud from 24 August 2015 was captured by the cross-section scan of the radar and compared to the stereo reconstruction (Figure 4.7). The reflectivity profiles of three scans between 14:03 UTC and 14:04 UTC show the vertical structure of the cloud with a clearly recognizable con-

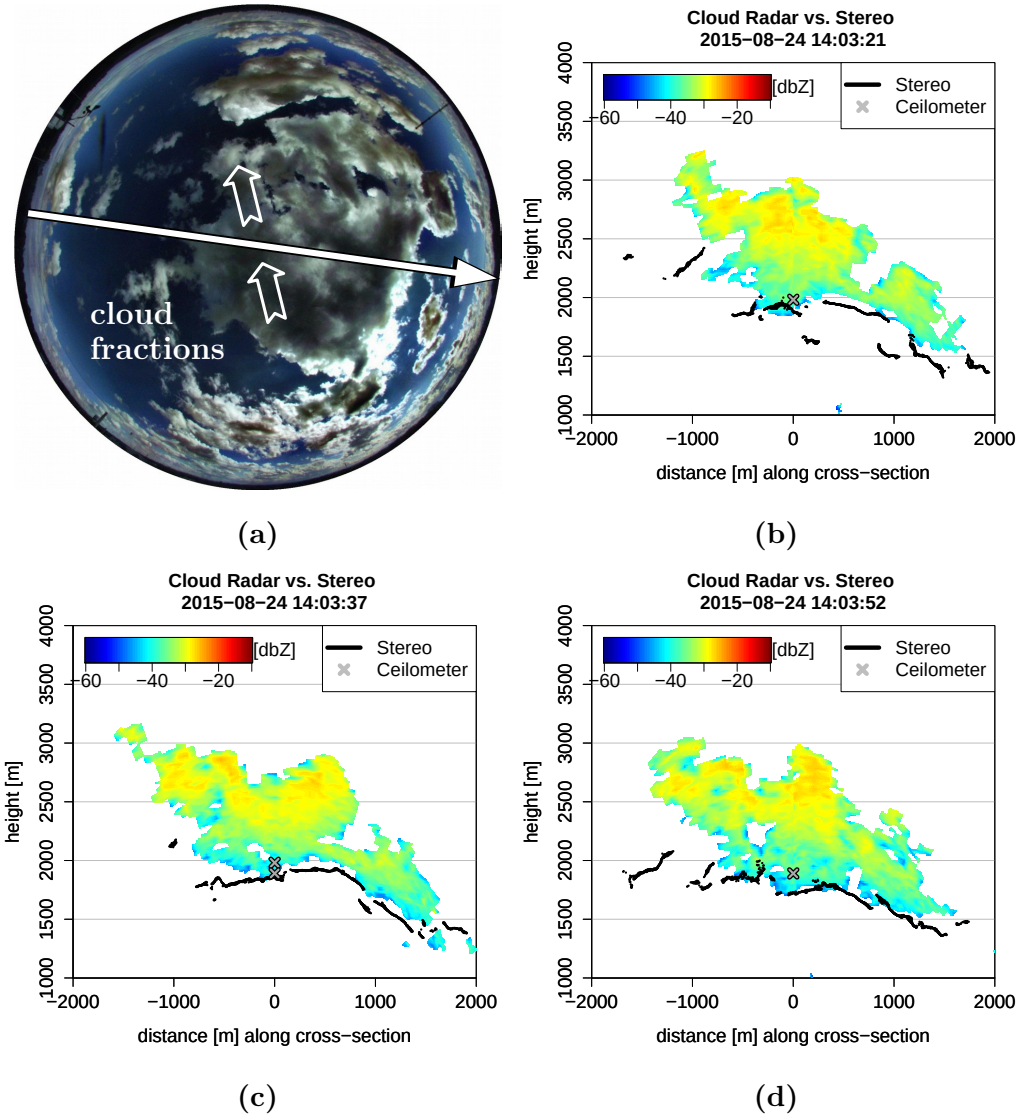


Figure 4.7: Comparison of cumulus stereo reconstruction with cloud radar reflectivity profiles from 24 August 2015 at 14:03 UTC. Shown is (a) the original fisheye image with marked direction of the cross-section (long white arrow) and comparisons between the reconstruction (black line) and the radar signal (b) at 14:03:21 UTC, (c) 14:03:37 UTC and (d) 14:03:52 UTC, respectively. The cloud motion direction is marked by the two white arrows.

vective top at around 3000 m. The radar reflectivity reaches its maximum in the upper half of the cloud due to droplet growth within ascending air parcels containing mostly undiluted cloud air. Towards the cloud base the signal strength decreases monotonously down to about -45 dBZ due to the small cloud droplets that are typical for continental cumulus.

Each reconstruction reproduces the principal irregularity and shape of the

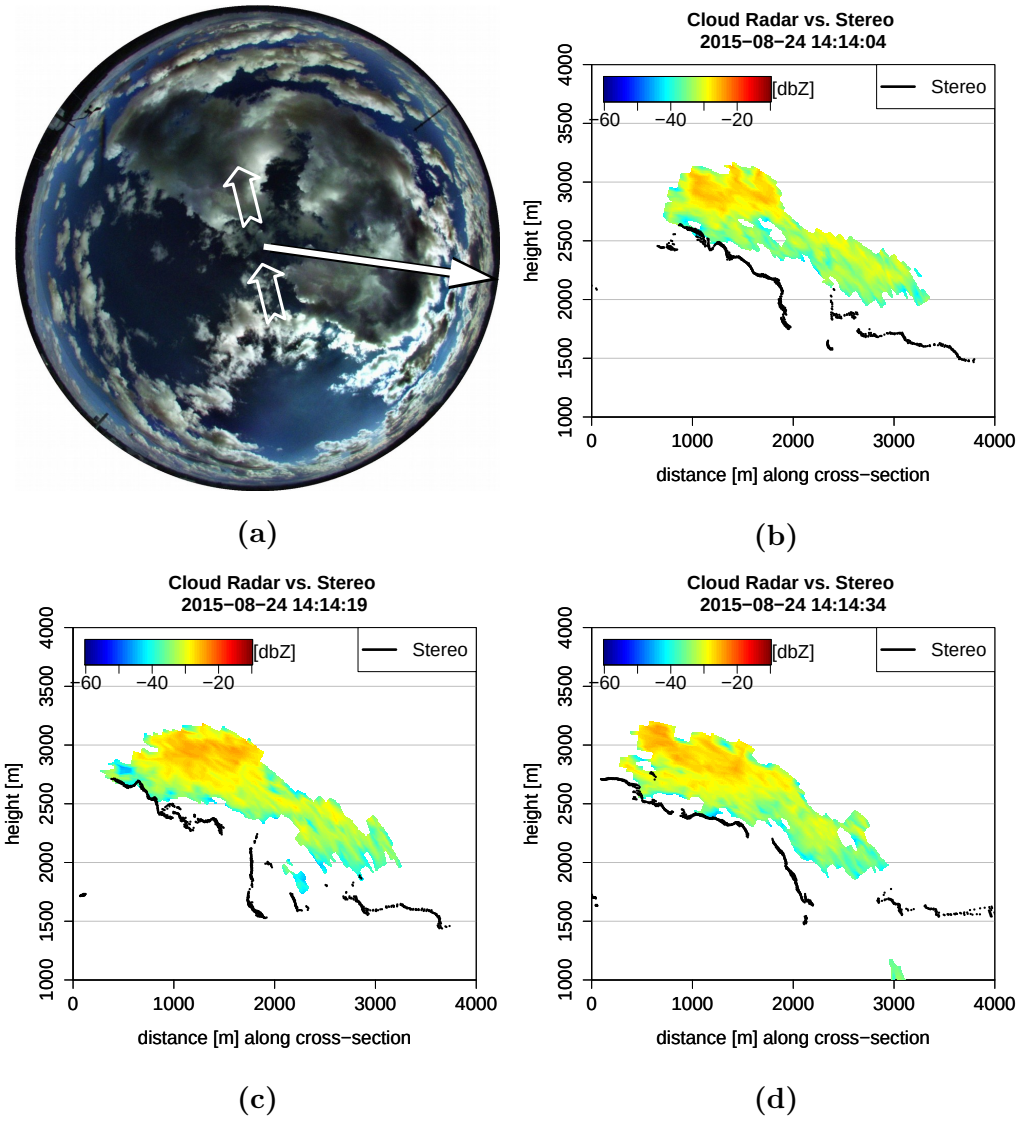


Figure 4.8: Comparison of cumulus stereo reconstruction to cloud radar reflectivity profiles from 24 August 2015 (b) at 14:14:04 UTC, at (c) 14:14:19 UTC, at (d) 14:14:34 UTC the same way as in Figure 4.7.

cloud. Median absolute distance to the radar-derived boundary amounts to 81 m at 14:03:21 UTC, 72 m at 14:03:37 UTC and 35 m at 14:03:52 UTC. The estimated cloud boundary tends to be closer to the cameras than the radar-derived boundary. Larger differences between the reconstruction and the reflectivity signal can most likely be explained by cloud fractions (Figure 4.7a), which consist of irregular distributed condensate. These fractions are already diluted by environmental air and probably contain small droplets, resulting in a low reflectivity signal, but were detected by the sky imagers, which then produced a 3-D signal below the radar reflectivity.

As the cloud crosses JOYCE at zenith, a comparison between the CBH

values from the lidar-ceilometer and the stereo cameras was possible, which showed differences between 41 and 179 meter; deviations are also recognizable between the radar signal and the CBH value from the lidar. Using a wider time interval between 14:00 and 14:15 UTC, however, reveals an average difference (ϵ_{rel}) of 0.25% and a standard deviation of 7.2%. Thus, the these inconsistencies can be interpreted as outliers, probably caused by the irregular cloud structure, observation time differences and the tilting angle of the lidar-ceilometer.

A different situation could be observed a few minutes later at 14:14 UTC when another cloud crossed the JOYCE site on an off-zenith path for which no CBH observations were available from the lidar-ceilometer (Figure 4.8). The reconstructed lower part of the convective top approximates the radar derived cloud boundary and showed typical geometric differences between 34 m and 170 m. The cloud base (at around 3000 m horizontal distance), however, was reproduced with high uncertainties and geometric differences of more than 800 m. The radar derived boundary suggests that the cloud base was located at about 2000 m height and only extended between 2500 m and 3500 m horizontal distance.

A likely explanation for these inconsistencies is the too low sensitivity of the radar, which is a known problem of scanning cloud radars and is also reported by Fielding et al. (2013) and Oue et al. (2016). On the other hand, any part of a cloud that shows poor contrast introduces considerable uncertainty to the stereo reconstruction. This can lead to missing 3-D information due to an invalid matching result or a discrepancy in the depth value. The geometric uncertainty for this cloud yielded standard deviations of up to 100 m, but mostly well below 50 m, which cannot explain the extreme difference of 800 m. An insufficient radar sensitivity is thus a likely explanation.

Since no reliable CBH data was available for off-zenith regions and both instruments, the stereo camera and the cloud radar, could have been affected by high uncertainties. A simple visual inspection could help to find the source of this inconsistency: Since 2-D cross-section scans of the radar can be converted from 2-D polar to 3-D Cartesian coordinates, according to Sec. 4.1.3.3, they can be projected onto the hemispheric image. Both signals should then cover the full extent of the cloud in the image along the respective cross-section. Figure 4.9 shows the situation one minute earlier at 14:13 UTC. The cross-section showed a large disagreement between the signal of the radar and the reconstruction, particularly for off-zenith regions. The radar was able to sense the vertical structure of the cloud at zenith, and the stereo camera could replicate the principal geometry of its lower boundary. The cloud structure at 2000 m distance, however, was barely sensed by the radar, while the stereo cameras identified multiple cloud fractions at a height of approximately 1600 m. Apparently, a significant part of the cloud was not covered by the projected radar signal (yellow line/gray circle in Figure 4.9). Even under assumption of a possible time shift between the two instruments, the signal would still have

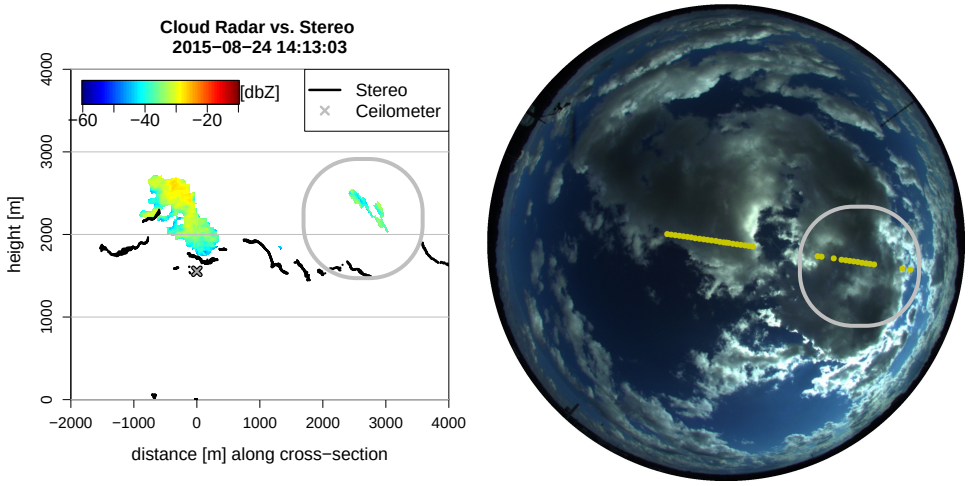


Figure 4.9: Left: Cross-section of the reconstruction and corresponding cloud radar reflectivity signal at 14:13 UTC. **Right**: Central-projection of the georeferenced cloud radar reflectivity profile to image of the reference camera (yellow line).

missed many parts of the cloud.

4.5 Accuracy Estimation via Time-Height Profiles of Zenith-CBH

The following analysis investigates the accuracy of the stereo method empirically by comparing the CBH observations from the lidar-ceilometer to the corresponding CBH values derived by the stereo cameras. Multi-hour stereo images of several days with different atmospheric conditions were processed, following the methodology described in Sec. 4.1.3.1. From the lidar-ceilometer, only the first detected CBH was used.

Figure 4.10 shows the results from 5 August 2014, which was dominated by layers of slowly moving altocumulus cloud layers at about 3 km height and sparse cumulus convection. The analysis reveals for the cloud layer a persistent vertical offset of -141 m and a standard deviation of 48 m, which corresponds to a relative distance bias and variability of -4.72% and 1.53%, respectively. The vertical offset could be the result of a flawed relative orientation estimation, but also an asynchronous recording, which could be observed at times within this experimental period (Sec. 3.7). A few cumulus clouds are visible with a CBH of 1300 m (between 10:36 and 11:12 UTC) and 1600 m (between 14:12 and 14:48 UTC). Between 13:36 UTC and 14:12 UTC several CBH values of about 3 km are obtained by the stereo cameras, which have no correspondence by the lidar-ceilometer. These represent the detected edge of an altocumulus cloud that could be detected due to the larger sampling area used by the stereo cameras to extract the CBH.

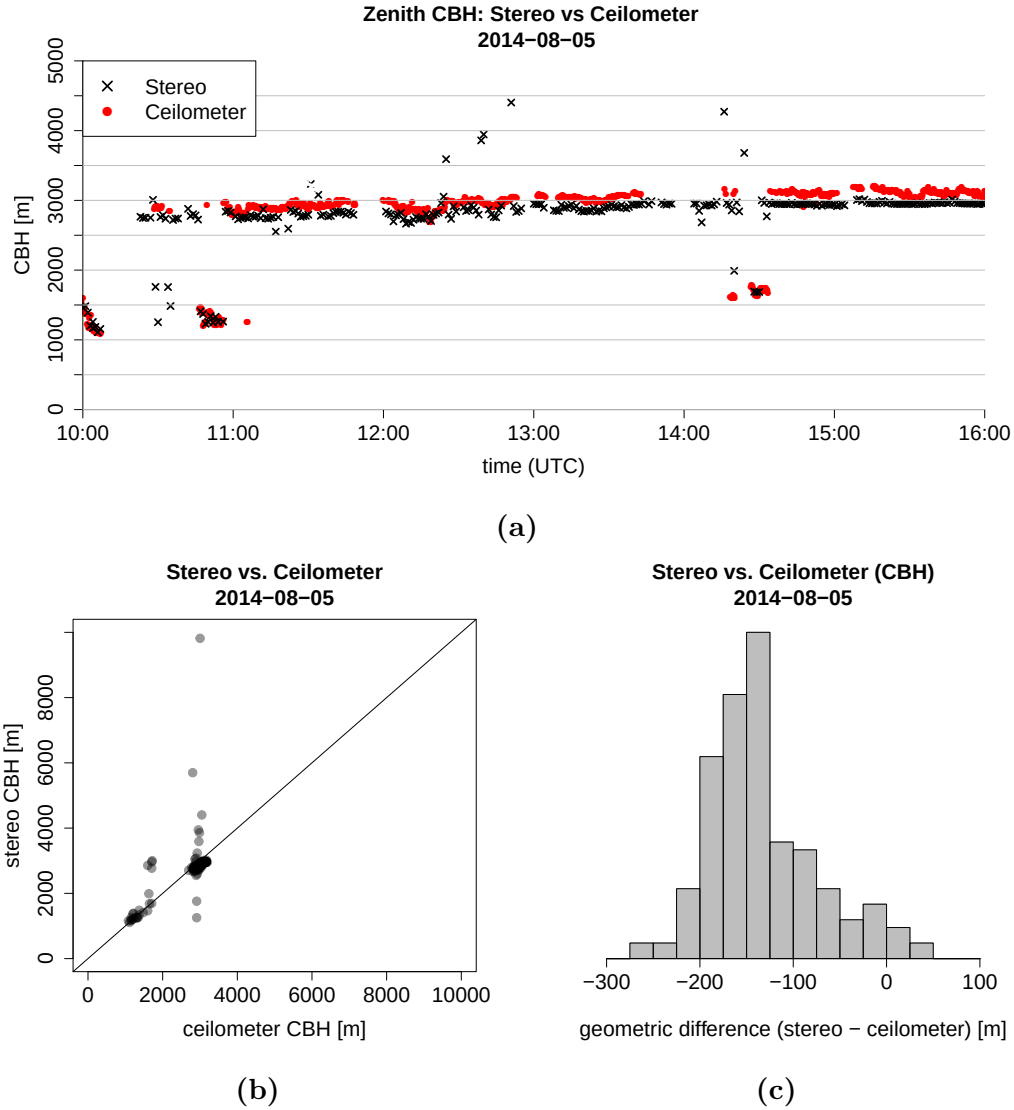


Figure 4.10: Comparison of zenith cloud base height (CBH) from the stereo cameras and the lidar-ceilometer from 5 August 2014. Shown is (a) a time-height plot of all observations (lidar: first detected CBH), (b) a scatter plot of synchronous observations and (c) the difference between synchronous observations.

Occasionally, the lidar-ceilometer and the stereo cameras sample different clouds or parts of a cloud because of their individual orientation and sampling area. The variable geometry of cumulus clouds or the simultaneous occurrence of low and high clouds may then cause higher deviations of the estimated CBH or even gross outliers, such as the inconsistent CBH observations between 14:12 UTC and 14:48 UTC: While the stereo camera detected the cloud layer at 3 km height, the lidar-ceilometer received a signal from the cumulus cloud at 1600 m height.

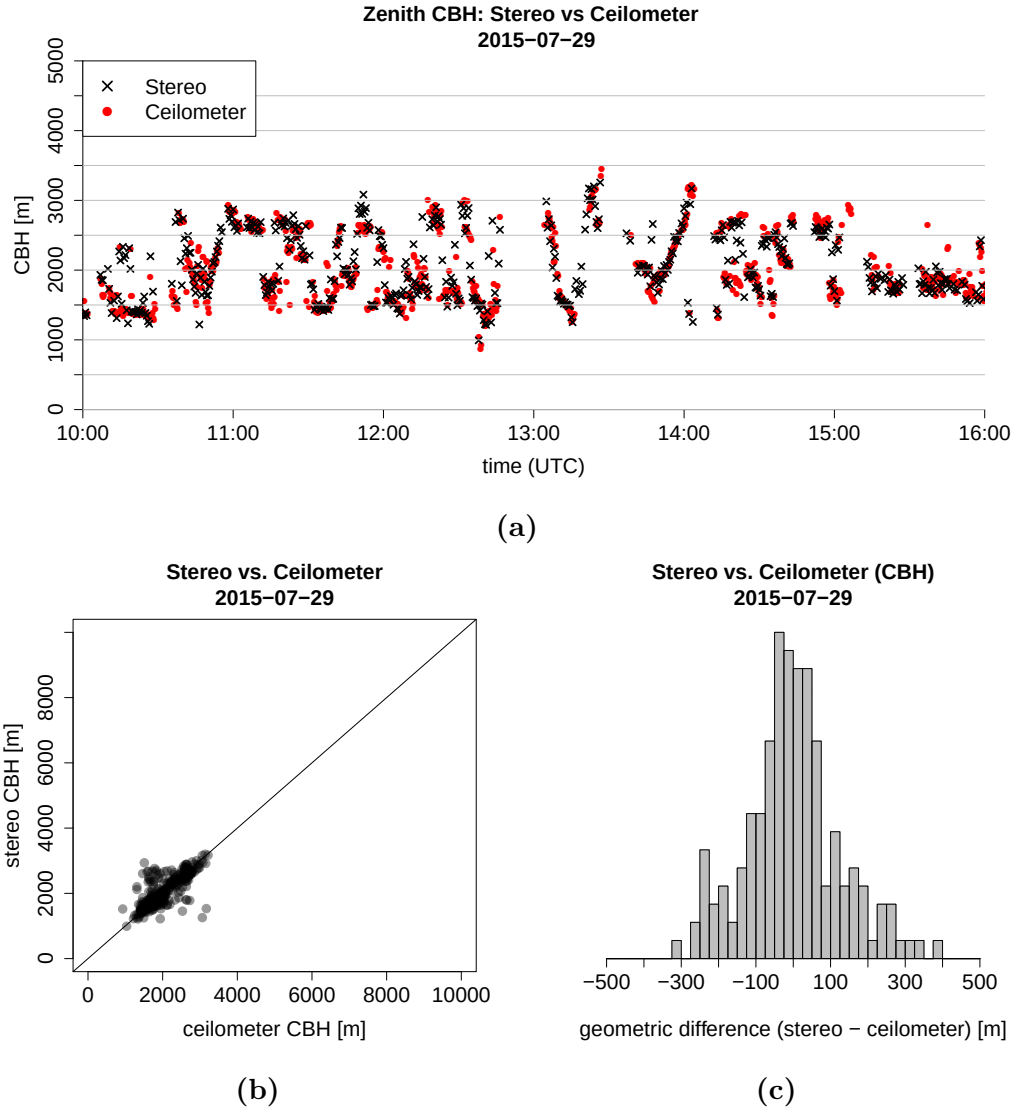


Figure 4.11: Comparison of CBH observations at zenith by the stereo cameras and the lidar-ceilometer from 29 July 2015 the same way as in Figure 4.10.

A different situation could be observed for the 29 July 2015, which was characterized by a broken cloud condition with quickly moving fields of strato-, altocumulus and cumulus (Figure 4.11). While the geometric distance showed a larger spread (138 m/6.67% of CBH) compared to the 5 August (48 m/1.53% of CBH), the mean estimated CBH agreed much better with a bias of -10 m (-0.6% of CBH). A likely reason for the higher spread is the spatial heterogeneity of these types of clouds, but also the differences in sampling area, time and orientation, which have a larger impact if the cloud field moves considerably faster. Table 4.2 summarizes the accuracy and precision of the stereo reconstructions for the experimental datasets for different cloud types. Note that CBH statistics for cumulus clouds may be affected by a smaller number

of observations

Date	Type	CBH	$\overline{\epsilon}_{proj}$	$\overline{\epsilon}_{rel}$	$\sigma(\epsilon_{proj})$	$\sigma(\epsilon_{rel})$
2014-08-05	Ac	3002 m	-143.1 m	-4.7%	46.3 m	1.4%
2014-08-11	Ac	2911 m	11.1 m	0.4%	36.5 m	1.2%
2014-08-05	Cu	1326 m	29.3 m	2.3%	111.8 m	7.7%
2015-08-24	Cu	1600 m	-2.6 m	0.3%	167.1 m	10.3%
2015-08-24	Ac,As	4846 m	0.3 m	0.1%	164.2 m	3.4%
2015-08-24	Sc	2918 m	36.1 m	1.2%	72.4 m	2.5%
2015-08-19	Sc	2271 m	81.8 m	-3.6%	22.8 m	0.9%

Table 4.2: Comparison of CBH obtained from the stereo cameras and a lidar-ceilometer for typical cloud types: the table lists the median CBH from the lidar-ceilometer CBH observations, mean geometric distance $\overline{\epsilon}_{proj}$, mean relative geometric distance $\overline{\epsilon}_{rel}$, standard deviation of the geometric distance $\sigma(\epsilon_{proj})$ and standard deviation of the relative geometric distance $\sigma(\epsilon_{rel})$.

4.6 Summary

In this chapter, the 3-D cloud reconstructions obtained from a single stereo camera pair with a baseline of 300 m were confronted to observations from a lidar-ceilometer and a cloud radar. If compared to a lidar-ceilometer, the obtained time-height profiles of cloud base height (CBH) often showed good agreement. The geometric difference distributions revealed a typical bias of less than 1-2% of CBH for all cloud types, but also some occasions with a significant bias of 3-5% of CBH. Standard deviation of geometric difference was generally smaller for well-structured clouds like altocumulus than for more dynamic and heterogeneous stratocumulus or cumulus clouds, with values of about 50 m (1.5% of CBH) and 100 to 150 m (7 - 10 % of CBH) respectively.

An analysis of individual cloud layers (Sec. 4.2) and cumulus clouds (Sec. 4.3 and Sec. 4.4) showed that the method was able to reconstruct large parts of the local cloud field and the principal geometry of clouds, including multi-layer clouds. A comparison to the reflectivity signal of the cloud radar showed typical absolute differences between both signals of 35 - 81 m for near-zenith cumulus clouds. For more distant clouds (3-5 km), typical differences ranged between 50 m and 200 m, but both signals also deviated significantly by several hundred meters near the cloud base, which was probably the result of an insufficient radar sensitivity for very small cloud droplets.

The reconstructions based on the accurate PMVS (Sec. 3.5.3) and the spatially dense SGBM (Sec. 3.5.2) stereo algorithms were compared for a convective turret and showed differences between the two methods of less than 100

m and mostly below 50 m. Since PMVS relies on the detection of distinctive interest points, which require high-contrast image areas, it does, however, not provide the same geometric completeness compared to a dense stereo approach like SGBM. The geometric uncertainty of the reconstruction in case of cumulus clouds yielded a standard deviation of distance that ranges from 10 to 20 meter for high-contrast areas, e.g. at the convective top, to more than 100 m in low-contrast areas, such as the cloud base.

5 Estimation of Cloud Motion Vectors

This chapter investigates the derivation of cloud motion vectors (CMV) using a pair of hemispheric sky imagers. Cloud motion vectors of stratiform clouds can be used to estimate wind speed and direction in the respective height. A cloud motion analysis of developing cumuli allows classification of macroscopic motion patterns and tracking for life-cycle analysis.

Cloud top motion estimation is a well-established technique in satellite remote sensing and an important source of information for numerical weather prediction (e.g. Seiz et al., 2007b; Mueller et al., 2017). If the cloud top height is known, detected cloud features (compare Sec. 3.5.1) can be tracked over a sequence of images and allow the estimation of horizontal cloud motion.

This chapter provides an analysis of cloud motion vector fields of cloud layers and convective clouds. Several cases of cloud layer motion vectors are confronted to the observations from a nearby wind-lidar to evaluate the reliability and accuracy of an operational retrieval. Within small time intervals, cloud layers are assumed rigid as they usually only show small variations of shape and appearance. This assumption is generally not valid for convective clouds, which are highly dynamical and subject to morphological changes. In many cases, however, the derived motion vectors can be used to describe the morphological motion pattern of cumulus clouds, as are produced by vertical wind-shear.

Sec. 5.1 introduces the concepts of *scene flow* and *optical flow*, while the specific optical flow algorithm used for the experiments is explained in Sec. 5.2. Sec. 5.3 explains the working principles of the wind-lidar system, which was used for an evaluation of the obtained cloud motion, according to Sec. 5.4.1. The experiments covered the estimation of atmospheric motion in cases of altocumulus cloud layers (Sec. 5.4.2) and deformation vectors of a cumulus cloud under vertical wind shear (Sec. 5.4.3).

5.1 Estimation of Cloud Motion Vectors via Scene Flow

The estimation of motion direction and velocity of an object is possible if its 3-D position at two different times is known. This assumes, however, that the motion path is linear and velocity is constant during this time period. In the

following, it is assumed that these assumptions are met within a sufficiently small range of time.

Let $\mathbf{X}_{t_1}$ be a 3-D object point at time t_1 and $\mathbf{X}_{t_2}$ the same point at time $t_2 = t + \mathrm{d}t$ with $\mathrm{d}t > 0$. If the object point is observed by a stationary hemispheric stereo camera, its estimated 3-D location at time t_1 and t_2 is determined by the triangulation function τ:

$$\hat{\mathbf{X}}_{t_1} = \tau(\mathbf{x}_{\mathbf{1}t_1}, \mathbf{x}_{\mathbf{2}t_1}, \mathbf{P}_1, \mathbf{P}_2),$$
$$\hat{\mathbf{X}}_{t_2} = \tau(\mathbf{x}_{\mathbf{1}t_2}, \mathbf{x}_{\mathbf{2}t_2}, \mathbf{P}_1, \mathbf{P}_2),$$

where $\mathbf{P}_1$ and $\mathbf{P}_2$ are the respective camera matrices, and $\mathbf{x}_{\mathbf{1}t_1}, \mathbf{x}_{\mathbf{2}t_1}$ and $\mathbf{x}_{\mathbf{1}t_2}, \mathbf{x}_{\mathbf{2}t_2}$ are the corresponding projections at time t_1 and t_2. The motion is then defined by

$$\mathbf{v} = \frac{\hat{\mathbf{X}}_{t_2} - \hat{\mathbf{X}}_{t_1}}{\mathrm{d}t}, \qquad v = \frac{\|\mathbf{v}\|}{\mathrm{d}t},$$

where $\mathbf{v}$ denotes the 3-D *scene flow* (Brox et al., 2004), i.e. the motion vector relative to the camera system, and v its magnitude (velocity).

The *optical flow* defines a 2-D motion vector field within an image, which describes the change in the image signal, e.g. due to observation of a moving object. Following Förstner (1993), matching of two (possibly time-varying) 2-D image intensity signals can be formulated as

$$I(u, v, t) = I(u + \mathrm{d}u, v + \mathrm{d}v, t + \mathrm{d}t),$$

which for a sufficiently small time interval $\mathrm{d}t$ and translation $[\mathrm{d}u, \mathrm{d}v]$ can be linearized as

$$I(u, v, t) \approx I(u, v, t) + I_u\,\mathrm{d}u + I_v\,\mathrm{d}v + I_t\,\mathrm{d}t$$

with $\mathrm{d}u = f_u\,\mathrm{d}t$, $\mathrm{d}v = f_v\,\mathrm{d}t$ and $\mathbf{f}(u, v, t) = [f_u, f_v, t]$ the optical flow field at time t. Often, the time interval $\mathrm{d}t$ is set to one time unit, which then results in an equivalence of the optical flow and the horizontal displacement in the image, i.e. $[f_u, f_v] = [\mathrm{d}u, \mathrm{d}v]$.

Since the optical flow field only reflects motion caused by a visually recognizable object, the presence of a flow field does not necessarily reflect real motion, while the absence of optical flow does not imply that the observed scene is static. For example, a lenticularis cloud generated by orographic overflow exhibits only little motion, appears mostly stationary and only provides minor changes in contrast or shape, but is nevertheless caused by a constant advection of air masses. Similarly, mutual shadowing of clouds can result in a flow field that reflects the shadows' motion rather than the ego-motion of the cloud it is projected on.

Similar to stereo analysis, optical flow relies on contrast in the image signal. The precision of optical flow is thus equivalent to the precision of a 2-D template matching and can be described by the structure tensor introduced in

Sec. 4.5. Without any prior information about object motion (or the camera motion), its computation requires for each image point $\mathbf{x}$ the determination of the horizontal and vertical displacement, $\mathrm{d}u$ and $\mathrm{d}v$, which usually renders optical flow computationally more demanding.

5.2 Variational Optical Flow

For the experiments, the optical flow routine of Liu et al. (2009) was used, which implements a variational approach based on the work of Brox et al. (2004). The optical flow problem is formulated in terms of an energy functional

$$E(u, v) = E_{data} + E_{smooth},$$

where E_{data} accounts for consistency based on the image data and E_{smooth} penalizes strong gradients of the resulting flow map. Estimation of an optical flow field relies on the following assumptions:

- *Brightness Constancy*: The intensity of two corresponding image points is conserved:
$$I(u, v, t + \mathrm{d}t) = I(u + \mathrm{d}u, v + \mathrm{d}v, t). \tag{5.1}$$
- *Gradient Constancy*: The gradient at two corresponding image points is conserved:
$$\nabla I(u, v, t) = \nabla I(u + \mathrm{d}u, v + \mathrm{d}v, t + \mathrm{d}t) \quad \text{,where} \quad \nabla = (\partial_u, \partial_v)^\top. \tag{5.2}$$
- *Smoothness Assumption*: The obtained optical flow field is assumed piecewise smooth. An additional smoothness term regulates the variability of the resulting flow map by penalizing the spatio-temporal gradients of each motion component

The first two assumptions are directly related to the image intensity signal and are included in the data term E_{data}. The gradient constancy assumption has a compensating effect in case the computed flow field might be affected by small illumination changes (e.g. an object texture with local ambiguity of intensity, but sufficient gradient information). The complete formulation of the energy with $\mathrm{d}t = 1$, $f_u = \mathrm{d}u$ and $f_v = \mathrm{d}v$ then reads as

$$E(u, v) = E_{data}(u, v) + \alpha E_{smooth}(u, v) \qquad , \alpha > 0, \tag{5.3}$$

where

$$\begin{aligned} E_{data}(u, v) = \int_\Omega \Psi(&|I(u + f_u, v + f_v, t + 1) - I(u, v, t)|^2 \\ &+ \gamma |\nabla I(u + f_u, v + f_v, t + 1) - \nabla I(u, v, t)|^2)\, \mathrm{dx}, \quad \gamma > 0, \\ E_{smooth}(u, v) = \int_\Omega \Psi(&|\nabla f_u|^2 + |\nabla f_v|^2)\, \mathrm{dx}. \end{aligned} \tag{5.4}$$

The computation involves all points $\mathbf{x} = [u, v]^\top$ within the image domain Ω; the function $\Psi(s^2)$ maps the highly convex term s^2 to a moderately convex term that is more robust to outliers. The input images are scaled multiple times down to a minimum resolution (image pyramid) in order to find a good first initial solution of the flow field. This step is important especially for larger motion magnitudes.

5.3 Instruments for Validation: Wind Lidar

The *streamline V4* Doppler wind lidar by Halo Photonics derives the wind component along the direction of a pencil beam based on the measured phase-shift of backscattered laser light. The energetic pulses are sent out with a frequency of 15 kHz and averaged over one second, which results in a temporal resolution of 1.67 seconds due to a data processing time of 0.67 seconds (Schween et al., 2014). Since the molecular scattering of laser-light in clear air is very weak, retrievals rely on the presence of aerosols or cloud droplets, which usually limits its operational range to the planetary boundary layer and low- to mid-level clouds, which offer a sufficiently high particle concentration. Its scanning capability allows to conduct cross-section scans (RHI) and derive radial wind speed up to 9000 m with a range resolution of 30 m. Resolvable velocities range between 0.0038 and 19.2 ms^{-1}.

Under the assumption of a homogeneous wind field, a *Velocity Azimuth Display* scan (VAD) allows to estimate the height-resolved horizontal components of the wind vector $\mathbf{v}(z) = [u(z), v(z), w(z)]^\top$.

During a VAD the pencil beam scans the wind field at a fixed elevation angle $\theta = const$ and 36 azimuth angles ($\varphi = 0°, ..., 350°$), thus forming a cone with the apex at the lidar instrument. Each measurement provides the radial velocity v_r along the beam direction $\mathbf{r} = [x_r, y_r, z_r]^\top$, i.e. the projection of the wind vector $\mathbf{v}$ onto $\mathbf{r}$, according to

$$v_r = -u \sin\varphi \cos\theta - v \cos\varphi \cos\theta - w \sin\theta.$$

Since the wind vector and the elevation are assumed constant, the radial velocity v_r takes the shape of a cosine depending on the azimuth angle φ:

$$v_r = a + b\ \cos(\varphi - \varphi_{max})$$

with amplitude a, offset b and angular phase shift φ_{max}, which can be used to fit the observations. At least three measurements in pairwise perpendicular directions are required to derive the components of the wind vector, which yields

$$u = -b\,\frac{\sin\varphi_{max}}{\cos\theta}, \quad v = -b\,\frac{\cos\varphi_{max}}{\cos\theta}, \quad w = \frac{-a}{\sin\theta}.$$

The 36 radial velocities v_r are used to compute a least-squares solution for $\mathbf{v}$ at each height.

5.4 Evaluation of Cloud Motion Vectors

Evaluation of obtained cloud motion vectors includes situations with cloud layers and a developing cumulus cloud under wind shear. Cloud layers, like any other cloud, are subject to internal and external dynamics, such as evaporation, condensation or turbulence. However, these dynamics may be neglected for small time intervals in the range of 15 or 30 seconds, which allows to interpret the clouds motion as the approximate atmospheric motion at the respective height. Under these circumstances, it is justified to compare the results to observational data from the local wind-lidar. Convective clouds pose a challenge to motion estimation due to their highly dynamic nature. In terms of optical flow, cumulus clouds severely violate the brightness-constancy assumption due to a permanent change of shape. The requirement of a pure translational change of the image signal is thus not satisfied. Hence, a motion estimation can only deliver information about the cloud boundary evolution rather than internal up- or down-drafts.

5.4.1 Methodology

The computation of motion direction must take into consideration that the concepts of mean or standard deviation are meaningless for an angle (the average of π and $-\pi$ represents the opposite direction). Instead of the average angle, the x- and y-component of the direction vector are analyzed independently: given a set of estimated motion directions $\boldsymbol{\varphi} = \{\varphi_1, \varphi_2, \ldots, \varphi_n\}$, the mean direction $\overline{\varphi}$ is defined as

$$\overline{\varphi} = \operatorname{atan2}\left(\overline{y_\varphi}, \overline{x_\varphi}\right), \qquad \text{where} \qquad \begin{aligned} \overline{x_\varphi} &= \frac{1}{n}\sum_{1}^{n} \cos\varphi_i, \\ \overline{y_\varphi} &= \frac{1}{n}\sum_{1}^{n} \sin\varphi_i \end{aligned}$$

and atan2 defined according to Eq. (2.3). Standard deviation and median are computed equivalently. Note that atan2 maps to the interval $\varphi \in \,]-\pi, \pi]$. Conversion between the wind direction from the wind lidar (clockwise from North) and the camera reference system is done according to

$$\varphi_{world} = -(\varphi_{lidar} - \frac{3}{2}\pi), \qquad \varphi_{lidar} = -\varphi_{world} + \frac{3}{2}\pi.$$

For the experimental setup, the optical flow routine by Liu et al. (2009) was applied with an image pyramid scaling set to 0.5 so that images are scaled to half resolution with each additional pyramid scale. The parameter α, which regulates the influence of the smoothing term, was set to a value of 0.008 to reduce the influence of small-scale random variabilities, while avoiding an oversmoothing of the flow field. Also, the input images were down-scaled to half

the original resolution, i.e. 864×864 pixels, before passed to the optical flow routine in order to reduce computational effort.

5.4.2 Estimation of Atmospheric Motion from Cloud Layers

Cloud motion vectors were estimated from a near-zenith area of cloud layers to obtain atmospheric motion direction and velocity. The results were compared to the corresponding observations from a nearby wind-lidar.

The sampling area for the evaluation was chosen large enough to average out small-scale variations of the cloud motion field, e.g. due to a cellular structure or internal dynamics of the cloud layer. An area of 2000 m × 2000 m was sampled; using a smaller sampling area, e.g. 1000 m × 1000 m, leads to similar results, but also to a more fragmented distribution of velocity and direction.

A cloud motion analysis was conducted for a cloud layer at approximately 3000 m height from 5 August 2014 at 15:47 UTC, which exhibited many cases of highly structured alto-cumulus and sporadic cumulus convection with directional wind shear (Figure 5.1).

For most areas of the cloud layer the average motion was found spatially homogeneous with a variability at spatial scales of individual turbulent cells of the cloud. Larger uncertainties could be found primarily at the circumference and in low-contrast areas, where a larger geometric uncertainty can be expected (Figure 5.1a, 1).

In general, motion estimation becomes increasingly uncertain along the motion direction of a cloud: In image areas where the viewing direction is aligned (approximately co-linear) to the motion direction of the cloud, the optical flow magnitude is lower compared to viewing directions perpendicular to cloud motion, because of the different projection of the cloud displacement depending on the view perspective. In these regions, uncertainties of the flow field then can have a larger impact on the estimated velocity (Figure 5.1a, 2).

The analysis shows that the median wind speed differed by less than 1 ms^{-1} with a value of 4.4 ms^{-1} by the cameras and 5.2 ms^{-1} by the lidar (Figure 5.1c). The estimated median motion direction was 265° by the stereo cameras and 258° by the lidar, resulting in a difference of 7° (Figure 5.1d).

An investigation of four case studies between 12:32 UTC and 15:47 UTC confirmed that the estimated median velocity deviated less than 1 ms^{-1} from the observations by the lidar (also median) with standard deviations between 0.45 ms^{-1} and 0.7 ms^{-1} (Figure 5.2). Each velocity estimation, however, showed a negative bias when compared to the lidar observations.

A possible explanation is the persistent negative bias of the stereo reconstruction for this day of about 150 m (compare Figure 4.10), which consequently has a negative effect on the estimation of cloud motion velocity. Assuming a CBH bias of -150 m (5% of 3000 m CBH) and a cloud motion speed of 5 ms^{-1}, the estimated motion must be corrected by +0.25 ms^{-1}, which does not explain the full bias. Another explanation could be internal dynamics of

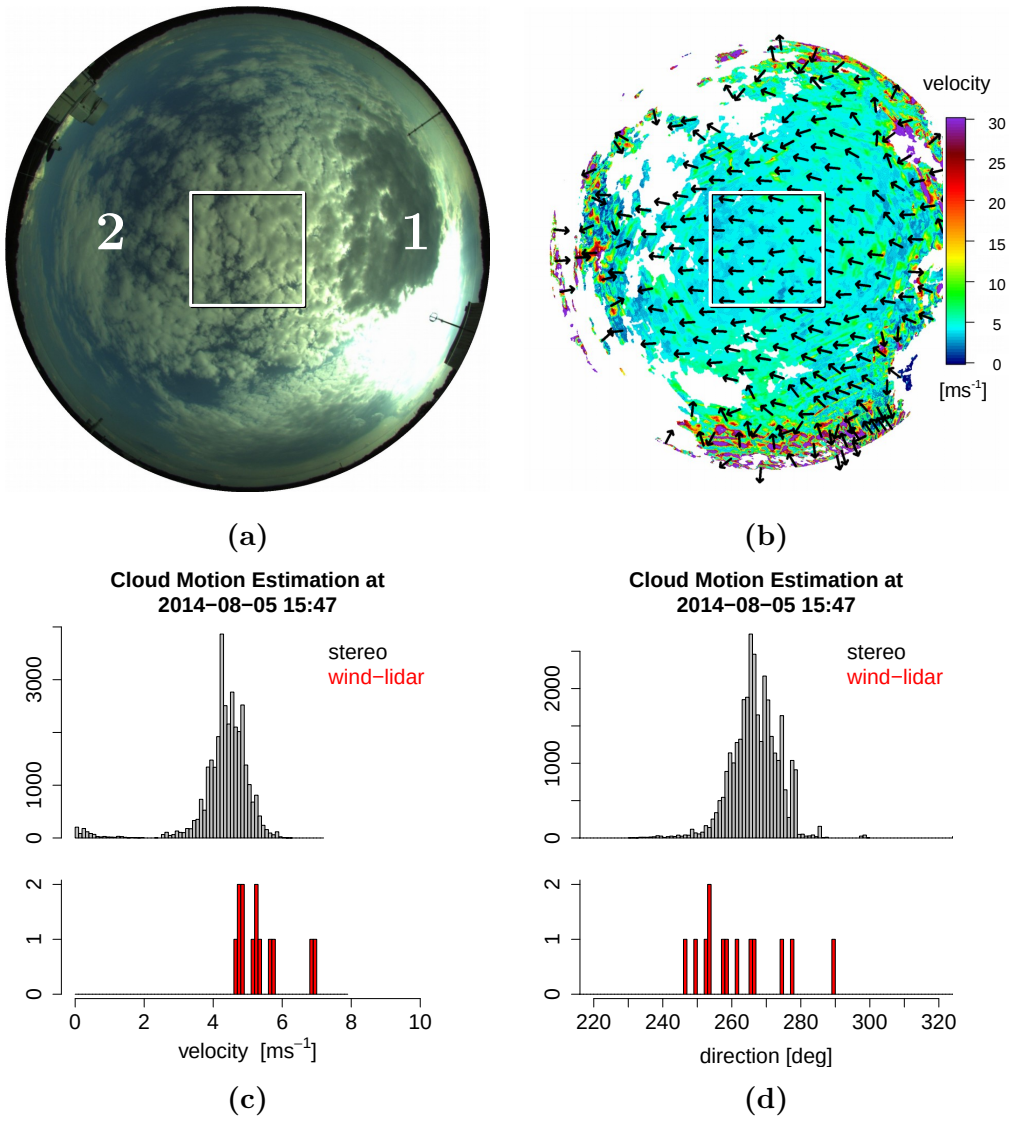
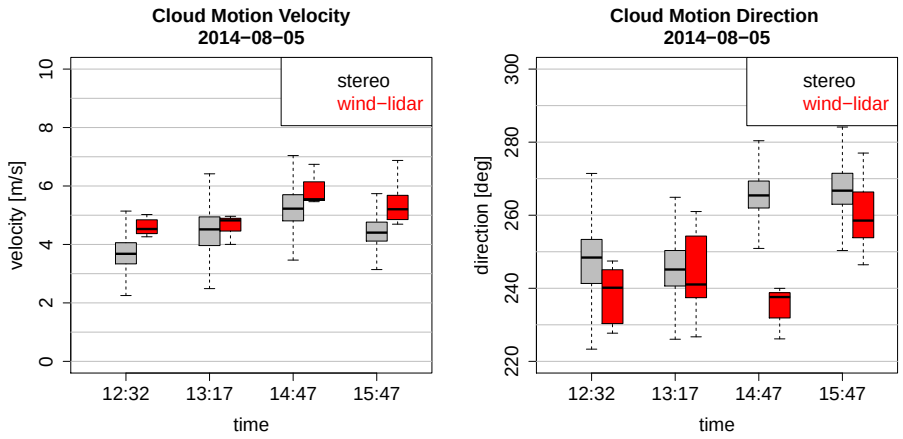


Figure 5.1: Cloud Motion Analysis from 5 August 2014 at 15:47 UTC. Shown is (a) the original fisheye image with marked region (rectangle) for which CMVs are evaluated, (b) a color map showing the horizontal motion velocity and direction for the whole image, (c) histogram of the estimated cloud velocity and (d) direction from the stereo cameras and the wind-lidar respectively.

the layer that affect the optical flow: In addition to general turbulent motion, which might have had a neutral net effect, condensation and evaporation within the cloud layer could have produced a bias in the optical flow to compensate sources and sinks of the image signal. The difference of estimated motion direction by the stereo cameras and the wind-lidar was between 1.29° and 10.06°; only for 14:47 UTC the comparison yields a larger disagreement of 30.8°.



(a) **(b)**

Figure 5.2: Cloud motion analysis from 5 August 2014 for altocumulus clouds and a comparison to the corresponding observations from a wind-lidar. Shown is (a) cloud motion velocity and (b) direction.

Another analysis was conducted for the 20 August 2014, which was dominated by a mixture of alto- and stratocumulus layers and thus offered a larger variability of CBH and cloud motion. While the uncertainty of estimated cloud motion velocity was larger in absolute terms compared to the 5 August 2014, the median velocity again deviated less than 1 ms^{-1}, resulting in a smaller relative error. A bias in motion velocity as for the 5 August 2014 was not detected. The difference between estimated motion direction by the stereo cameras and the wind-lidar was within 7°, which is in agreement with the results from the 5 August 2014.

5.4.3 Estimation of Cumulus Deformation under Wind Shear

This section evaluates derived cloud deformation vectors for cumulus clouds. Convective clouds pose a challenge to the computation of cloud motion vectors due to their highly dynamic nature. The permanent shape evolution often violates the brightness-constancy assumption that optical flow computation relies on (Sec. 5.2). Hence, cloud motion vectors do not necessarily reflect the internal dynamics of the cloud, especially not near the cloud base, where the vertical component of the sub-cloud inflow can hardly be identified visually and geometrically.

Cloud deformation vectors and vertically resolved velocity and direction was derived for a cumulus cloud at 5 August 2014, 10:40:30 UTC, which was developing under vertical wind shear. The cloud motion analysis captured the dynamic evolution of the cloud, which has almost reached its maximum height and is topped by an altocumulus cloud at approximately 2800 m height (Fig-

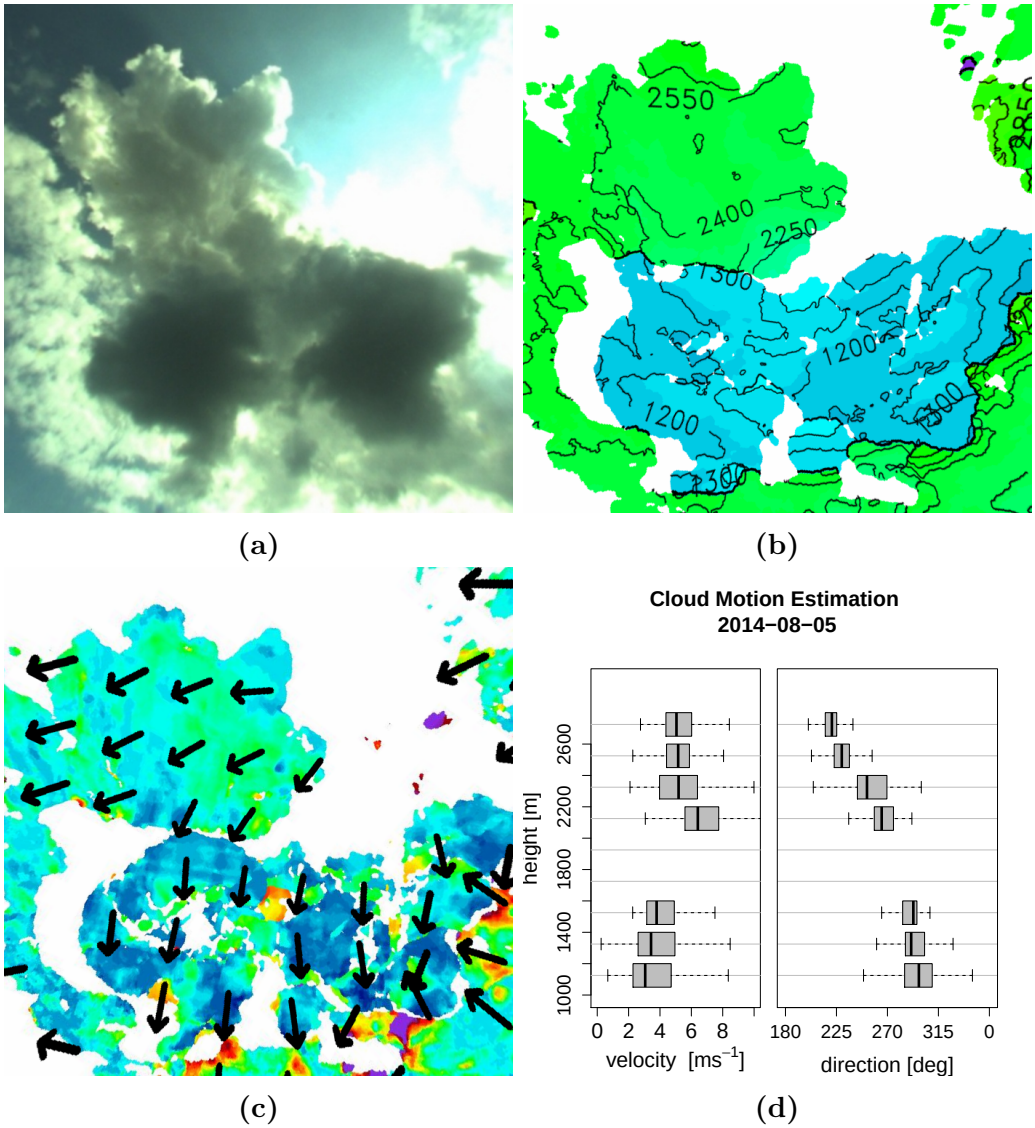


Figure 5.3: Cloud motion analysis of a developing cumulus cloud from 5 August 2014. Detail view of (a) the original fisheye image on the cloud, (b) height profile from the stereo reconstruction, (c) corresponding map of motion magnitude and direction and (d) vertically resolved velocity and direction of the cloud.

ure 5.3). The ascending turret was strongly influenced by the vertical wind shear, which resulted in a significant change of motion direction by approximately 80° between 2000 m height and the cloud top at about 2872 m. The motion velocity, however, increased only slightly with a value of about 4 ms^{-1} near the cloud base and about 5 ms^{-1} at the ascending turret.

The vertical component of the cloud motion vectors was also investigated with the aim to quantify ascent velocity of cloud thermals or detect downward flux of cloudy air as part of the subsiding shell (Heus and Jonker, 2008), which

were observed at the boundary of cumulus clouds. A distinctive relationship, however, was not found, most probably due to the constant change of shape and the small signal-to-noise-ratio of the vertical component.

5.5 Summary

In this chapter, a cloud motion analysis of cloud layers and convective clouds was presented. Sequences of stereo images from a pair of hemispheric sky cameras were analyzed by an optical flow routine, which conducts a dense visual tracking of image features. The combination of tracking information and a dense 3-D cloud reconstruction allows to identify corresponding 3-D cloud features and thus an estimation of cloud motion. The estimated motion vectors of several cases of strato- and altocumulus cloud layers were compared to the observations of a nearby Doppler wind-lidar. It was shown that the estimated velocity deviates by no more than 1 ms^{-1} and the direction by no more than 7° to 10° if compared to the wind-lidar. If applied to convective clouds, the derived motion vectors were accurate enough to reflect macroscopic motion patterns of such clouds, as was shown for a developing cumulus under wind-shear. Such deformation vector fields can be used for an accurate and robust cloud tracking and life-cycle analysis, and could be useful for identifying detrainment in cumulus clouds, i.e. outflow and mixing of cloud air with environmental air. The deformation vectors should, however, interpreted with care, as they usually do not reflect internal dynamics of such clouds.

6 Analysis of Cumulus Life-Cycle and Morphology

This chapter exploits the dense 3-D cloud reconstructions obtained from two pairs of separated hemispheric stereo cameras to estimate the spatial dimension and life-cycle of cumulus clouds. Such statistics are valuable for the validation of cloud populations generated by Large-Eddy-Simulations (LES), which are a major tool to derive cloud population statistics required for parametrizations in large-scale atmospheric models.

Relevant geometric aspects of cumulus clouds and the methodology to derive them from the 3-D stereo reconstructions are introduced in Sec. 6.1. In Sec. 6.2, a moving cumulus cloud evolving under the influence of vertical wind shear is reconstructed and tracked over a period of 30 minutes. The two viewing perspectives allow to reconstruct both the front and back of the highly tilted cloud body, which are then used to assess vertical structure, diameter and tilting angle of the cloud. The second case study is presented in Sec. 6.3, which analyzes the evolution of a cumulus cloud over its whole life-cycle of about 50 minutes. The reconstruction allows to compute geometric statistics, such as cloud base height, thickness, area, perimeter and further derived measures.

6.1 Methodology

6.1.1 Geometric Analysis of Cumulus Clouds

Key geometric properties of cumulus clouds are cloud base height (CBH), cloud top height (CTH), surface-projected area (A) and perimeter (p). Further derivatives are cloud height (h), equivalent diameter (D), linear size (L) and perimeter scaling (s_p), which are explained in the following (Figure 6.1).

These are defined as

$$\begin{aligned} CTH &= \max\{z_X \mid \mathbf{X} \in \Omega\}, & D &= 2\sqrt{\frac{A}{\pi}}, \\ CBH &= \min\{z_X \mid \mathbf{X} \in \Omega\}, & L &= \sqrt{A}, \\ h &= CTH - CBH, & s_p &= \frac{\ln p}{\ln L}, \end{aligned} \tag{6.1}$$

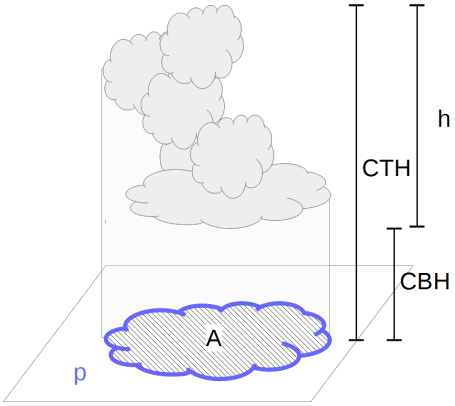


Figure 6.1: Basic geometric measures of a cumulus cloud include its cloud base height (CBH), cloud top height (CTH), height h, surface-projected area A and perimeter p.

where $\mathbf{X}$ denotes a point from the set Ω of reconstructed cloud points.

The often irregular geometry of cumulus clouds deviates from basic Euclidean geometries like circles (spheres) or rectangles (cubes). It is convenient to convert the measured cloud area to its equivalent diameter and linear size. The equivalent diameter D corresponds to a circular cloud with area A. Similarly, the linear size L relates to the edge length of a cubic cloud with area A. The perimeter p is defined as the edge length of the domain occupied by the surface projection of a cloud. In combination with the linear size of the cloud it determines its fractal dimension, i.e. the anomalous scaling of cloud surface with increasing size. The fractal nature of clouds is an important aspect for turbulent mixing, which is assumed to be proportional to the contact surface between two fluids, i.e. the cloud boundary to the gaseous atmosphere. The scaling dimension s_p relates the cloud perimeter to the linear size via

$$p = L^{s_p}.$$

Since clouds with a pronounced vertical dimension produce significant 3-D radiative effects, such as photon leakage through cloud sides or an increased mutual shadowing, the estimation of cloud aspect ratio, i.e. the cloud height in relation to its diameter, is important for radiative transfer of a cloud field. According to (Benner and Curry, 1998; Plank, 1969), the relation between cloud height and equivalent diameter is assumed to follow a power-law relationship:

$$h = a + bD^c.$$

If generalized for cloud populations under different conditions, this relationship shows a significant uncertainty in the estimated parameters. A detailed 3-D cloud reconstruction allows to investigate whether an individual cloud follows the same relationship over its life-cycle.

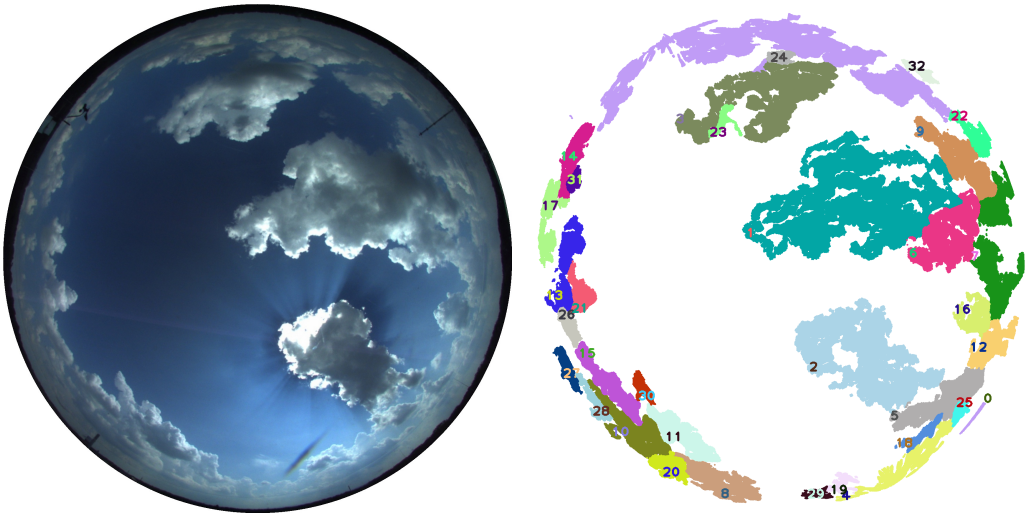


Figure 6.2: Extraction of individual clouds is done via segmentation of a 3-D point cloud. Shown is an example result of a segmentation for 13 May 2016. Each color defines one detected cluster.

6.1.2 Data Processing

A geometric analysis of individual clouds requires processing and analyzing a large, unorganized set of 3-D object points or oriented patches from the 3-D reconstruction. The identification of individual clouds can be challenging as clouds are often subject to splitting and merging events throughout their life-cycle: Buoyant parcels may constantly reach the lifting condensation level (LCL) near an existing cloud and form new cloud, which leads to an ensemble of neighboring clouds that are often hard to distinguish. Also, a large portion of a cloud will remain unresolved if reconstructed from only one view perspective. Hence, a comprehensive statistical life-cycle analysis should concentrate on clouds that are observed by at least two camera pairs from different views so that each pair contributes complementary geometric information w.r.t. vertical and horizontal extent, and morphology.

The identification of individual clouds requires a segmentation of the 3-D point cloud into subsets or clusters that are assumed to represent distinct clouds. This can be accomplished by image-based segmentation or by analyzing the 3-D data directly using a clustering algorithm. In general, image-based segmentation is not recommended because clouds often show a wide range of contrast, texture and brightness. Since mutual occlusion occurs frequently within the cloud field, using a cloud mask for segmentation, as is done by Crispel and Roberts (2017), is usually limited to clouds at near zenith and situations with distinct clouds.

In the following analysis, a Euclidean clustering algorithm is applied for cloud segmentation, which is implemented by the open-source *Point Cloud Library* (Library, 2016; Rusu, 2009). The algorithm analyzes an unorganized

set of 3-D points and returns a set of clusters based on a minimum distance between two clusters. A distance threshold r_{thresh} determines whether a point is added to an existing cluster or defines an own cluster. A minimum and maximum size of clusters can be defined to remove clutter and clusters that are too small to be of interest. Since the segmentation relies only on Euclidean distance it is sensitive to outliers or clusters that are located close to each other, which is usually the case during splitting and merging of cumulus clouds or when detrained cloud condensate approaches a neighboring cloud. In such cases, the segmentation results in fewer and larger clusters that do not exactly reflect the real cloud field, and thus affect cloud population and life-cycle statistics. In many cases, however, the clustering is able to identify individual clouds correctly (Figure 6.2), allowing to derive cloud top height, cloud base height and location directly.

While CBH and CTH can be calculated directly from a given point cloud, further derived measures, such as cloud area or perimeter, are sensitive to missing data points, as caused by occlusions, sharp surface gradients that cannot be resolved by the template matching, or invalid disparities due to poor contrast. As a result, the surface-projected cloud area yields holes and rifts that directly affect the computation of cloud area and perimeter. Therefore a cloud boundary estimation is desired which approximates the given projection, while satisfying a certain degree of geometric smoothness.

An implementation of the *active contours* technique is used (Xu and Prince, 1998) to estimate a cloud boundary: Starting from an initial boundary enclosing the projected data points, the contour is iteratively adjusted to minimize an energy functional, which penalizes discrepancies to the data points and a too strong boundary curvature. The routine operates on a binary image that provides a regular 2-D raster for the projected point cloud. The chosen resolution of the raster has a strong influence on the final contour, since a finer resolution allows a (potentially) more realistic contour, while a coarse resolution tends to increase the cloud area and decrease the area-perimeter-ratio as fine scale structures disappear. For the analysis, a resolution of 20 meters per pixel was chosen, which is similar to the grid resolution of current LES. Once a closed cloud boundary is available, the computation of cloud area, perimeter and location (centroid) is done by computational geometry routines, which generally make use of the Green formula.

6.2 Estimating Cumulus Vertical Structure under Wind Shear

This section analyses cumulus morphology reconstruction from two view perspectives with a focus on tilted clouds. Vertical wind shear often results in highly tilted cloud shapes. Simulations suggest that the higher effective cloud fraction of tilted clouds could result in a decreased domain averaged solar flux

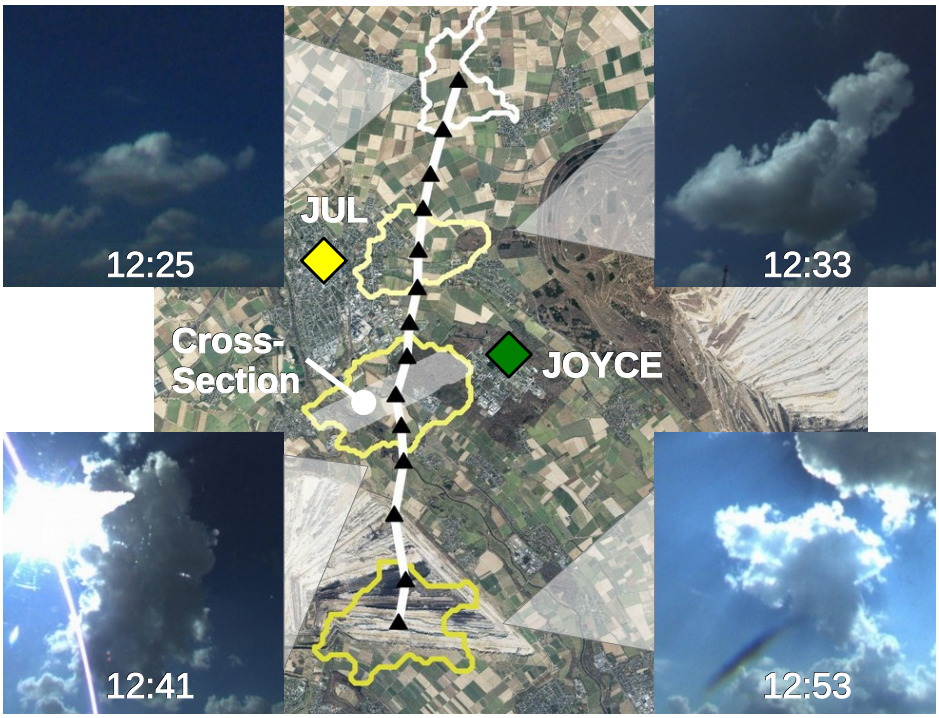


Figure 6.3: Overview of tracked cumulus cloud from 13 May 2016 between 12:25 UTC and 12:53 UTC with marked location, perimeter and path of the cloud, which was determined by the centroid of the surface-projected cloud area. Map data: *Land NRW (2017) Datenlizenz Deutschland - Namensnennung -Version 2.0 (www.govdata.de/dl-de/by-2-0).*

of up to -39.9 Wm^{-1}, depending on the solar zenith angle and tilting direction (Hinkelman et al., 2007). Vertically pointing instruments like lidars or radars in soda-straw scanning mode have a low chance to observe a representative slice of tilted clouds and thus their true height, size and shape. Statistics of cumulus populations obtained from space- or airborne photography derive the size of individual clouds by means of the surface-projected cloud area (Joseph and Cahalan, 1990; Plank, 1969). The tilted body of a cloud under wind shear thus increases the surface-projected cloud area, leading to a decreased cloud aspect ratio and a different cloud size distribution as used for the parametrization of convection in GCMs. The following case study analyzes the first of a sequence of developing clouds from 13 May 2016, which were monitored by the stereo cameras at JOYCE and JUL, and allowed a simultaneous reconstruction from two different view perspectives for over 30 minutes.

The cloud formed at around 12:20 UTC approximately 10 km north of JOYCE and moved south. Except at 12:49 UTC, when the view was severely deteriorated by the sun, the cloud could be identified, tracked and geometrically analyzed (Figure 6.3).

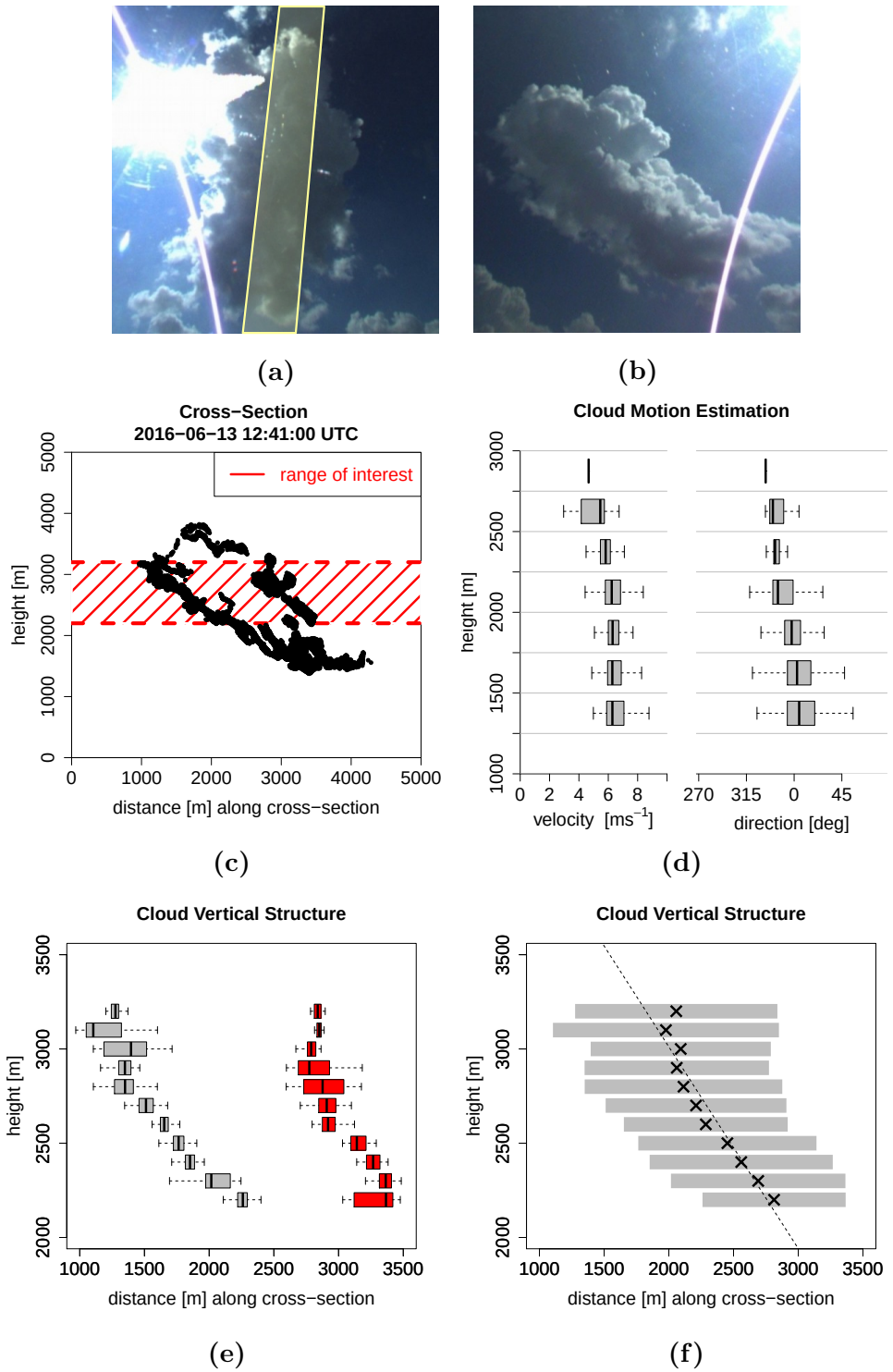


Figure 6.4: Analysis of cumulus vertical structure from 13 May 2016 at 12:41 UTC: The top row shows the close-up view from (a) JOYCE and (b) JUL. The cross-section marked in (a) is plotted in (c) and a cloud motion analysis is given in (d). The cross-section in (c) is used to estimate (e) front (gray) and back (red) of the cloud by splitting the data and computing the respective median distance along the cross-section. Cloud overlap in (f) is then represented by the tilting angle of the cloud derived from the horizontal shift of the respective layers.

Cloud development was accompanied by a single rising plume at 12:27 dragged eastward by wind shear. Its motion path - defined by the centroids of the surface-projected cloud area - was slightly bent eastwards, most likely caused by a combination of wind shear and cloud growth followed by a slowly decaying updraft. This assumption is supported by the vertical profile of estimated cloud motion vectors at 12:39 UTC, which indicates that the motion direction at the cloud base differed by approximately 30° to that of the cloud top (Figure 6.4d).

Each stereo camera provided complementary geometric information about the cloud: While the tilted cloud top was monitored by JUL, stereo camera JOYCE was able to capture the cloud base and the horizontal extent of the cloud. The combination of both stereo reconstructions at 12:39 UTC allowed a reasonable estimation of cloud geometric statistics, such as CBH(1500 m), CTH (3834 m) or equivalent diameter (2420 m).

A cross-section along the clouds tilting direction revealed a highly skewed cloud geometry (Figure 6.4c). The analysis of the vertically resolved cross-sections (provided in intervals of 100 m) included reconstruction data within a 600 m wide stripe and was limited to heights for which reliable and complete reconstructions for both cloud top and base were available. While the total cross-section of the cloud was 2271 m, the vertically resolved cross-sections increased from about 1100 m at 2200 m height to 1750 m at 3100 m height (Figure 6.4f). Cloud vertical overlap was investigated via the clouds tilting angle derived from the relative horizontal shifts of adjacent cross-section layers. One can observe the decreasing overlap up to a height of 2800 m, where the cloud was significantly deformed by the wind shear. Above 2800 m vertical overlap increased due to the convective structure near the cloud top. A line fit through the centroids of each layer yielded a tilting angle of about 46° (Figure 6.4f), which was, however, significantly influenced by the convective cloud top.

In this case study, the four-camera stereo setup allowed to capture the principal geometry of the cloud for a detailed geometric shape analysis. A dense spatial coverage from both perspectives was sometimes impeded by small cloud fractions blocking the view, but also by the size of the cloud in relation to the distance between cameras JUL and JOYCE. Since the path of the cloud examined here was almost parallel to the baselines of both stereo cameras, the reconstruction precision was generally lower compared to a path along the line connecting JUL and JOYCE (Sec. 3.6), which would have allowed to estimate cloud perimeter and surface-projected area with highest precision (compare Figure 3.8). The larger baseline of about 700 m, however, also lead to a better geometric conditioning compared to the stereo setup with a 300 m baseline.

6.3 Analysis of Cumulus Life-Cycle

A cumulus life-cycle analysis from 20 August 2015 is presented as a show case example using stereo images from the camera pairs at JOYCE and RWE. The different view perspectives allowed a survey of the base and one side of the cloud, providing information about cloud area, perimeter, diameter and cloud height during its life-cycle. The evolving cloud was located between the camera pairs approximately 4 km South-East of JOYCE over a timespan of 50 minutes, from 13:10 UTC to 14:00 UTC.

Cloud formation started at about 13:10 UTC when thermals reached the LCL at 1436 m and ascended up to 1861 m within the first 15 minutes. At this stage, cloud development could be characterized as forced (compare Stull, 1988) as no significant vertical updraft could was observed. Several buoyant parcels reached the LCL which resulted in a constant merging of newly created cloud condensate and a steady increase of equivalent diameter from 1031 m at 13:10 UTC to 1545 m at 13:25 UTC.

At 13:26 UTC cloud development became more active as several ascending thermals of different size and height can be observed. At 13:30 UTC a larger thermal became visible, almost 20 minutes after initial cloud formation, and reached its terminal height of 2990 m at 13:39 UTC, which is also the maximum cloud top height. A larger secondary thermal started its ascent at around 13:33 UTC, accelerated quickly and reached its terminal height of 2750 m at 13:41 UTC.

Observations and numerical simulations suggest that the typical cumulus life-cycle consists of a series of pulses or active thermals rather than a single convective event or a constant plume. According to Zhao and Austin (2005), the first thermal tends to determine the maximum height of the cloud during its full life-cycle; parcels that start their ascent after this primary convective event only reach a lower terminal height. Heus (2008), however, notes that this is not always the case. The height evolution of the cloud examined here is largely marked by a single convective turret until 13:39 UTC, and seems to support the general case stated by Zhao and Austin (2005). The first active thermal, however, appeared at 13.26 UTC with a significantly lower height than the maximum of 2990 m. The principal phenomenon of a sharp increase of CTH and a gradual decay could be observed here.

The shape of the cloud was influenced by moderate vertical wind shear, which exerted a drag on ascending thermals and separated them from the main cloud (Figure 6.7). These splitting events - at 13:39 UTC for the first thermal and at 13:41 UTC for the second one - are visible as sudden drops in the time series of cloud height and area (Figure 6.5). Since thermals kept emerging close by, the cloud was not significantly changing its location during its entire life-cycle.

The aspect ratio of the cloud ranged between 0.25 and 0.77 with an average of 0.35 from 13:10 UTC to 13:25 UTC and 0.5 from 13:26 UTC to 13:50 UTC.

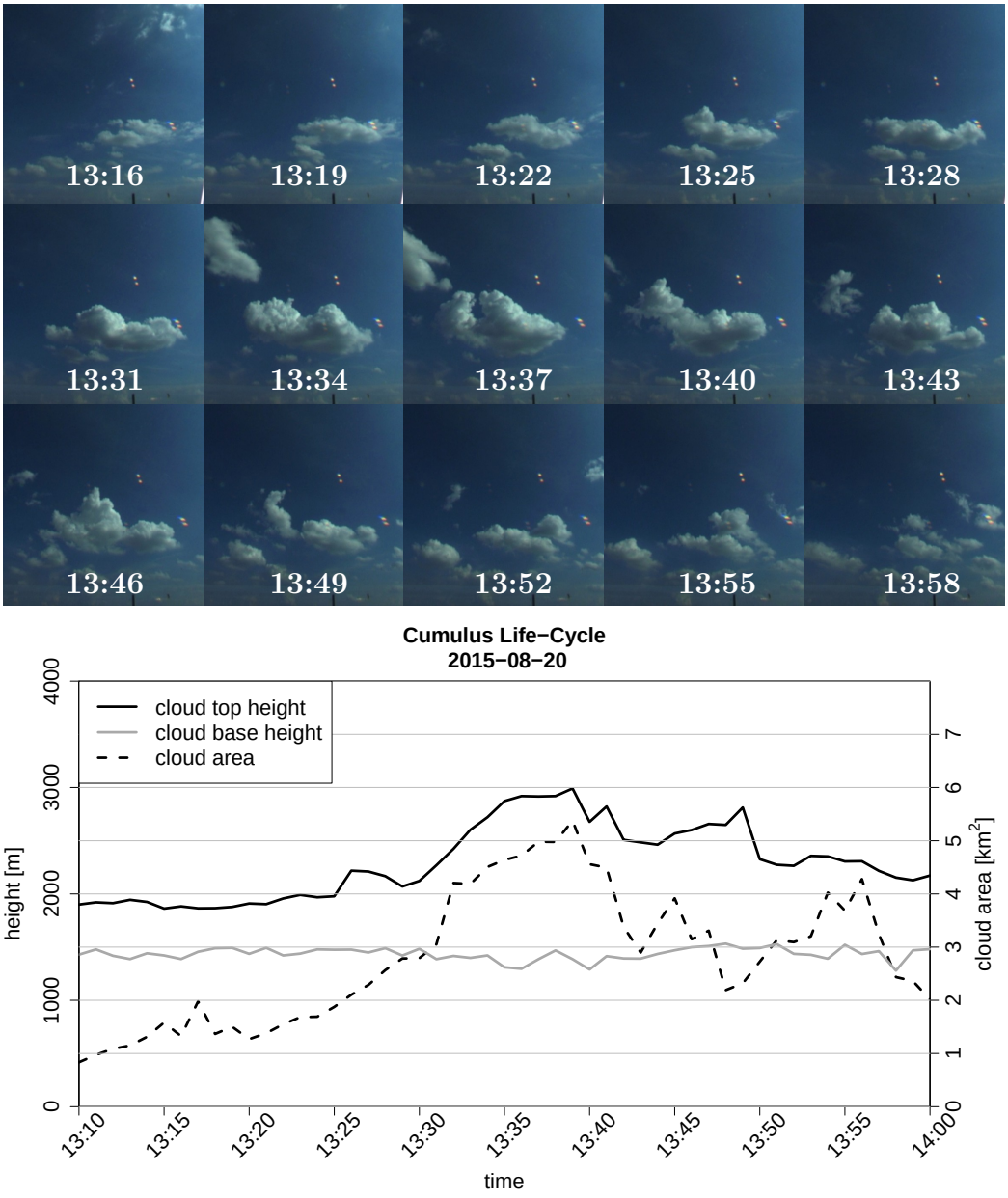
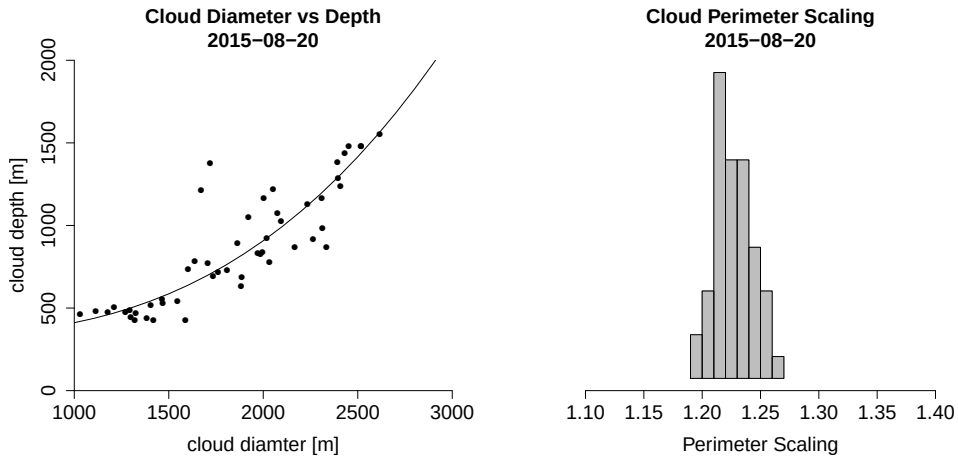


Figure 6.5: Life-cycle analysis of cumulus cloud from 20 August 2015. Shown is a sequence of detail views from JOYCE (top) and a summary of geometric statistics including CBH, CTHand cloud area (bottom).

The data also seems to support the observation of Plank (1969) that the aspect ratio follows a power-law functional relationship (Figure 6.6). The dataset contains a few outliers, which are most probably linked to the transition phase between individual convective pulse events.

The close-up view of stereo camera RWE allowed to estimate the perimeter of the clouds' vertical projection as seen from a satellite (Figure 6.7), which



(a) **(b)**

Figure 6.6: Cumulus life-cycle analysis from 20 August 2015: cloud aspect ratio (equivalent diameter vs cloud height) and fitted power-law function (a), perimeter scaling (b).

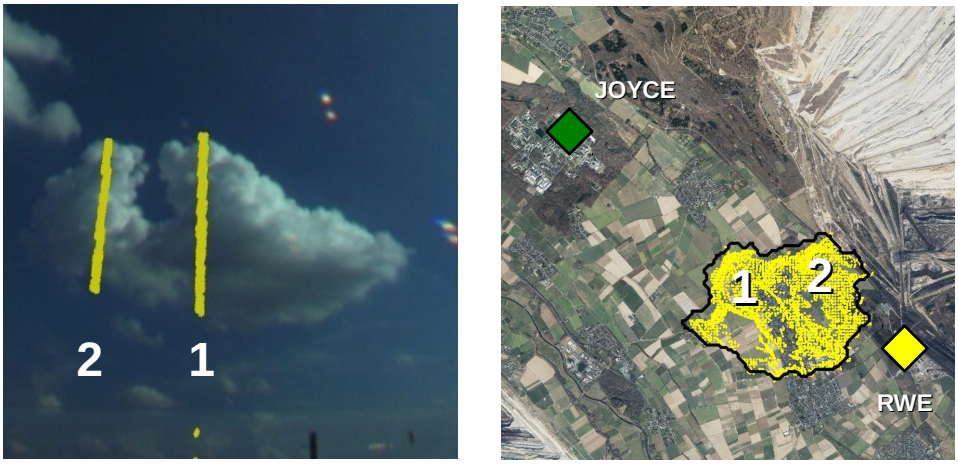


Figure 6.7: Location and horizontal extent of cumulus cloud at 13:37 UTC. **Left**: Detail view of original fisheye image from JOYCE. **Right**: Vertical projection and perimeter of the segmented cloud reconstruction and its location between JOYCE and RWE. The numbers (1) and (2) indicate the location of the convective turrets. Map data: *Land NRW (2017) Datenlizenz Deutschland - Namensnennung -Version 2.0 (www.govdata.de/dl-de/by-2-0).*

yielded a perimeter scaling of 1.23 with a standard deviation of 0.15. This value is smaller than those reported by Siebesma and Jonker (2000) (1.32) and Cahalan and Joseph (1989) (1.33, for clouds smaller than 0.5 km in diameter). A possible explanation for this difference could be the active contours algorithm used to compute the perimeter, as it tends to produce slightly smoothed

contours which then lead to a smaller perimeter.

6.4 Summary

In this chapter, the potential of ground-based stereo cameras for studies of cloud evolution and morphology was investigated. A dense stereo analysis was applied to the image data of two spatially separated pairs of hemispheric cameras. The obtained geometries were used to segment and track a cloud during its life-cycle and estimate its vertical structure and dimension. The case of a highly tilted cumulus cloud was investigated and its vertical overlap analyzed via its tilting angle of more than 46°. In a future analysis, such results could be valuable for a comparison to respective statistics derived from LES (Corbetta et al., 2015) or the sensed reflectivity profiles of cloud radars if available(Oue et al., 2016).

A geometric life-cycle analysis was conducted on a developing cumulus cloud observed from two separated camera pairs. The result was a time series of key geometric aspects, such as cloud top height, equivalent diameter, perimeter and further derived measures like aspect ratio or perimeter scaling. The results indicate that during a clouds' life-cycle the aspect ratio follows a power-law functional relationship. A future application to a larger number of cumuli might include an extended set of distributed stereo cameras and more sophisticated cloud segmentation tools (compare Chapter 7). Also, the analysis could be extended to estimate the nearest neighbor spacing of clouds (Joseph and Cahalan, 1990), which is still subject of current research.

7 Conclusion and Outlook

7.1 Conclusions

In this thesis, the 3-D stereo reconstruction of clouds with ground-based hemispheric sky cameras was investigated to evaluate its suitability for studies of cloud evolution, LES validation and radiative transfer under cloudy conditions. The experimental setup included a large-scale stereo configuration of sky cameras to derive 3-D cloud boundaries for large parts of the local cloud field and to monitor the life-cycle of individual clouds from different view perspectives. State-of-the-art stereo analysis tools were used to analyze the collected image data showing various cloud types and also multi-layer cloud situations. The obtained 3-D cloud boundaries were compared to the corresponding observations from a nearby lidar-ceilometer and a cloud radar to evaluate the accuracy and reliability of such a camera system. The computation of optical flow from sequences of stereo images allowed to derive cloud motion vectors and characterize macroscopic cloud motion and dynamics. The technique was then used to conduct a life-cycle analysis of individual clouds and it was shown how the technique can be applied to derive relevant geometric information for LES validation and cloud evolution studies.

For the experimental camera setup, two and later four sky cameras were assembled and composed to form two spatially separated stereo cameras, each providing complementary geometric data about the local cloud field. Calibration of the internal camera projection parameters was accomplished by a calibration software from the Department for Photogrammetry and Geodesy of the University of Bonn, which was designed to handle fisheye projections. An important aspect of such a large-scale stereo application lies in the estimation of the absolute and relative orientation of two cameras.

A sufficiently accurate absolute orientation for each camera was achieved by recording night sky images and use stars as reference points. The relative orientation between a camera pair, which is crucial for an accurate stereo matching and subsequent reconstruction, was refined using a large number of distinctive cloud feature correspondences, such as convective towers, cloud fractions or the cellular structure of an altocumulus layer (Sec. 3.2).

Before a stereo analysis was conducted, the input images were preprocessed by an epipolar image rectification designed for fisheye projections, which was used to transform the original stereo images in such a way that they reflected an optimal camera configuration, i.e. both cameras having equal orientation, separated along the x-axis of the reference camera and having equal projec-

tion parameters (Sec. 3.3). This allowed to apply state-of-the-art dense stereo matching algorithms that seek an image correspondence for each image pixel rather than only a sparse set of distinctive image features. The obtained 3-D cloud models showed a high degree of geometric completeness and consistence, which was advantageous for segmentation and tracking of individual clouds (compare Sec. 4.3 and Sec. 6.2). A method inter-comparison between the dense stereo matching algorithm using epipolar rectified images and a high-accuracy multi-view reconstruction algorithm that operates on the original, unrectified images showed differences of mostly below 70 m for a cumulus cloud 3 km distant (Sec. 4.3).

In Chapter 4 the cloud reconstructions were compared to observations from co-located lidar-ceilometer and a 35-GHz cloud radar. It was shown that the method is able to reconstruct large parts of the local cloud field and the principal geometry of clouds.

Average estimated cloud base height often showed good agreement to the lidar-derived values with a difference bias of mostly below 1 - 2% of CBH, but also a few occasions with a significant bias of 3-5% of CBH (Sec. 4.5). Standard deviation of geometric difference was found smaller for well-structured clouds like altocumulus than for heterogeneous stratocumulus or cumulus clouds, yielding values of about 50 m (1.5% of CBH) and 100 to 150 m (7 - 10 % of CBH) respectively.

While a validation of zenith CBH was done in previous studies, a comparison to the reflectivity profile by a cloud radar was not conducted before. The comparison of cross-section profiles between the stereo cameras and the sensed reflectivity profiles by a cloud radar showed good agreement for clouds crossing at zenith with average differences well below 100 m (Sec. 4.4). Larger inconsistencies were observed for more distant clouds (3-5 km distance), for which the decreasing sensitivity of the radar is a likely explanation. While the stereo method suffers from a higher geometric uncertainty in low-contrast areas, such as the cloud base, the analysis also indicates that the cameras are better able to detect smaller cloud patches, edges and parts of the cloud base than the radar due to the higher sensitivity in the visible spectrum.

The camera system can be used to capture cloud motion by tracking cloud features over sequences of images (Sec. 5). For cloud layers, which are geometrically relatively homogeneous and show little deformation, the estimated velocity deviated less than 1 ms^{-1} from the observations by a nearby wind-lidar, while motion direction was accurate within 7° to 10°. A cloud motion analysis also allowed to detect and quantify macroscopic deformation patterns and distinguish areas of different dynamics, such as changes of direction due to vertical wind-shear (Sec. 5.4.3). Since the method is limited to the visual information given by a cloud, it is generally not suited to quantify its internal dynamics.

This thesis was motivated by the potential of 3-D stereo reconstruction to provide cloud envelope reconstructions suitable for LES validation and cloud

evolution studies. In Sec. 6.2 a developing cumulus cloud was reconstructed and tracked by two spatially separated camera pairs to analyze the 3-D geometry of its highly tilted cloud body. The tilt was caused by moderate vertical wind shear, which could be confirmed by the estimated cloud motion vectors. The observation from two different view perspectives allowed to estimate the clouds' height-resolved diameter, which was used to compute the tilting angle of the cloud. In Sec. 6.3, the technique was used to monitor the evolution of a developing cumulus and derive key geometric statistics, such as effective diameter, perimeter, CBH and cloud top height (CTH), during its entire life-cycle.

7.2 Outlook

The results presented in this thesis show that stereo reconstruction can provide access to key geometric information about cumulus clouds. In a next step, the method should be applied to a statistically meaningful number of clouds and atmospheric conditions. Also, the analysis should be extended to cloud field statistics, which requires a larger camera network and some improvements to the 3-D reconstruction and analysis methods. Also, there are further applications that could benefit from such a stereo camera system, but were not part of this thesis.

Extended Cloud Field Analysis

Since hemispheric sky imagers allow to monitor the local cloud field, a further analysis should include the spatial organization of cumulus cloud populations and its development over the diurnal cycle. Besides parameters like diameter, cloud base height, cloud top height and perimeter (compare Chapter 6), the cloud horizontal spacing (*nearest neighbor spacing*, Joseph and Cahalan, 1990) is an important aspect that affects the radiative impact of a cloud field and is subject of current research (Figure 7.1). Such an analysis relies on a largely automatized collection of image data and subsequent 3-D reconstruction, which can be achieved with the methodology used in this thesis. A subsequent analysis, however, would benefit from a larger camera network and an improved post-processing.

Extending the Camera Setup

The number of stereo cameras and their spatial configuration determine the capabilities of a ground-based stereo camera system: A larger network covers a greater area, which is important for statistical reasons, and also allows to observe individual clouds from multiple view perspectives, each providing complementary 3-D data.

The current network consists of four operational cameras organized in two stereo pairs (JOYCE and JUL). The network covers an area of approximately

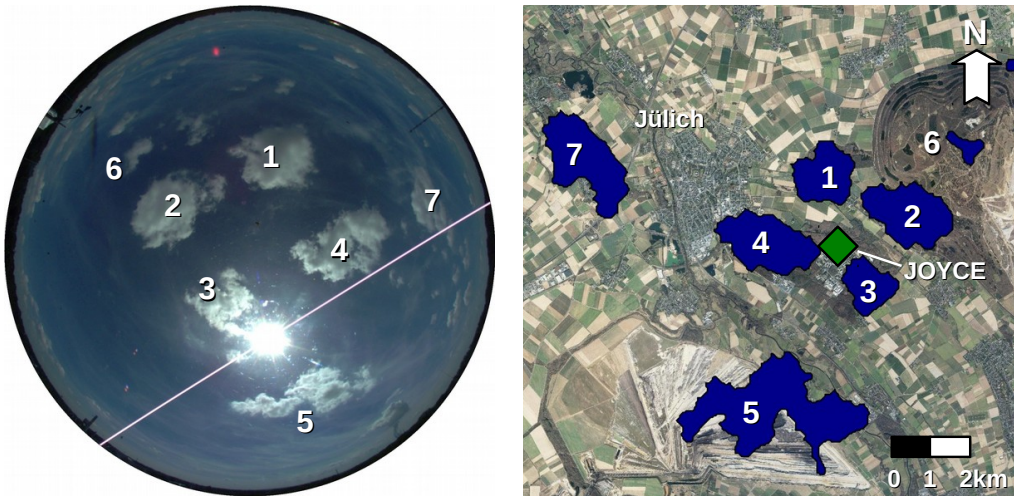


Figure 7.1: An operational stereo camera network at JOYCE can be used to derive cloud field statistics, such as spatial organization and spacing. **Left**: Original fisheye image. **Right**: A segmented and georeferenced cloud field from 31 July 2015. Number index relates to the respective cloud. Map data: *Land NRW (2017) Datenlizenz Deutschland - Namensnennung -Version 2.0 (www.govdata.de/dl-de/by-2-0).*

15×10 km^2, since each stereo camera provides reasonable reconstructions up to a distance of 5 km and both stereo cameras overlap by 50%. An extended network may consist of additional stereo cameras equally distributed around and oriented towards (w.r.t. geometric precision) the JOYCE supersite to achieve a quasi-complete 3-D reconstruction of clouds near JOYCE and a high degree of synergy among the available instruments (Figure 7.2). The stereo camera at JOYCE could be extended to a 3-camera setup, for example in a right-angled configuration in order to guarantee maximum reconstruction precision in all directions (compare Sec. 3.6). For clouds that are not in the vicinity of JOYCE, a partial reconstruction by the network still allows to derive cloud diameter, perimeter, height as well as spatial organization (e.g. nearest neighbor spacing) of the cloud field within an area of at least 20×20 km^2. Also, clouds crossing the site from arbitrary directions could be tracked more reliably over the course of their life-cycle. Note, however, that finding suitable (long-term) locations can be challenging because each has to satisfy certain requirements, such as a baseline length between 300 and 500 meters, an alignment towards JOYCE and a distance of about 5 km, but also installation permissions, power supply and an unblocked field of view (Sec. 4.1.1).

Segmentation and Tracking

Currently, the segmentation and tracking of individual clouds is still time-consuming and requires manual user intervention. Common problems are the

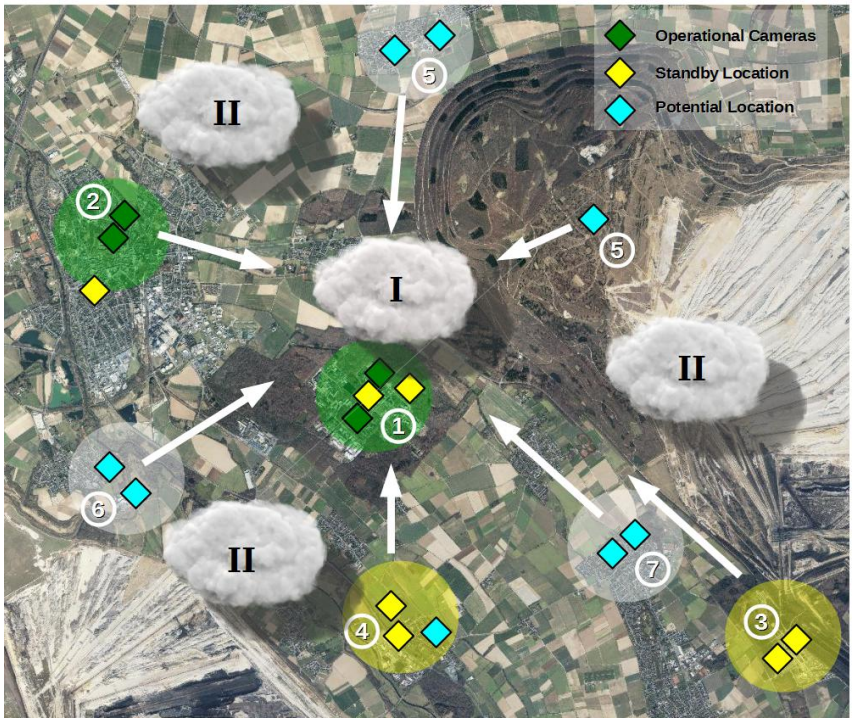


Figure 7.2: Potential future stereo camera network: Each marker represents a single sky camera; two devices can form one stereo camera setup (color-filled disks). Operational cameras are highlighted in green and include the location JOYCE (①) and JUL (②), while yellow markers represent former camera locations that could be reactivated for future use. Cyan markers indicate potential camera locations that have not been examined yet, but seem suited. Each stereo camera, except of camera ①, is aligned towards JOYCE. Hence, clouds near the center (marked as **I**) could be fully reconstructed from all view-perspectives, while other clouds (marked as **II**) could be assessed in terms of their spatial extent (equivalent diameter, height) and distance to other clouds (cloud spacing). The large area covered by the network would allow to track individual clouds over their entire life-cycle. Map data: *Land NRW (2017) Datenlizenz Deutschland - Namensnennung -Version 2.0 (www.govdata.de/dl-de/by-2-0).*

splitting and merging of individual clouds, a limited geometric completeness of the cloud reconstruction compared to its size, occlusions or geometric irregularity. An extension of the camera network would add complementary geometric information that would also make segmentation and tracking more reliable. While a simple Euclidean segmentation - as used in this thesis - or a cluster-based approach can deliver reasonable results for cloud fields with clearly separable clouds, they are not very reliable if these conditions are not met. Also, obtained cloud motion vectors allow monitoring the evolution of clouds and could be used for a Lagrangian analysis similar to Kuhn et al.

(2012) to provide a 4-D cloud life-cycle information.

Refining the Stereo Analysis

In this thesis, stereo analysis relied primarily on the semi-global block matching algorithm (SGBM) implemented in the *OpenCV* computer vision library, which offers high geometric completeness and sufficient accuracy while being computationally very efficient. A typical problem of dense stereo matching algorithms, however, are over-smoothed object boundaries when using a fixed template window geometry to compute the pixel-wise matching cost. An adaptive template approach (Kanade and Okutomi, 1991; Stein et al., 2006) based on the intensity gradient and distortion produced by the epipolar rectification could better exploit high-frequency features in the image signal, such as the cellular structure at the cloud top or the edges of a convective turret. Michael et al. (2013) proposed an adaptive weighting of scan-lines during matching cost accumulation depending on their orientation and achieved a more consistent disparity estimation in image areas with a sharp depth gradient, i.e. an object surface orientation perpendicular to the viewing direction in the image. Cloud reconstruction would benefit from an improved result at the base of distant clouds, which is currently still subject to missing and flawed disparity values.

Multi-View Cloud Reconstruction

While a large, distributed camera network allows to reconstruct different parts of a cloud from multiple view perspectives independently, the combined reconstruction should be geometrically consistent. A degrading precision of stereo reconstruction or geometric artifacts caused by a flawed stereo analysis, e.g. due to regularization or visual deteriorations, can produce detectable inconsistencies and affect cloud segmentation and derived statistics, such as cloud diameter or perimeter.

In many cases, the application of multi-view reconstruction techniques could reduce such geometric inconsistencies and artifacts by computing a 3-D cloud model, which is consistent with the image signal of each camera rather than only a single image pair (stereo). Within a distributed camera network, the cloud mask, i.e. the image classification into *cloud* and *sky* areas, provides for each camera the cloud silhouettes. The set of all 3-D points that project to a *cloudy* pixel on each cloudmask then defines the *Visual Hull* (Kutulakos and Seitz, 1999), which can be used to filter out invalid 3-D object points or restrict the range of valid disparity values during stereo analysis (Li et al., 2002; Fan and Ferrie, 2006).

Feature-based multi-view reconstruction algorithms like PMVS (Sec. 3.5.3), could be used to improve the geometric consistency and detail of convective turrets. It is, however, unlikely that the full potential of such techniques could be exploited in this application, as they work best if an object point is visible in several images at the same time. This only applies to a small fraction of

visible cloud boundaries because the number of cameras is likely limited to a few or maybe a dozen, distributed within an area of about 20 km × 20 km.

Further Potential Applications

The motivation of this thesis was to explore the potential of a ground-based, distributed system of stereo cameras for LES validation, studies of radiation closure and cloud evolution. Its area of application is, however, not restricted to these. The technique delivers information about the 3-D cloud morphology in most atmospheric situations and allows to study cloud related phenomena in general.

The investigation and forecast of deep convection (*convective initiation*) could benefit from a high-resolution ground-based survey of cumulus clouds. Size and growth rate of individual clouds, but also the local wind field are important parameters in current forecast schemes (Walker et al., 2012). These and other parameters are currently obtained from imagers on-board geostationary satellites like *Meteosat* and are affected by the limited spatial ($\approx$ 1 km) and temporal ($\approx$ 5 minutes) resolution, cloud tracking inaccuracies and distorted cloud top height estimations due to advected cirrus clouds. An adequately configured stereo camera system allows to observe cloud formation within intervals of seconds and could provide useful 3-D data, especially in the early stages of deep convection.

As renewable energy becomes increasingly important, solar energy production is significantly affected by the small-scale variability of solar irradiance caused by the local cloud field. Approaches exist that use a hemispheric sky camera for a local short-term forecast of solar irradiation (Schmidt et al., 2016). Significant errors can be attributed to missing 3-D information about the local cloud field, including cloud base height, cloud vertical extent and cloud motion vectors, which a dense stereo analysis could deliver. The proposed technique is cost-effective, fast enough for real-time analysis and can be fully automated. An operational application should, however, consider an optimization of the cloud segmentation and tracking routine.

The cloud radar is a common instrument to derive cloud fraction profiles (Oue et al., 2016), cloud overlap (Corbetta et al., 2015) or cloud vertical structure in general. Simulations (Fielding et al., 2013) and the comparisons provided in this thesis have shown that the radar sensitivity can be a significant problem if used to derive cumulus population statistics for global climate models. The current camera network at Jülich allows to further investigate the accuracy and representativeness of cloud radar reflectivity profiles.

Mathematical Notation

Symbol	Type	Description
$\mathbf{\Omega_{\{W,C,I,S\}}}$	basis ($\mathbb{R}^3$)	(world/camera/image/sensor) coordinate system
$\mathbf{X}, \mathbf{X_C}$	3-D vector	representation of an object point within the world/camera coordinate system
$\mathbf{x}$	2-D vector	representation of a point on the sensor plane in pixel coordinates
$\mathbf{x}'$	2-D vector	representation of a point on the image plane in normalized image coordinates
$\mathbf{x}''$	3-D vector	unit vector representing the projection ray related to $\mathbf{\Omega_C}$, spherically normalized version of $\mathbf{X_C}$
$(.\tilde{.})$	(N+1)-D vector	homogeneous form of an N-dimensional vector, e.g. $\tilde{\mathbf{x}}, \tilde{\mathbf{X}}$
$(\ldots)_\mathbf{V}$	-	entity defined within a virtual camera reference system, e.g. $\mathbf{x}_\mathrm{V}$ is the virtual equivalent of $\mathbf{x}$
$\mathbf{C}$	3-D vector	camera center in world coordinates ($\mathbf{\Omega_W}$)
$\mathbf{R}$	3×3 matrix	rotation matrix, in the context of camera orientation it maps from camera ($\mathbf{\Omega_C}$) to world coordinates ($\mathbf{\Omega_W}$), i.e. $\mathbf{X} = \mathbf{R}\,\mathbf{X_C}$
$\mathbf{K}$	3×3 matrix	calibration matrix, homogeneous entity that maps $\tilde{\mathbf{x}}'$ to $\tilde{\mathbf{x}}$
$\mathbf{P}$	3×4 matrix	camera matrix, homogeneous entity that maps $\tilde{\mathbf{X}}$ to $\tilde{\mathbf{x}}$
$\mathbf{E}$	3×3 matrix	essential matrix, homogeneous entity, encapsulates the epipolar geometry and relative orientation between two cameras
$\mathbf{b}$	3-D vector	distance vector between two camera centers $\mathbf{C_1}$ and $\mathbf{C_2}$; unit vector if used for relative orientation
f	scalar	focal length, distance between the image plane to the camera center $\mathbf{C}$

c	scalar	camera constant, product of focal length and number of pixels per unit length
φ, θ	angle	azimuth/zenith angle of a projection ray within the camera reference system
β, ψ	angle	angular coordinates of a projection ray within the virtual camera reference system
$\mathbf{x}_\mathrm{C}/\mathbf{x}_0$	2-D vector	principal point on the image/sensor plane in normalized/pixel coordinates
$r(\theta)$	scalar function	determines the distance between a point $\mathbf{x}'$ to the principal point $\mathbf{x}_\mathrm{C}$ on the model image plane depending on θ
${}^a(\ldots)$	-	approximate value of the respective entity, e.g. ${}^a\mathbf{R}, {}^a\mathbf{x}''$
$\hat{(\ldots)}$	-	estimated value of the respective entity, e.g. $\hat{\mathbf{R}}$
$\mathbf{I}$	3×3 matrix	identity matrix
$\mathbf{S_b}$	3×3 matrix	skew-symmetric matrix of a 3-D vector $\mathbf{b}$
SVD	-	Singular Value Decomposition

List of Abbreviations

Abbreviation	Term	Description
LCL	lifting condensation level	height at which a mechanically lifted air parcel from the ground starts to condensate
LFC	level of free convection	height at which an air parcel is more buoyant than its surrounding air
CBH	cloud base height	height of the lower cloud boundary
CTH	cloud top height	height of the upper cloud boundary
Cu	cumulus	low-altitude convective cloud, visually mixed-contrast
Sc	stratocumulus	low-altitude stratiform cloud, visually mixed-contrast
Ac	altocumulus	mid-altitude cloud with cellular structure, visually high-contrast
As	altostratus	mid-altitude stratiform cloud, visually low-contrast
Ci	cirrus	high-altitude ice cloud, visually mixed-contrast
Cc	cirrocumulus	high-altitude ice cloud with cellular structure, visually high-contrast
Cs	cirrostratus	high-altitude stratiform ice cloud, vsually low-contrast
Cb	cumulonimbus	low-altitude convective cloud of large size and high vertical extent

List of Figures

List of Tables

Bibliography

Abraham, S. and Förstner, W. (2005). Fish-eye-stereo calibration and epipolar rectification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(5):278–288.

Abraham, S. and Hau, T. (1997). Towards autonomous high-precision calibration of digital cameras. In El-Hakim, S. F., editor, *Videometrics V*. SPIE.

Alexandrov, M. D., Cairns, B., Emde, C., Ackerman, A. S., Ottaviani, M., and Wasilewski, A. P. (2016). Derivation of cumulus cloud dimensions and shape from the airborne measurements by the research scanning polarimeter. *Remote Sensing of Environment*, 177:144–152.

Allmen, M. C. and Kegelmeyer, W. P. (1996). The computation of cloud-base height from paired whole-sky imaging cameras. *Journal of Atmospheric and Oceanic Technology*, 13(1):97–113.

Bay, H., Ferrari, V., and Gool, L. V. (2005). Wide-baseline stereo matching with line segments. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. IEEE.

Benner, T. C. and Curry, J. A. (1998). Characteristics of small tropical cumulus clouds and their impact on the environment. *Journal of Geophysical Research: Atmospheres*, 103(D22):28753–28767.

Birchfield, S. and Tomasi, C. (1998). Depth discontinuities by pixel-to-pixel stereo. In *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*. Narosa Publishing House.

Blender-Foundation (2016). Blender. `https://www.blender.org/`. [Online; accessed March-2016].

Blyth, A. M. (1993). Entrainment in cumulus clouds. *Journal of Applied Meteorology*, 32(4):626–641.

Boykov, Y., Veksler, O., and Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239.

Bradbury, D. and Fujita, T. (1968). Computation of height and velocity of clouds from dual, wholesky, time-lapse picture sequences. *SMRP Research Paper No. 70*.

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

Brast, M., Neggers, R. A. J., and Heus, T. (2016). What determines the fate of rising parcels in a heterogeneous environment? *Journal of Advances in Modeling Earth Systems*, 8(4):1674–1690.

Brown, A. R., Cederwall, R. T., Chlond, A., Duynkerke, P., Golaz, J.-C., Khairoutdinov, M., Lewellen, D. C., Lock, A. P., MacVean, M. K., Moeng, C.-H., Neggers, R. A. J., Siebesma, A. P., and Stevens, B. (2002). Large-eddy simulation of the diurnal cycle of shallow cumulus convection over land. *Quarterly Journal of the Royal Meteorological Society*, 128(582):1075–1093.

Brown, D. C. (1966). Decentering distortion of lenses. *Photogrammetric Engineering and Remote Sensing*.

Brox, T., Bruhn, A., Papenberg, N., and Weickert, J. (2004). *High Accuracy Optical Flow Estimation Based on a Theory for Warping*, pages 25–36. Springer Berlin Heidelberg, Berlin, Heidelberg.

Cahalan, R. F. and Joseph, J. H. (1989). Fractal statistics of cloud fields. *Monthly Weather Review*, 117(2):261–272.

Cahalan, R. F., Ridgway, W., Wiscombe, W. J., Gollmer, S., and Harshvardhan (1994). Independent pixel and monte carlo estimates of stratocumulus albedo. *Journal of the Atmospheric Sciences*, 51(24):3776–3790.

Chambers, L. H., Wielicki, B. A., and Evans, K. F. (1997). Independent pixel and two-dimensional estimates of landsat-derived cloud field albedo. *Journal of the Atmospheric Sciences*, 54(11):1525–1532.

Chum, O., Matas, J., and Obdrzalek, S. (2004). Enhancing ransac by generalized model optimization. In *Proc. of the ACCV*, volume 2, pages 812–817.

Corbetta, G., Orlandi, E., Heus, T., Neggers, R., and Crewell, S. (2015). Overlap statistics of shallow boundary layer clouds: Comparing ground-based observations with large-eddy simulations. *Geophysical Research Letters*, 42(19):8185–8191.

Crispel, P. and Roberts, G. (2017). All-sky photogrammetry techniques to georeference a cloud field. *Atmospheric Measurement Techniques Discussions*, pages 1–31.

Damiani, R., Geerts, B., Demko, J., Haimov, S., French, J., Zehnder, J., Razdan, A., Hu, J., Petti, J., Leuthold, M., and Poulos, G. S. (2008). The cumulus, photogrammetric, in situ, and doppler observations experiment of 2006. *Bulletin of the American Meteorological Society*, 89(1):57–73.

Damiani, R., Vali, G., and Haimov, S. (2006). The structure of thermals in cumulus from airborne dual-doppler radar observations. *Journal of the Atmospheric Sciences*, 63(5):1432–1450.

Davies, R. (1978). The effect of finite geometry on the three-dimensional transfer of solar irradiance in clouds. *Journal of the Atmospheric Sciences*, 35(9):1712–1725.

de Rooy, W. C., Bechtold, P., Fröhlich, K., Hohenegger, C., Jonker, H., Mironov, D., Siebesma, A. P., Teixeira, J., and Yano, J.-I. (2012). Entrainment and detrainment in cumulus convection: an overview. *Quarterly Journal of the Royal Meteorological Society*, 139(670):1–19.

Dufresne, J.-L. and Bony, S. (2008). An assessment of the primary sources of spread of global warming estimates from coupled atmosphere–ocean models. *Journal of Climate*, 21(19):5135–5144.

El Gamal, A. (2002). High dynamic range image sensors. In *Tutorial at International Solid-State Circuits Conference*, volume 290.

Evans, K. F., McFarlane, S. A., and Wiscombe, W. (2001). The importance of three-dimensional solar radiative transfer in small cumulus cloud fields derived from the nauru, mmcr and mwr. In *Eleventh ARM Science Team Meeting Proceedings, Atlanta, Georgia*, volume 3, page 2001.

Fan, S. and Ferrie, F. (2006). Photo hull regularized stereo. In *The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)*. IEEE.

Felzenszwalb, P. and Huttenlocher, D. (2006). Efficient belief propagation for early vision. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.* IEEE.

Fielding, M. D., Chiu, J. C., Hogan, R. J., and Feingold, G. (2013). 3d cloud reconstructions: Evaluation of scanning radar scan strategy with a view to surface shortwave radiation closure. *Journal of Geophysical Research: Atmospheres*, 118(16):9153–9167.

Förstner, W. (1993). Image matching. In Haralick, R. M. and Shapiro, L. G., editors, *Computer and Robot Vision*, pages 289–379.

Forsyth, D. A. and Ponce, J. (2012). *Computer Vision: A Modern Approach*. Pearson Education, Limited, New York, 2 edition.

Förstner, W. (2011). Minimal representations for uncertainty and estimation in projective spaces. In *Computer Vision – ACCV 2010*, pages 619–632. Springer Berlin Heidelberg.

Förstner, W. and Wrobel, B. P. (2016). *Photogrammetric Computer Vision*. Springer International Publishing.

Furukawa, Y. and Ponce, J. (2010). Accurate, dense, and robust multiview stereopsis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(8):1362–1376.

Genkova, I., Seiz, G., Zuidema, P., Zhao, G., and Girolamo, L. D. (2007). Cloud top height comparisons from ASTER, MISR, and MODIS for trade wind cumuli. *Remote Sensing of Environment*, 107(1-2):211–222.

Giorgetta, M. A., Brokopf, R., Crueger, T., Esch, M., Fiedler, S., Helmert, J., Hohenegger, C., Kornblueh, L., Köhler, M., Manzini, E., Mauritsen, T., Nam, C., Raddatz, T., Rast, S., Reinert, D., Sakradzija, M., Schmidt, H., Schneck, R., Schnur, R., Silvers, L., Wan, H., Zängl, G., and Stevens, B. (2018). ICON-a, the atmosphere component of the ICON earth system model: I. model description. *Journal of Advances in Modeling Earth Systems*, 10(7):1613–1637.

Griffith, E. J., Post, F. H., Koutek, M., Heus, T., and Jonker, H. J. J. (2005). Feature tracking in vr for cumulus cloud life-cycle studies. IPT/EGVE 2005.

Harris, C. and Stephens, M. (1988). A combined corner and edge detector. In *In Proc. of Fourth Alvey Vision Conference*, pages 147–151.

Hartley, R. and Zisserman, A. (2000). *Multiple View Geometry in Computer Vision*. Cambridge University Press.

Hartley, R. I. (2001). An investigation of the essential matrix. In *GE internal report*.

Heinle, A., Macke, A., and Srivastav, A. (2010). Automatic cloud classification of whole sky images. *Atmospheric Measurement Techniques*, 3(3):557–567.

Heus, T. (2008). *On the edge of a cloud*. PhD thesis, Technische Universiteit Delft.

Heus, T. and Jonker, H. J. J. (2008). Subsiding shells around shallow cumulus clouds. *Journal of the Atmospheric Sciences*, 65(3):1003–1018.

Heus, T., Jonker, H. J. J., den Akker, H. E. A. V., Griffith, E. J., Koutek, M., and Post, F. H. (2009). A statistical approach to the life cycle analysis of cumulus clouds selected in a virtual reality environment. *Journal of Geophysical Research*, 114(D6).

Hinkelman, L. M., Evans, K. F., Clothiaux, E. E., Ackerman, T. P., and Stackhouse, P. W. (2007). The effect of cumulus cloud field anisotropy on domain-averaged solar fluxes and atmospheric heating rates. *Journal of the Atmospheric Sciences*, 64(10):3499–3520.

Hirschmuller, H. (2005). Accurate and efficient stereo processing by semi-global matching and mutual information. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR05)*. IEEE.

Hirschmuller, H. and Gehrig, S. (2009). Stereo matching in the presence of sub-pixel calibration errors. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE.

Hu, J., Razdan, A., and Zehnder, J. A. (2010). Geometric calibration of digital cameras for 3d cumulus cloud measurements. *Journal of Athmospheric and Oceanic Technology*, 26(2):200–214.

IDS (2013). *manual for uEye cameras*. [Online; accessed March-2016].

Joseph, J. H. and Cahalan, R. F. (1990). Nearest neighbor spacing of fair weather cumulus clouds. *Journal of Applied Meteorology*, 29(8):793–805.

Kanade, T. and Okutomi, M. (1991). A stereo matching algorithm with an adaptive window: theory and experiment. In *Proceedings. 1991 IEEE International Conference on Robotics and Automation*. IEEE Comput. Soc. Press.

Kannala, J. and Brandt, S. (2006). A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(8):1335–1340.

Kassianov, E., Long, C. N., and Christy, J. (2005). Cloud-base-height estimation from paired ground-based hemispherical observations. *Journal of Applied Meteorology*, 44(8):1221–1233.

Kollias, P., Bharadwaj, N., Widener, K., Jo, I., and Johnson, K. (2014). Scanning ARM cloud radars. part i: Operational sampling strategies. *Journal of Atmospheric and Oceanic Technology*, 31(3):569–582.

Koppe, C. (1896). *Photogrammetrie und internationale Wolkenmessung*. Vieweg.

Öktem, R., Prabhat, Lee, J., Thomas, A., Zuidema, P., and Romps, D. M. (2014). Stereophotogrammetry of oceanic clouds. *Journal of Atmospheric and Oceanic Technology*, 31(7):1482–1501.

Öktem, R. and Romps, D. M. (2015). Observing atmospheric clouds through stereo reconstruction. In Sitnik, R. and Puech, W., editors, *Three-Dimensional Image Processing, Measurement (3DIPM), and Applications 2015*. SPIE.

Kuhn, A., Senst, T., Keller, I., Sikora, T., and Theisel, H. (2012). A lagrangian framework for video analytics. In *2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP)*. IEEE.

Kutulakos, K. and Seitz, S. (1999). A theory of shape by space carving. In *Proceedings of the Seventh IEEE International Conference on Computer Vision*. IEEE.

Lamer, K. and Kollias, P. (2015). Observations of fair-weather cumuli over land: Dynamical factors controlling cloud size and cover. *Geophysical Research Letters*, 42(20):8693–8701.

Löhnert, U., Schween, J. H., Acquistapace, C., Ebell, K., Maahn, M., Barrera-Verdejo, M., Hirsikko, A., Bohn, B., Knaps, A., O'Connor, E., Simmer, C., Wahner, A., and Crewell, S. (2015). JOYCE: Jülich observatory for cloud evolution. *Bulletin of the American Meteorological Society*, 96(7):1157–1174.

Li, M., Schirmacher, H., Magnor, M., and Siedel, H.-P. (2002). Combining stereo and visual hull information for on-line reconstruction and rendering of dynamic scenes. In *2002 IEEE Workshop on Multimedia Signal Processing.* IEEE.

Library, P. C. (2016). Point cloud library. `http://pointclouds.org`. [Online; accessed March-2016].

Liu, C. et al. (2009). *Beyond pixels: exploring new representations and applications for motion analysis*. PhD thesis, Massachusetts Institute of Technology.

Longuet-Higgins, H. C. (1981). A computer algorithm for reconstructing a scene from two projections. *Nature*, 293(5828):133.

Loop, C. and Zhang, Z. (1999). Computing rectifying homographies for stereo vision. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*. IEEE Comput. Soc.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110.

Luhmann, T. (2018). *Nahbereichsphotogrammetrie*. VDE Verlag GmbH.

Malkus, J. S. and Ronne (1954). On the structure of some cumulonimbus clouds which penetrated the high tropical troposphere. *Tellus*, 6(4):351–366.

Martucci, G., Milroy, C., and O'Dowd, C. D. (2010). Detection of cloud-base height using jenoptik CHM15k and vaisala CL31 ceilometers. *Journal of Atmospheric and Oceanic Technology*, 27(2):305–318.

Matas, J., Chum, O., Urban, M., and Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing*, 22(10):761–767.

Michael, M., Salmen, J., Stallkamp, J., and Schlipsing, M. (2013). Real-time stereo vision: Optimizing semi-global matching. In *2013 IEEE Intelligent Vehicles Symposium (IV)*. IEEE.

Micušık, B. (2004). *Two-view geometry of omnidirectional cameras*. PhD thesis, PhD. Thesis, Czech Technical University.

Mueller, K. J., Wu, D. L., Horváth, Á., Jovanovic, V. M., Muller, J.-P., Girolamo, L. D., Garay, M. J., Diner, D. J., Moroney, C. M., and Wanzong, S. (2017). Assessment of MISR cloud motion vectors (CMVs) relative to GOES and MODIS atmospheric motion vectors (AMVs). *Journal of Applied Meteorology and Climatology*, 56(3):555–572.

Muller, J.-P., Denis, M.-A., Dundas, R. D., Mitchell, K. L., Naud, C., and Mannstein, H. (2007). Stereo cloud-top heights and cloud fraction retrieval from ATSR-2. *International Journal of Remote Sensing*, 28(9):1921–1938.

Neggers, R. A. J., Jonker, H. J. J., and Siebesma, A. P. (2003). Size statistics of cumulus cloud populations in large-eddy simulations. *Journal of the Atmospheric Sciences*, 60(8):1060–1074.

Nguyen, D. A. and Kleissl, J. (2014). Stereographic methods for cloud base height determination using two sky imagers. *Solar Energy*, 107:495–509.

Olmo, F. J., Cazorla, A., Alados-Arboledas, L., López-Álvarez, M. A., Hernández-Andrés, J., and Romero, J. (2008). Retrieval of the optical depth using an all-sky CCD camera. *Applied Optics*, 47(34):H182.

Onishi, R. and Sugiyama, D. (2017). Deep convolutional neural network for cloud coverage estimation from snapshot camera images. *SOLA*, 13(0):235–239.

Orville, H. D. and Kassander, A. R. (1961). Terrestrial photogrammetry of clouds. *Journal of Meteorology*, 18(5):682–687.

Oue, M., Kollias, P., North, K. W., Tatarevic, A., Endo, S., Vogelmann, A. M., and Gustafson, W. I. (2016). Estimation of cloud fraction profile in shallow convection using a scanning cloud radar. *Geophysical Research Letters*, 43(20):10,998–11,006.

Pizer, S. M., Johnston, R. E., Ericksen, J. P., Yankaskas, B. C., and Muller, K. E. (1990). Contrast-limited adaptive histogram equalization: speed and effectiveness. In *Visualization in Biomedical Computing, 1990., Proceedings of the First Conference on*, pages 337–345. IEEE.

Plank, V. G. (1969). The size distribution of cumulus clouds in representative florida populations. *Journal of Applied Meteorology*, 8(1):46–67.

Pruppacher, H. and Klett, J. (1996). *Microphysics of Clouds and Precipitation (Atmospheric and Oceanographic Sciences Library)*. Springer.

Román, R., Antón, M., Cazorla, A., de Miguel, A., Olmo, F. J., Bilbao, J., and Alados-Arboledas, L. (2012). Calibration of an all-sky camera for obtaining sky radiance at three wavelengths. *Atmospheric Measurement Techniques*, 5(8):2013–2024.

Romps, D. M. and Öktem, R. (2015). Stereo photogrammetry reveals substantial drag on cloud thermals. *Geophysical Research Letters*, 42(12):5051–5057.

Rusu, R. B. (2009). *Semantic 3D Object Maps for Everyday Manipulation in Human Living Environments*. PhD thesis, Computer Science department, Technische Universitaet Muenchen, Germany.

Scaramuzza, D. (2008). Omnidirectional vision: from calibration to robot motion estimation. *ETH Zurich, PhD Thesis*, 17635.

Scharstein, D., Szeliski, R., and Zabih, R. (2001). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. In *Proceedings IEEE Workshop on Stereo and Multi-Baseline Vision (SMBV 2001)*. IEEE Comput. Soc.

Schmidt, T., Kalisch, J., Lorenz, E., and Heinemann, D. (2016). Evaluating the spatio-temporal performance of sky-imager-based solar irradiance analysis and forecasts. *Atmospheric Chemistry and Physics*, 16(5):3399–3412.

Schneider, J., Schindler, F., Läbe, T., and Förstner, W. (2012). Bundle adjustment for multi-camera systems with points at infinity. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, I-3:75–80.

Schneider, J., Stachniss, C., and Forstner, W. (2016). On the accuracy of dense fisheye stereo. *IEEE Robotics and Automation Letters*, 1(1):227–234.

Schween, J. H., Hirsikko, A., Löhnert, U., and Crewell, S. (2014). Mixing-layer height retrieval with ceilometer and doppler lidar: from case studies to long-term assessment. *Atmospheric Measurement Techniques*, 7(11):3685–3704.

Seiz, G. (2003). *Ground- and Satellite-Based Multi-View Photogrammetric Determination of 3D Cloud Geometry*. PhD thesis, ETH Zurich.

Seiz, G., Baltsavias, E. P., and Gruen, A. (2002). Cloud mapping from the ground: use of photogrammetric methods. *Photogrammetric Engineering & Remote Sensing*, pages 941–951.

Seiz, G. and Davies, R. (2006). Reconstruction of cloud geometry from multi-view satellite images. *Remote Sensing of Environment*, 100(2):143–149.

Seiz, G., Shields, J., Feister, U., Baltsavias, E. P., and Gruen, A. (2007a). Cloud mapping with ground-based photogrammetric cameras. *International Journal of Remote Sensing*, 28(9):2001–2032.

Seiz, G., Tjemkes, S., and Watts, P. (2007b). Multiview cloud-top height and wind retrieval with photogrammetric methods: Application toMeteosat-8hrv observations. *Journal of Applied Meteorology and Climatology*, 46(8):1182–1195.

Sherwood, S. C., Roca, R., Weckwerth, T. M., and Andronova, N. G. (2010). Tropospheric water vapor, convection, and climate. *Reviews of Geophysics*, 48(2).

Siebesma, A. P., Bretherton, C. S., Brown, A., Chlond, A., Cuxart, J., Duynkerke, P. G., Jiang, H., Khairoutdinov, M., Lewellen, D., Moeng, C.-H., et al. (2003). A large eddy simulation intercomparison study of shallow cumulus convection. *Journal of the Atmospheric Sciences*, 60(10):1201–1219.

Siebesma, A. P. and Jonker, H. J. J. (2000). Anomalous scaling of cumulus cloud boundaries. *Physical Review Letters*, 85(1):214–217.

Slesareva, N., Bruhn, A., and Weickert, J. (2005). Optic flow goes stereo: A variational method for estimating discontinuity-preserving dense disparity maps. In *Lecture Notes in Computer Science*, pages 33–40. Springer Berlin Heidelberg.

Sonka, M., Hlavac, V., and Boyle, R. (2007). *Image Processing, Analysis, and Machine Vision*. CL Engineering.

Stein, A., Huertas, A., and Matthies, L. (2006). Attenuating stereo pixel-locking via affine window adaptation. In *Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006.* IEEE.

Stevens, B., Ackerman, A. S., Albrecht, B. A., Brown, A. R., Chlond, A., Cuxart, J., Duynkerke, P. G., Lewellen, D. C., Macvean, M. K., Neggers, R. A. J., Sánchez, E., Siebesma, A. P., and Stevens, D. E. (2001). Simulations of trade wind cumuli under a strong inversion. *Journal of the Atmospheric Sciences*, 58(14):1870–1891.

Stevens, B., Farrell, D., Hirsch, L., Jansen, F., Nuijens, L., Serikov, I., Brügmann, B., Forde, M., Linne, H., Lonitz, K., and Prospero, J. M. (2016). The barbados cloud observatory: Anchoring investigations of clouds and circulation on the edge of the ITCZ. *Bulletin of the American Meteorological Society*, 97(5):787–801.

Stevens, B., Giorgetta, M., Esch, M., Mauritsen, T., Crueger, T., Rast, S., Salzmann, M., Schmidt, H., Bader, J., Block, K., Brokopf, R., Fast, I., Kinne, S., Kornblueh, L., Lohmann, U., Pincus, R., Reichler, T., and Roeckner, E. (2013). Atmospheric component of the MPI-m earth system model: ECHAM6. *Journal of Advances in Modeling Earth Systems*, 5(2):146–172.

Stull, R. B., editor (1988). *An Introduction to Boundary Layer Meteorology*. Springer Netherlands.

Tiedtke, M. (1989). A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Monthly Weather Review*, 117(8):1779–1800.

Tohsing, K., Schrempf, M., Riechelmann, S., Schilke, H., and Seckmeyer, G. (2013). Measuring high-resolution sky luminance distributions with a ccd camera. *Applied optics*, 52(8):1564–1573.

Tuominen, P. and Tuononen, M. (2017). Cloud detection and movement estimation based on sky camera images using neural networks and the lucas-kanade method. In *AIP Conference Proceedings*, volume 1850, page 140020.

Tuytelaars, T. and Mikolajczyk, K. (2007). Local invariant feature detectors: A survey. *Foundations and Trends® in Computer Graphics and Vision*, 3(3):177–280.

Urquhart, B., Kurtz, B., Dahlin, E., Ghonima, M., Shields, J. E., and Kleissl, J. (2015). Development of a sky imaging system for short-term solar power forecasting. *Atmospheric Measurement Techniques*, 8(2):875–890.

Urquhart, B., Kurtz, B., and Kleissl, J. (2016). Sky camera geometric calibration using solar observations. *Atmospheric Measurement Techniques*, 9(9):4279–4294.

Vant-Hull, B., Marshak, A., Remer, L. A., and Li, Z. (2007). The effects of scattering angle and cumulus cloud geometry on satellite retrievals of cloud droplet effective radius. *IEEE Transactions on Geoscience and Remote Sensing*, 45(4):1039–1045.

Veikherman, D., Aides, A., Schechner, Y. Y., and Levis, A. (2015). Clouds in the cloud. In *Computer Vision – ACCV 2014*, pages 659–674. Springer International Publishing.

Walker, J. R., MacKenzie, W. M., Mecikalski, J. R., and Jewett, C. P. (2012). An enhanced geostationary satellite–based convective initiation algorithm for 0–2-h nowcasting with object tracking. *Journal of Applied Meteorology and Climatology*, 51(11):1931–1949.

Weitkamp, C., editor (2005). *Lidar*. Springer New York.

Xu, C. and Prince, J. (1998). Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 7(3):359–369.

Zehnder, J. A., Zhang, L., Hansford, D., Radzan, A., Selover, N., and Brown, C. M. (2006). Using digital cloud photogrammetry to characterize the onset and transition from shallow to deep convection over orography. *Monthly Weather Review*, 134(9):2527–2546.

Zhang, Y. and Klein, S. A. (2013). Factors controlling the vertical extent of fair-weather shallow cumulus clouds over land: Investigation of diurnal-cycle observations collected at the ARM southern great plains site. *Journal of the Atmospheric Sciences*, 70(4):1297–1315.

Zhang, Z. (2000). A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334.

Zhao, M. and Austin, P. H. (2003). Episodic mixing and buoyancy-sorting representations of shallow convection: A diagnostic study. *Journal of the Atmospheric Sciences*, 60(7):892–912.

Zhao, M. and Austin, P. H. (2005). Life cycle of numerically simulated shallow cumulus clouds. part II: Mixing dynamics. *Journal of the Atmospheric Sciences*, 62(5):1291–1310.

BONNER METEOROLOGISCHE ABHANDLUNGEN

Herausgegeben vom Institut für Geowissenschaften der Universität Bonn, Abteilung Meteorologie, durch Prof. Dr. H. FLOHN (Hefte 1-25), Prof. Dr. M. HANTEL (Hefte 26-35), Prof. Dr. H.-D. SCHILLING (Hefte 36-39), Prof. Dr. H. KRAUS (Hefte 40-49), ab Heft 50 durch Prof. Dr. A. HENSE.

Heft 1-69: siehe `http://www.meteo.uni-bonn.de/bibliothek/bma`

70-88: open access, verfügbar unter `https://uni-bn.de/kpSDaQffe1`

Heft 70: ***A S M Mostaquimur Rahman***: Influence of subsurface hydrodynamics on the lower atmosphere at the catchment scale, 2015, 98 S. + XVI.

Heft 71: ***Sabrina Wahl***: Uncertainty in mesoscale numerical weather prediction: probabilistic forecasting of precipitation, 2015, 108 S.

Heft 72: ***Markus Übel***: Simulation of mesoscale patterns and diurnal variations of atmospheric CO_2 mixing ratios with the model system TerrSysMP-CO_2, 2015, [erschienen] 2016, 158 S. + II

Heft 73: ***Christian Bernardus Maria Weijenborg***: Characteristics of Potential Vorticity anomalies associated with mesoscale extremes in the extratropical troposphere, 2015, [erschienen] 2016, 151 S. + XI

Heft 74: ***Muhammad Kaleem***: A sensitivity study of decadal climate prediction to aerosol variability using ECHAM6-HAM (GCM), 2016, 98 S. + XII

Heft 75: ***Theresa Bick***: 3D Radar reflectivity assimilation with an ensemble Kalman filter on the convective scale, 2016, [erschienen] 2017, 96 S. + IX

Heft 76: ***Zied Ben Bouallegue***: Verification and post-processing of ensemble weather forecasts for renewable energy applications, 2017, 119 S.

Heft 77: ***Julia Lutz***: Improvements and application of the STatistical Analogue Resampling Scheme STARS, 2016, [erschienen] 2017, 103 S.

Heft 78: ***Benno Michael Thoma***: Palaeoclimate Reconstruction in the Levant and on the Balkans, 2016, [erschienen] 2017, XVI, 266 S.

Heft 79: ***Ieda Pscheidt***: Generating high resolution precipitation conditional on rainfall observations and satellite data, 2017, V, 173 S.

Heft 80: ***Tanja Zerenner***: Atmospheric downscaling using multi-objective genetic programming, 2016, [erschienen] 2017, X, 191 S.

Heft 81: ***Sophie Stolzenberger***: On the probabilistic evaluation of decadal and paleoclimate model predictions, 2017, IV, 122 S.

Heft 82: ***Insa Thiele-Eich***: Flooding in Dhaka, Bangladesh, and the challenge of climate change, 2017, V, 158 S.

Heft 83: ***Liselotte Bach***: Towards a probabilistic regional reanalysis for Europe, 2017 [erschienen] 2018, VI, 114 S.

Heft 84: ***Yen-Sen Lu***: Propagation of land surface model uncertainties in terrestrial system states, 2017, [erschienen] 2018, X, 120 S.

Heft 85: ***Rüdiger Hewer***: Stochastic physical models for wind fields and precipitation extremes, 2018, 99 S.

Heft 86: ***Sebastian Knist***: Land-atmosphere interactions in multiscale regional climate change simulations over Europe, 2018, 147 S.

Heft 87: ***Jessica Keune***: Integrated terrestrial simulations at the continental scale: Impact of groundwater dynamics and human water use on groundwater-to-atmosphere feedbacks during the European heatwave in 2003, 2019, IX, 172 S.

Heft 88: ***Christoph Beekmans***: 3-D Cloud Morphology and Evolution Derived from Hemispheric Stereo Cameras, 2019, [erschienen] 2020, VIII., 118 S.

Institut für Geowissenschaften
Abteilung Meteorologie
Mathematisch Naturwissenschaftliche Fakultät
Universität Bonn

institut für geowissenschaften
METEOROLOGIE