Article

# Low-cost optical Camera Systems for real-time Mapping Applications

FRANZ KURZ, SEBASTIAN TÜRMER, OLIVER MEYNBERG, DOMINIK ROSENBAUM, HARTMUT RUNGE, PETER REINARTZ, Oberpfaffenhofen & JENS LEITLOFF, Karlsruhe

Summary: Real-time monitoring of natural disasters, mass events, and large accidents with airborne optical sensors is an ongoing topic in research and development. The performance of experimental airborne low-cost camera systems for real-time mapping applications is exemplarily evaluated in terms of geometrical accuracy, radiometric performance, and processing times. Focus lies on the analysis of the geometrical stability of low-cost camera systems within several years of operation and on the evaluation of the limits of direct georeferencing. Further a real-time processing chain including a real-time GPU (graphical processing unit) based orthorectification method is proposed taking into account the maximum possible frame rates of the cameras of five frames per second.

Zusammenfassung: *Preisgünstige Kamerasysteme für Echtzeitanwendungen*. Die Beobachtung von Naturkatastrophen, Großereignissen und Unfällen mit flugzeuggestützten optischen Sensoren in Echtzeit ist ein derzeit wichtiges Thema in Forschung und Entwicklung. In diesem Zusammenhang wird die Leistungsfähigkeit von preisgünstigen Kamerasystemen für Echtzeitanwendungen in Hinblick auf geometrische Genauigkeit, radiometrische Eigenschaften und Prozessierungszeiten evaluiert. Der Schwerpunkt liegt bei der Analyse der geometrischen Stabilität von preisgünstigen Kameras im langjährigen Betrieb und den Grenzen der direkten Georeferenzierung. Weiterhin wird eine echtzeitfähige Prozessierungskette mit einer GPU (graphical processing unit) basierten Orthorektifizierungsmethode für eine maximal mögliche Aufnahmerate von 5 Hz vorgestellt.

## 1 Introduction

Compared to using optical satellite data, airborne monitoring of natural disasters, mass events, and large accidents has the advantage of flexible data acquisition and higher spatial resolution. With the rise of new airborne platforms, there is an increasing demand for low-cost, low-weight and small optical camera systems. These aspects become even more important as the payload of these flying platforms is limited and end users such as police and rescue forces want to equip their proprietary flight squadrons at limited costs. Compared to high-end photogrammetric systems, 36 mm format cameras are very cheap and offer new applications due to their higher frame rate, smaller size and weight. Nevertheless,

the geometrical accuracy and radiometric performance compared to reference photogrammetric systems will also be decisive for the acceptance of low-cost cameras. Different small low-cost camera systems are in use for gyrocopters, helicopter, UAVs (Unmanned Aerial Vehicles), and airships (HAARBRINK & KOERS 2006, ZONGJIAN 2008, NAGAI et al. 2009, LEHMANN et al. 2011). The range of camera types goes from small industrial cameras (ERNST et al. 2003) via middle format cameras (SCHARDT et al. 2010) to large format cameras (PETRIE 2003, CRAMER 2010). Real-time processing of airborne optical camera images in combination with high frame rates paves the way for innovative applications. It is possible to monitor highly dynamic processes like traffic (ROSENBAUM et al. 2008, LEITLOFF et al. 2010)

or persons (SIRMACEK & REINARTZ 2011). Airborne video camera systems are also used for person and traffic monitoring e.g. in SHASTRY & SCHOWENGERDT (2005). DSMs (Digital Surface Models) generated in real-time (ZHU et al. 2010) and real-time orthophoto maps are a valuable data source in different scenarios. A real-time monitoring system based on airborne video cameras which includes the whole processing chain from the airplane to the ground was designed in SCHARDT et al. (2010) and SRINIVASAN & LATCHMAN (2004).

The general requirements for the development of optical camera systems at DLR (Deutsches Zentrum für Luft- und Raumfahrt) were to acquire images with high resolution and wide FOV (field of view) in real-time. Besides, the products like ortho mosaics and traffic parameters should be generated with sufficient geometric accuracy; 3 m absolute horizontal position accuracy are assumed as sufficient in particular for the import into GIS or road databases. For real-time monitoring of highly dynamic processes, the overall processing should avoid data pile-ups and the time between acquisition and delivery should not exceed five minutes. Finally, the sensor systems including the real-time processing chain should be cheap, small and light-weighted enough to be flexibly mountable on airplanes, gliders, and UAVs.

The consideration of all requirements resulted in the sensor and hardware architecture described in this paper. The proposed architecture is exemplarily evaluated based on already existing sensor systems and the experience made in five years of experimental application is presented. The focus lies on the geometrical and radiometric performance of the real-time processing chain, which includes image processors on board, data downlink to the ground station, and further processing and distribution on the ground. Section 2 presents the flight, camera and viewing configuration as well as the software and hardware architecture required for real-time processing. To fulfill the low-cost requirement, the proposed system architecture consists of three 36 mm cameras which are cheap compared to photogrammetric systems. The other components are also relatively cheap, except for the GPS/IMU (inertial measurement unit) from IGI

(IGI 2011). Thus, efforts are made to replace the IMU by a software solution e.g. by optical navigation (KOZEMPEL & REULKE 2009), but in the proposed processing chain the GPS/IMU remains included to allow real-time processing. Moreover, the real-time requirement is not easy to realize, as the number of airborne processing units is restricted. On the other hand, full processing at the ground station in real-time is also impossible due to the limited downlink capacity. Thus, new processors with high parallelization or GPU (graphical processing unit) implementations are designed as described in section 2.4 for the orthorectification process. In section 3, the geometrical properties of the sensor systems are evaluated, which is followed by a comprehensive performance evaluation in section 4.

## 2 The Sensor Hardware and Processing Chain

### 2.1 *The Cameras*

There are three low-cost real-time optical sensor units operated at DLR, the 3K and 3K+ camera system licensed for the DLR airplanes Cessna and Do228 as well as a sensor unit called CHICAGO integrated in a motorized DLR glider powered by a hydrogen-oxygen fuel cell (COPPINGER 2010). The important components are the cameras, which are described in more detail in this chapter.

Each system consists of three non-metric Canon cameras. For the 3K system the Canon EOS 1Ds Mark II camera with Canon lenses is used, whereas the successor model 3K+/CHICAGO uses the CANON EOS 1Ds Mark III camera with Zeiss lenses. The nominal focal length for 3K/3K+ is 50 mm and for the CHICAGO system 35 mm in the side-look and 50 mm in forward / backward direction. The 3K and 3K+ systems are mounted on a ZEISS aerial shock mount ready for the DLR airplanes. The main differences between 3K and 3K+/CHICAGO are the cameras and lenses, the rest of the hardware and software components remain the same. The properties of the sensors of the 3K and 3K+/CHICAGO systems are listed in Tab. 1.

**Tab. 1:** Properties of 3K and 3K+/CHICAGO camera system.

|  | **3K** | **3K+/CHICAGO** |
|---|---|---|
| Cameras | 3 × EOS 1Ds Mark II | 3 × EOS 1Ds Mark III |
| Sensor / Pixel size | Full frame CMOS / 7.21 μm | Full frame CMOS / 6.41 μm |
| Image size | 3 × 4992 × 3328 (16.7 MPix) | 3 × 5616 × 3744 (21.0 MPix) |
| Max. frame rate (Max. images) | 3 Hz (~50 images) | 5 Hz (63 images) |
| File size | 20 MByte (RAW) 5.5 MByte (JPEG level 8) | 25 MByte (RAW) 6.5 MByte (JPEG level 8) |
| ISO | 100 – 1600 | 50 – 3200 |
| Aperture | 1.4 – 22 | 1.4 – 22 |
| Lenses | Canon EF 1.4 50 mm | Zeiss Makro-Planar 2/50 mm Zeiss Distagon T* 2/35 mm |
| Data rate (1 cam, JPEG level 8) at 0.5 Hz | 8.3 MByte/s | 9.8 MByte/s |
| Interface | Firewire IEEE 1394a | USB 2.0 |

The Mark III camera delivers 21.0 MPix compared to 16.7 MPix of the Mark II camera. Thus, the GSD (ground sample distance) of an image taken from 1000 m above ground in nadir direction is 15 cm and 13 cm for the 3K and the 3K+ systems, respectively. The Mark III camera also outperforms the Mark II by a higher frame rate and the maximum number of images, which is important for photogrammetric applications where high overlaps are necessary. The camera offers 5 Hz maximum frame rate resulting e.g. in 97 % overlap at 1000 m above ground at a typical flying speed of 140 knots. According to the specification, the Mark III takes up to 63 images with 5 Hz. The number of images is limited by an overflow of the internal memory, which depends on the data rate produced by the cameras. The listed data rates of 8.3 and 9.8 MByte/s, respectively, at 0.5 Hz can be used for unlimited continuous acquisition. Therefore, different acquisition modes are required depending on the application.

The different acquisition modes are illustrated in Fig. 1. In principle, the cameras of the 3K and 3K+ systems are arranged to provide one nadir view and two oblique views. The oblique viewing angle is configurable freely with a maximum of 32°. The camera system can be installed across track or along track, which results in a wide FOV of up to 104° across or along track, respectively (see Tab. 2).

Typical applications for along-track acquisition are the monitoring of roads and places up to one minute since certain target areas remain longer in the FOV. For applications which require higher frame rates, e.g. traffic monitoring, the camera systems are in a recording mode that is called "burst mode". In this mode, the cameras take a series of e.g. three exposures with a frame rate of up to 5 Hz, and then it pauses for several seconds. During this pause, the plane moves significantly over ground. Then, with an overlap of about 10 % to 20 % to the first exposure "burst", the second exposure sequence is started. By this periodical shift between exposure sequences and brakes, we are able to perform an area-wide traffic monitoring while avoiding data overflow in the cameras. In the continuous mode, the frame rate is lower but with constant overlap, e.g. for 3D applications.

The CHICAGO system has a different configuration of the three cameras, providing two acquisition modes, the along-track continuous mode and the circling mode. In the first mode, the glider follows a certain route which is monitored by the forward and backward cameras. The side-looking camera is used for the circling mode, where the glider is circling around a fixed position, thus, allowing a continuous monitoring of a static position on ground.

Typical flying heights of the camera systems are in the range of 500 m to 3000 m above ground, which leads to a GSD between 6.5 cm and 45 cm. In Tab. 2, all relevant properties of the viewing configurations are listed based on 35 mm lenses for the CHICAGO circling mode and 50 mm lenses for all other modes. The viewing directions of the 3K/3K+ side-looking cameras are variable up to 32°. The FOV at the maximum angle is 2×52° across track and 2×13° in the other direction. For the CHICAGO system, the camera viewing angles are fixed. The coverage is defined as the largest possible rectangle in the ortho-rectified images from all three cameras assuming a horizontal airplane orientation. This assumption is not valid for the CHICAGO circling modus, as the glider rolls with a roll angle up to 30° in the direction of the side-looking camera.
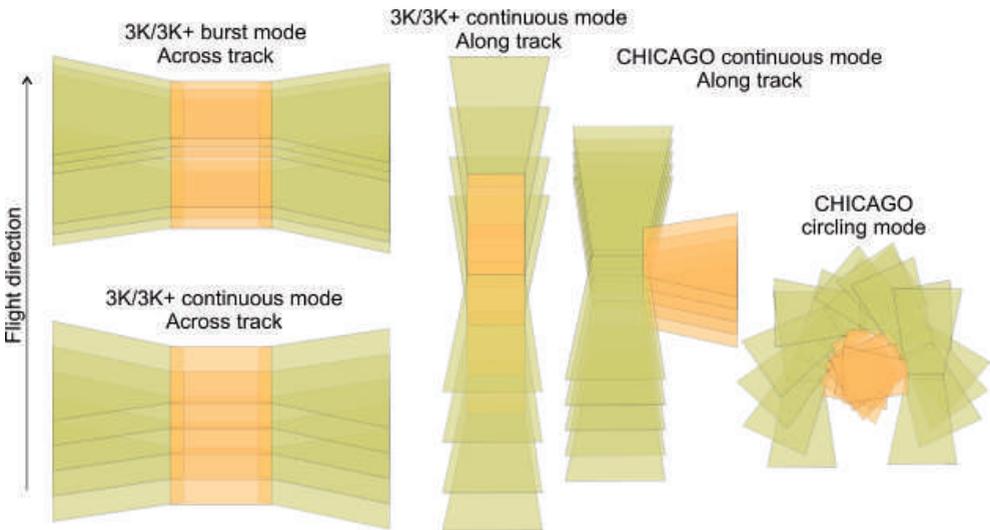


**Fig. 1:** Acquisition modes of 3K, 3K+ and CHICAGO system. Orange: Nadir camera (3K/3K+) / side- looking camera (CHICAGO). Green: side-looking cameras (3K/3K+) / forward and backward camera (CHICAGO).

**Tab. 2:** Properties of the different viewing configurations.

|  | **3K** | **3K+** | **CHICAGO Forward-Backward** | **CHICAGO Circling (assuming +20° roll angle)** |
|---|---|---|---|---|
| Viewing directions | 1× nadir, 2× ±32° /variable | 1× nadir, 2× ±32° / variable | 2× ±13° along track | 1× 11° right look (31° install. angle) |
| FOV | ±52° across, ±13° along | ±52° across, ±13° along | ±20° across (50 mm lenses) ±27° along (50 mm lenses) | -8° to +30° right look (35 mm lenses) ±27° along (35 mm lenses) |
| Coverage / GSD @ 500m | 1280 m × 240 m / 7.5 cm nadir | 1280 m × 240 m / 6.5 cm nadir | 363 m × 509 m / 6.5 cm nadir | 358 m × 519 m 9.2 cm |
| Coverage / GSD @ 1000m | 2560 m × 480 m / 15 cm nadir | 2560 m × 480 m / 13 cm nadir | 726 m × 1018 m / 13 cm nadir | 716 m × 1038 m 18.4 cm |
| Coverage / GSD @ 3000m | 7680 m × 1440 m / 45 cm nadir | 7680 m × 1440 m / 39 cm nadir | 2178 m × 3054m / 39 cm nadir | 2148 m × 3114 m 54.9 cm |

## 2.2 The real-time Processing Chain: Hardware

For the real-time processing and the distribution of airborne images, an on-board processing chain including data transfer from the airplane to the ground is currently under development. The on-board system consists of the optical sensors, the GPS/IMU, the processing units, and a C-band microwave data link with a downlink capacity of up to 54 MBit/s depending on the distance and bandwidth (Fig. 2). Upon receiving the pre-processed data from the airplane, the mobile ground station processes the data and provides them to the end users via web-based portals (KURZ et al. 2011).

Tab. 3 lists the sizes and weights of all necessary hardware components as they are installed in the DLR glider. Further improve-
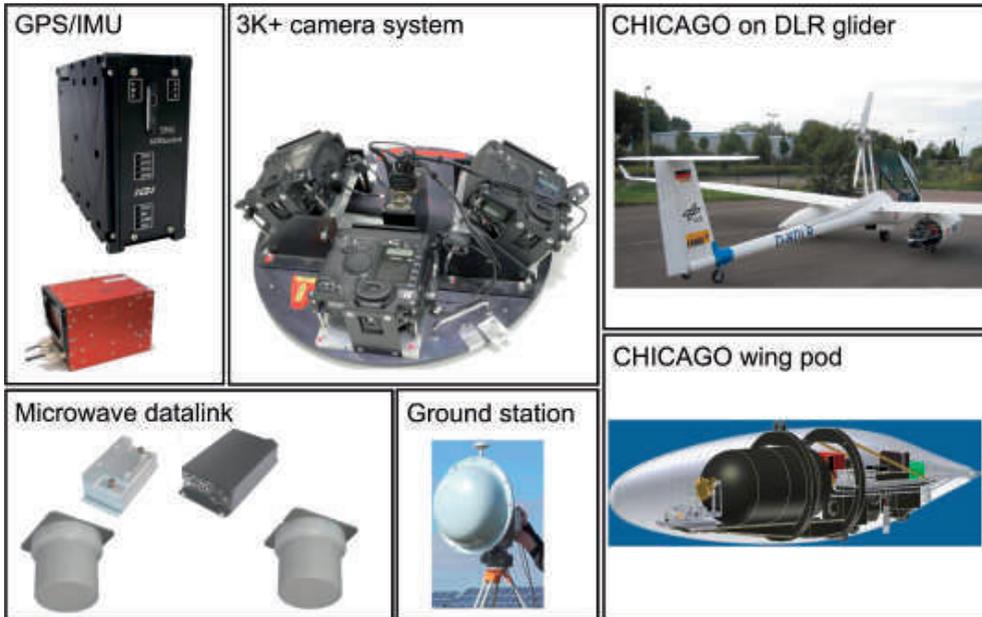


**Fig. 2:** 3K+/CHICAGO airborne and ground based hardware components.

**Tab. 3:** Size and weight of CHICAGO airborne hardware components.

| Nr | Component | Size (mm) | Weight (kg) |
|---|---|---|---|
| 2× | Microwave antenna SRS (SRS 2011) | 120×120×113 | 2×0.75 |
| 1× | Network radio SRS | 58×120×230 | 1.00 |
| 2× | Amplifier SRS | 78×108×220 | 2×1.60 |
| 3× | Canon EOS 1Ds cameras | 156×159×80 | 3×1.21 |
| 3× | Zeiss lenses | ø72 length 65 | 3×0.53 |
| 1× | GPS Antenna | | 0.20 |
| 1× | IMU (IGI Aerocontrol IId) | 200×132×137 | 2.10 |
| 1× | GPS/IMU processor (IGI) | 65×140×205 | 1.80 |
| 1× | PC unit | 360×240×30 | 2.5 |
| 1× | System mount + cables + screws etc. | | 9.8 |
| | | | ∑ 27.32 |

ments with respect to weight, volume, and price could be achieved by substituting the IMU with a smaller but less accurate version or by an optical navigation software solution.

## 2.3 *The real-time Processing Chain: Software Architecture*

The real-time image processing chain on-board the airplane is illustrated in Fig. 3. Each of the cameras is connected via FireWire (3K) or USB (3K+) to a PC on which a Canon soft-ware interface (EDSDK) and the orthorectification process run. The external flash output of the camera triggers the GPS/IMU. Every time a flash signal is received by the GPS/IMU it creates a hardware event and sends the exterior orientation parameters at the time of this event to the camera PC via a TCP connection. The camera control module runs a TCP client and assigns the received exterior orientation data to the image just downloaded by the camera. Using a DEM (digital elevation model), e.g. from the Shuttle Radar Topographic Mission in 2000 (SRTM), together
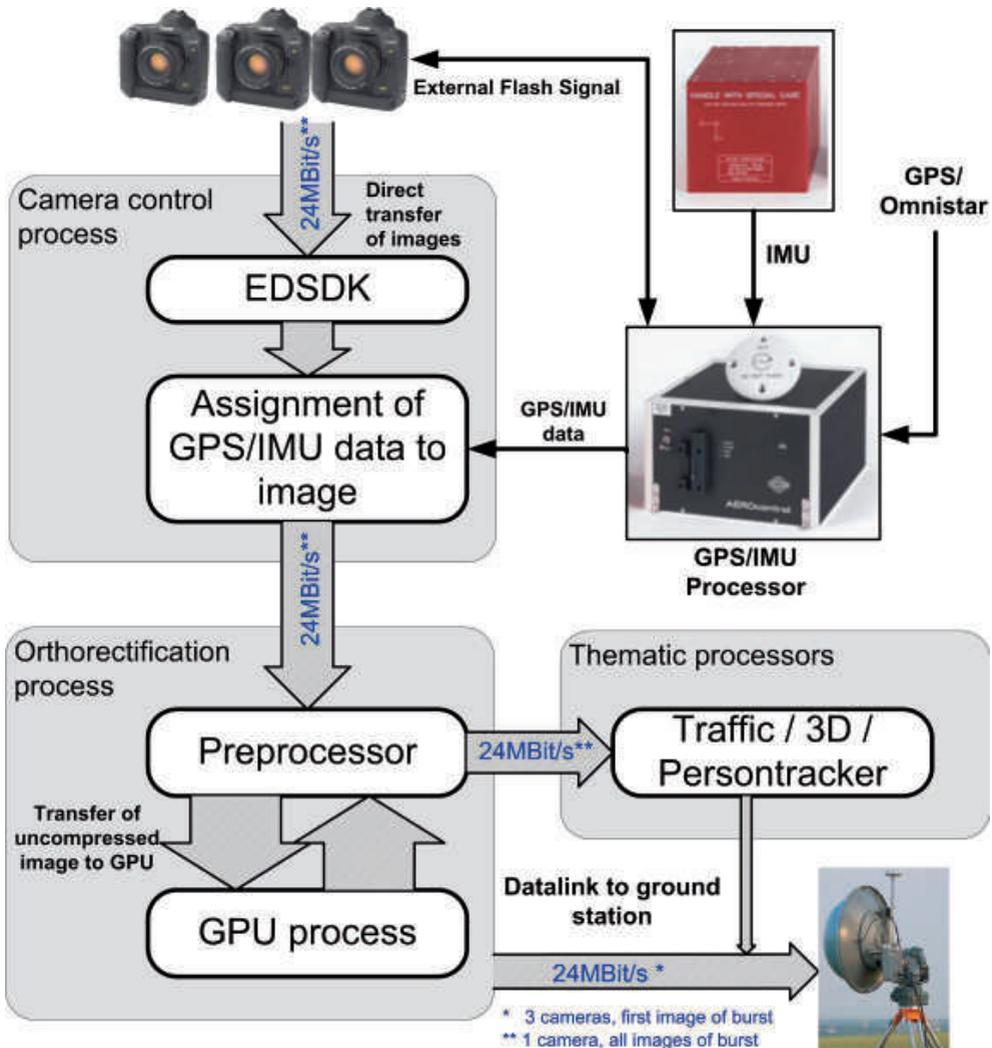
**Fig. 3:** On-board image processing system.

with the calibration and orientation parameters, each image will be orthoprojected on the GPU (section 3.4). Further thematic processors, like automatic traffic processors, person trackers, or 3D processors, are attached to the processing chain. Assuming the 3K+ burst mode (section 3.1.) as the standard configuration, the data rate at the cameras is about 24 Mbit/s, which results from taking three pictures with a file size of 9 MB in two seconds and then pausing for another seven seconds. The data streams of the three cameras are unified at the datalink. To avoid data pile-ups and to fulfill the real-time requirement, the orthorectification process, the thematic processors, and the datalink should be ready for data rates of 24 Mbit/s.

## 2.4 *GPU-enabled Orthorectification*

The orthorectification needs to meet the required timing constraints of 24 Mbit/s. As each pixel can be orthorectified independently, the orthorectification is well-suited for GPU architectures. Our implementation runs on CUDA-enabled (Compute Unified Device Architecture) GPUs which hold a large number of so-called *thread processors*. The basic idea is that each thread processor calculates the geo- and colour information for a single pixel of the orthoimage in one *thread*. This calculation runs in 768 threads simultaneously resulting in a large speedup compared to a CPU-only implementation. The threads are executed on the physically separated GPU, which operates as a coprocessor to the host's CPU where the rest of the program runs. This is done by well documented CUDA extensions to the C language. The main programming challenge is the correct usage of the different kinds of processing units on the GPU and the different kinds of memory areas to maximize the performance gain (NVIDIA 2011).

As illustrated in Fig. 4, the host does all preparatory steps for the direct georeferencing such as calculating the ground-to-image rotation matrix or creating a bounding box for the output image (step 1 in Fig. 4). This information only needs to be calculated once for each image, so there would not be a considerable speed-up if the GPU performed these steps.
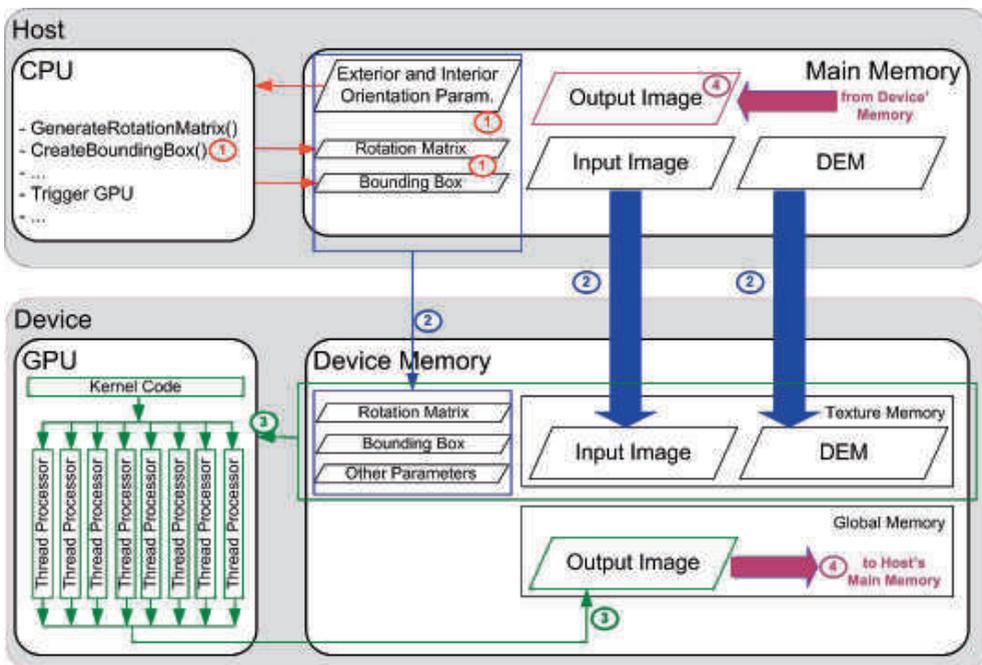


**Fig. 4:** Instruction throughput and memory utilization of the orthorectification process. The encircled numbers refer to the processing steps described in the text.

Step 2) in Fig. 4 loads the input image and the DEM of this region into the host memory and from there into the texture memory which is part of the device memory. The texture memory is a special memory area; it is cached and supports 2D memory layouts, 2D indexing, and automatic resampling. All of these features suit perfectly well for image data. The memory is read-only, though, so we cannot use it for storing the output image. As a side note, newer architectures do support writeable texture memory, which is called surface memory.

After loading the rotation matrix and camera parameters to the cached constant memory, the kernel function is called. This means that all available thread processors on the GPU start to execute at the same time (step 3 in Fig. 4). Each thread processor runs the same kernel with the same algorithm in parallel and calculates the UTM coordinates for one pixel in the output image. The processing is separated in blocks which consist of a constant number of threads, so every block of threads calculates one particular region. In combination with a structured block index the image coordinates in the whole output image are uniquely determined, so that we can now focus on the processing steps within the kernel.

Basically, the orthorectification kernel implements the collinearity equations in direction from object to image space. Starting from the object space the pixel position in the input image is calculated based on the interior

and exterior orientation parameters of the image. In image space, we use a self-written bilinear interpolation function for determining the output colour vectors because CUDA only supports automatic resampling if the pixels are stored as 32-bit floats, which would waste too much memory in our case. Step 4 in Fig. 4 transfers the complete orthorectified output image back to host memory. Due to the slow PCI bus this step takes as much time as all preceding steps but cannot be avoided.

## 3  Sensor Calibration

In this section, the geometric properties of the 3K/3K+/CHICAGO camera systems are evaluated. For this purpose, the interior camera parameters are determined by in-flight calibration. For the 3K camera system, the interior camera parameters were repeatedly evaluated in the last years. The results showed that the overall accuracy of these low-cost non-metric cameras is significantly lower than for standard photogrammetric cameras (Kurz 2009). The reasons are thermal effects, non-flatness of the image plane, aberrations, as well as defocusing problems. The latter problem occurred because the lenses were not adjusted to the camera gauges, i.e. points in the infinity were not imaged in the focal plane correctly. The next generation camera systems 3K+ and CHICAGO have been improved by using more appropriate lenses (Zeiss Makro Planar) with
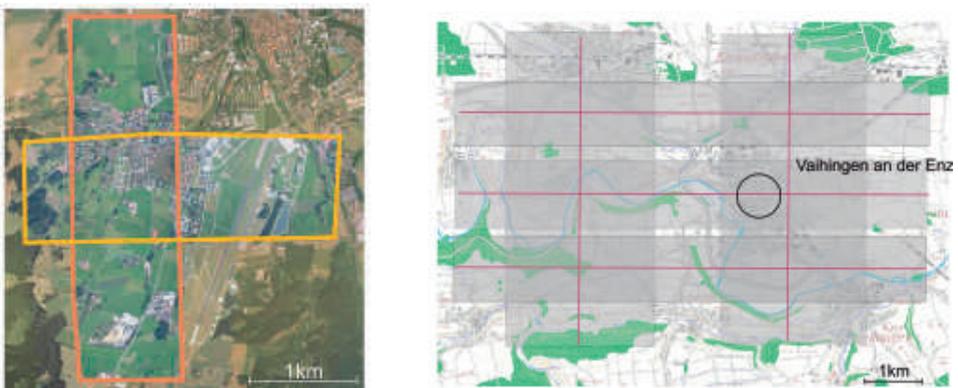


**Fig. 5:** Left: Calibration site in Kaufbeuren on 21st September 2010 with 40 ground control points. Flying height was 1350 m above ground. Right: Calibration site in Vaihingen on 15th July 2008 with 200 ground control points. Flying heights were 1350 m and 2600 m above ground, respectively.

lower distortions and by adjusting the lenses to the camera gauges.

For the calibration of the 3K and 3K+ systems, calibration flights were performed in 2008 and 2010 at the calibration sites in Vaihingen and Kaufbeuren (Fig. 5). The goal of the calibration is to estimate the interior orientation parameters, to validate the interior camera model chosen, to estimate boresight misalignment angles as well as drift and offset parameters of the IMU, and to compare the performance of 3K and 3K+ cameras in terms of accuracy and lens distortions.

For each campaign, a self calibrating bundle adjustment is performed to estimate the interior orientation parameters of each camera. Tie points are matched automatically and all control and check points are measured manually. Additionally, the GPS positions of the projection centres are introduced into the bundle adjustment. A subset of a physical distortion model (FRASER 1997) is chosen for the calibration. Up to nine additional parameters are introduced into the calibration: the focal length c, the coordinates of principal point $x_0$ and $y_0$, two radial symmetric distortion parameters $A_1$ and $A_2$, two asymmetric parameters $B_1$ and $B_2$, and finally a scaling $C_1$ and an affine shearing parameter $C_2$. At the radius $R_0$ the radial symmetric distortion is zero by definition, which avoids too high distortion values at the edges and reduces the correlation with the focal length. The radius $R_0$ was set to 0.014 m, which corresponds to a third of the sensor diagonal. Assuming x and y to be the distorted image coordinates, the corrections $\Delta x$ and $\Delta y$ are then calculated by the following equations.

$$\Delta x = x_p + A_1 x_* \left( r^2 - R_0^2 \right) + A_2 x_* \left( r^4 - R_0^4 \right)$$
$$+ B_1 \left( r^2 + 2x_*^2 \right) + B_2 2 x_* y + C_2 y \qquad (1)$$

$$\Delta y = y_p + A_1 y \left( r^2 - R_0^2 \right) + A_2 y \left( r^4 - R_0^4 \right)$$
$$+ B_1 \left( r^2 + 2y^2 \right) + B_2 2 x_* y \qquad (2)$$

with $r = \sqrt{x_*^2 + y^2}$ and $x_* = \dfrac{x}{C_1}$. The undistorted image coordinates $x'$ and $y'$ are then calculated by $x' = x + \Delta x$ and $y' = y + \Delta y$.

Fig. 6 shows distortion curves based on the results of the self calibrating bundle adjustment. The parameters $B_1$, $B_2$, $C_1$ and $C_2$ are omitted as $B_1$, $B_2$, $C_2$ do not differ significantly from zero and the scaling parameter $C_1$ is close to one. As expected, the radial distortion of the 3K+ lenses is lower than of the 3K lenses. The maximum distortion at the chip corners of the 3K is 20 pixels compared to 8 pixels for the 3K+. Comparing cameras of the same type, the differences in radial distortion are below one pixel, whereas the principal point moves significantly for each camera and even moves for a single camera between the years. This movement gives a hint, firstly, that the estimation of the principal point may be biased by strong correlations with the exterior orientation parameters, and secondly, that the connection lens to camera body is not stable enough. The comparison of the calibration parameters obtained for the years 2008 and 2010 shows that there is a movement of the principal point of 157 μm in the y-coordinate. In fact, the principal point is weakly determined and changes in the lens-camera connection can be neither confirmed nor ruled out.

Additionally, the measured GPS positions of the projection centres must also be corrected in terms of the leverarm errors, GPS offsets and drifts, and the latency between image exposure and the registration at the GPS processor. All these errors influence the accuracy of the determined positions of the projection centres. By using differential GPS measurements in combination with Kalman filtering, the GPS offsets and drifts are negligible. The estimation of the GPS latency is highly correlated to interior camera parameters and is therefore neglected in further processing. Assuming latency not to be longer than 3 ms, which corresponds to a movement of the airplane of about 20 cm, and assuming a lever arm correctness of 10 cm, the resulting positioning error of the projection centre of 30 cm will move the principal point by about two pixels. In contrast to the principal point, the geometrical properties of the lenses and the focal length did not change significantly within the years. In particular, the radial distortion parameters only vary up to 10 % for the same lens type.

The chromatic aberrations of the lenses were estimated by comparing the results of the bundle adjustments using the single bands separately with the results using the RGB images. For this purpose, all tie and control

points measurements were performed separately in the RGB image and in the images corresponding to the single bands. As visualized in Fig. 6, no significant differences in the radial distortion parameters between the red and blue bands were encountered (the green band is between the red and the blue band).

We also compared two models for radial distortion. In the first model, we only used the parameter $A_1$, whereas in the second model we used both $A_1$ and $A_2$. Comparing the distortion curves achieved for these two models (Fig. 6, centre), differences of up to three pixels can be observed at the edges. An additional analysis of the residuals shows that the one-parameter model cannot replace the two parameter model, as this will cause up to three pixels error. The two-parameter model of radial distortion seems to be sufficient, as on the one hand other distortion parameters are not significantly estimated and on the other hand the image residuals do not show any systematic behaviour. In Fig. 7, the residual prediction errors are plotted, which are averaged in a grid defined in the images. Systematic patterns of these averaged errors which would reveal model errors of the chosen calibration model are not visible. The maximum error encountered for 3K+ lies on the chip corners and is about 1.5 pixels.

## 4   Performance Evaluation

### 4.1   *Geometrical Accuracy*

In this section, the overall accuracy of the optical sensors is evaluated. The evaluation has two parts. The first part is the evaluation of
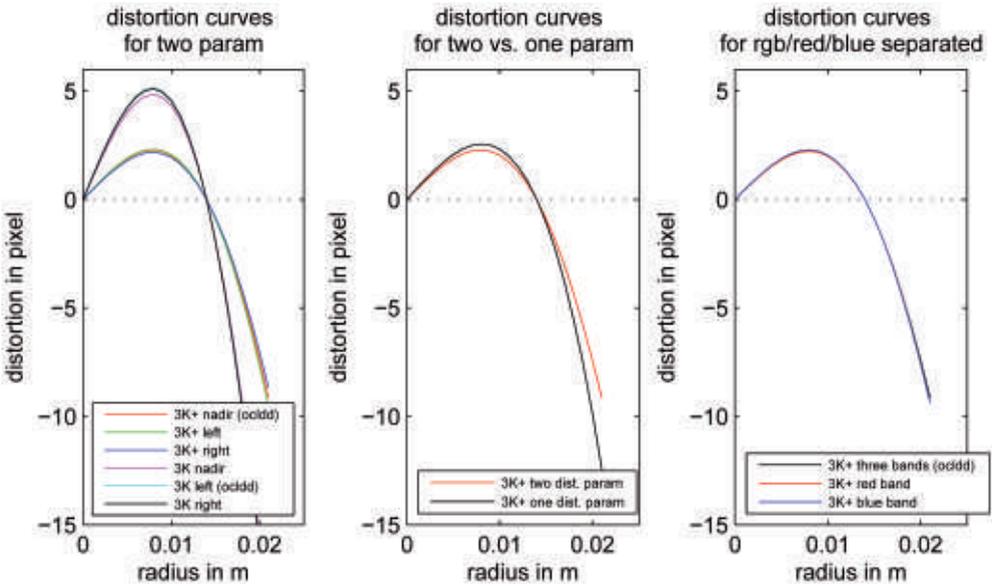


**Fig. 6:** Visualization of the 3K/3K+ lens distortion parameters, ocldd = occluded.
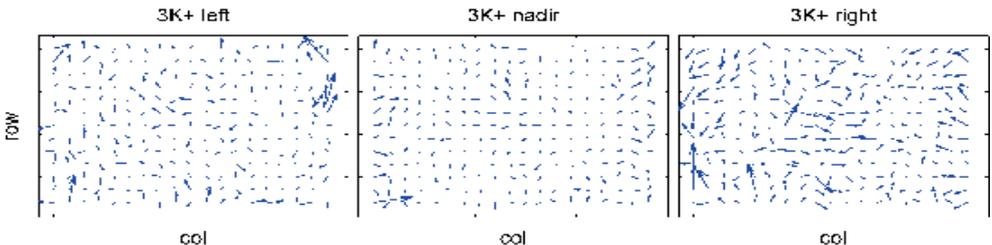


**Fig. 7:** Separate plots of averaged residual prediction errors for each 3K+ camera.

the performance of bundle adjustment based on ground control points and the second part is the performance of direct georeferencing, which is only based on the GPS/IMU data without any additional control information.

The accuracy potential of the sensor in the bundle adjustment is assessed using the theoretical standard deviations at the check points from bundle adjustment and the empirical RMSE (root-mean-square error) using the differences between estimated and measured coordinates of the check points. These values are listed in Tab. 4, separately for the 3K and 3K+ datasets from the years 2008 and 2010. In general, the errors are smaller in 2010 than in 2008, when the images were slightly defocused and thus the control and tie points were not determined precisely enough. The performance of the 3K and 3K+ sensors in 2010 is comparable in the order of around 0.3 m in the XY-coordinate and 0.6 m in the Z-coordinate. Taking into account the errors of GPS mea-

sured check points and targeting errors, the empirical RMSEs for all datasets are slightly higher than the theoretical ones, but the differences are still not significant.

The accuracy of the direct georeferencing depends mainly on the performance of the positions and attitudes determined by GPS/IMU measurements. Most crucial is the determination of the image attitudes with respect to the reference coordinate system, which implies several transformation steps from the image coordinate system to the reference coordinate system (MÜLLER et al. 2002). One important transformation step rotates the image coordinate system into the IMU coordinate system, which is called boresight misalignment. This misalignment is usually determined by a bundle adjustment using the attitudes determined from IMU measurements and additional control information. Another transformation step is the rotation from the IMU coordinate system into a local tangential earth fixed coordi-

**Tab. 4:** Theoretical and empirical accuracy from the bundle adjustment (NI: number of images, NP: number of object points, q: across overlap, p: along overlap) at NC check points.

|  | 3K (*) $\sigma_0$ = 2.39 µm; NI=281; NP=3e5; NC=70; GSD=23 cm; q=65%; p=60% | | 3K Nadir (**) $\sigma_0$ =2.5 µm; NI=35; NP=1e5; NC=10; GSD=22 cm; q=n.a.; p=60% | | 3K+ (**) $\sigma_0$ =2.37 µm; NI=60; NP=1e5; NC=10; GSD=20 cm; q=n.a.; p=50% | |
|---|---|---|---|---|---|---|
|  | $\sigma_{theor}$ | $RMSE_{empir}$ | $\sigma_{theor}$ | $RMSE_{empir}$ | $\sigma_{theor}$ | $RMSE_{empir}$ |
| X | 0.128 m | 0.647 m | 0.094 m | 0.280 m | 0.083 m | 0.138 m |
| Y | 0.147 m | 0.651 m | 0.101 m | 0.388 m | 0.078 m | 0.365 m |
| Z | 0.325 m | 0.576 m | 0.523 m | 0.767 m | 0.400 m | 0.452 m |

(*) Values from calibration 2008 (KURZ 2009)
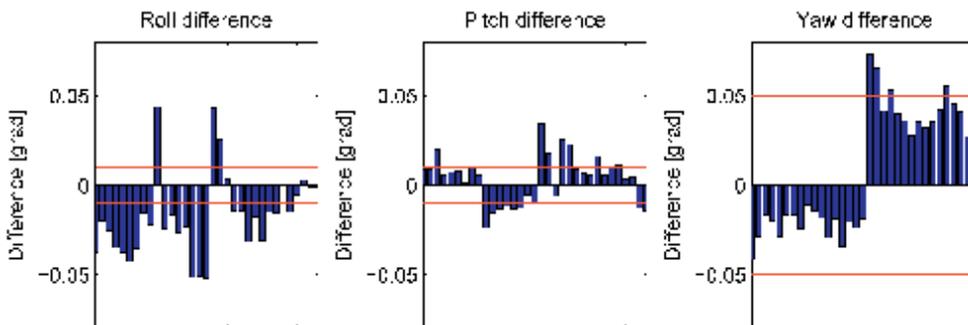(**) Values from calibration 2010



**Fig. 8:** Differences between reference image attitudes from bundle adjustment with image attitudes determined from IMU measurements and corrected for boresight misalignment. Red lines signalize the expected accuracy from the real-time IMU data.

nate system, which is described by the IMU-determined angles, roll, pitch, and yaw, whose accuracy depends on the type of the IMU and the way the angles are determined from the original IMU measurements. The real-time solutions required for the real-time processing chain are less accurate by up to a factor of two than the solutions obtained after post-processing.

Additional errors deteriorate the accuracy of the final rotation. Fig. 8 shows the differences between the image attitudes determined by direct georeferencing and the image attitudes estimated by bundle adjustment. The differences encountered are partly larger than the accuracy of the IMU data. For instance, the maximum difference of the roll angle is 0.05°, whereas the accuracy of a roll angle determined from IMU measurements is 0.01°. The main reasons for these differences are a decoupling of the IMU and sensors due to vibrations, shocks, and platform deformations, and errors both in the boresight misalignment determination and in the reference bundle adjustment. The decoupling effects are mainly caused by an unfavourable mounting of the IMU in the airplane (relatively far away from the sensors) due to the special configuration of the DLR airplanes. Mounting the IMU directly on the sensor platform may reduce the errors and is planned for future integrations of the system into the aircrafts. Nevertheless, the performance of high-end sensors (CRAMER et al. 2000) cannot be reached as in the low-cost environment, the optical sensor and the other hardware components except the GPS/IMU do not achieve the standards of high-end systems.

By projection of the maximum roll errors of around 0.05°, the maximum pitch errors of around 0.03°, and the maximum yaw errors of 0.05°, the maximum positioning error on the ground is calculated for a flying height of 1000 m above ground. The range goes from 1.45 m in nadir to 3.25 m and 3.33 m at the edges of the FOV for the 3K/3K+ and CHICAGO configurations, respectively. There is a linear dependence of these positioning errors from the flying height. In Fig. 9 the resulting positioning errors are illustrated for the two flight configurations with a flying height of 1000 m above ground. Compared to these er-
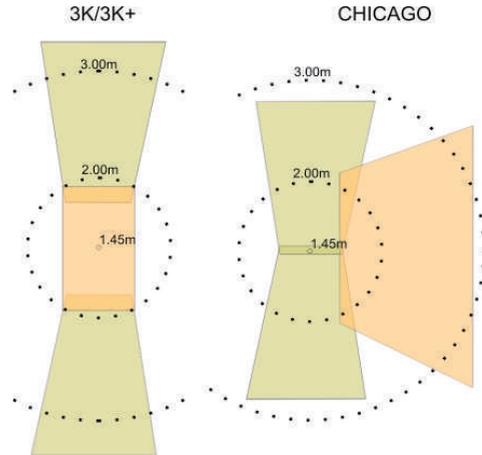


**Fig. 9:** Maximum positioning error by direct georeferencing for the 3K/3K+ and CHICAGO flight configuration.

rors caused by inaccurate image attitudes, the positioning errors due to inaccurate projection centres are negligible, as lever arm deviations, latencies, and GPS errors together do not exceed 0.3 m (assuming a differential GPS solution, latencies < 3 ms, and precisely measured leverarms).

## 4.2  *Radiometric Performance*

This section evaluates the radiometric properties of the 3K/3K+ camera system. The goal is to optimize the use of off-the-shelf cameras for airborne monitoring purposes, i.e. to acquire images with best resolution and contrast in the presence of forward motion blurring and changing incoming radiance. In contrast to high level photogrammetric systems, the forward motion blurring of off-the-shelf cameras is reduced by short exposure times, which worsens the conditions for achieving radiometrically optimal images. As the internal processing of the camera has no changeable parameters, it works as a black box and there is no further influence on how they affect the radiometric quality. The remaining free configurable parameters are the f-number and the ISO speed which are dependent on each other, so that only an appropriate combination allows the best possible imaging result. Concise information about the influence of these

parameters on the radiometric performance is given in the following paragraphs.

The f-number is the focal length divided by the "effective" aperture diameter. A low f-number (e.g. 2.0) passes a lot of light to the sensor but also results in blurring due to the larger circle of confusion. However, the image sharpness in the focal plane varies with the relative aperture size. Additionally, there is optical vignetting which is sensitive to the f-number and lens architecture. In general the blurring can be cured by a reduction in aperture of 2 steps. Due to the lens properties of 3K/3K+ f-numbers greater than 4.0 are capable to produce satisfying results. For instance Zeiss Makro Planar 2 has aperture range of f/2 to f/22.

The shutter speed is indirectly proportional to the light reaching the sensor. As mentioned, short exposure times are aspired to reduce forward motion blurring. A flying velocity of e.g. 70 m/s at 1000 m altitude, with a shutter speed of 1/2000 s results in 3.5 cm blurring which approximately corresponds to 1/4 pixel. Higher shutter speed values reduce the incoming light and thus enforce the f-number and film speed to inappropriate values. Our test supports the assumption that a shutter speed of 1/2000 s is an acceptable compromise. Fig. 10 shows the edge spread functions (ESF) and their corresponding line spread functions (LSF) based on an image with shutter speed 1/2000 s in Fig. 10a and one with shutter speed 1/8000 s in Fig. 10b. Comparing both LSFs shows that the image with lower shutter speed has sharper edges (sigma 0.73 respectively sigma 0.92). The reason is obviously that a higher shutter speed is not able to compensate the lower f-number.

ISO speed is the measure of the sensor's sensitivity to light. Higher values result in noisy images. Therefore our aim is a low ISO speed. But this can be an impossible requirement, especially on cloudy days. The impact of different ISO speed parameters can be observed in Figs. 11a and 11b, which illustrates that higher ISO values cause noisy images.

A test campaign with 3K/3K+ sensor was performed with different f-number, shutter speed values and ISO settings to find out
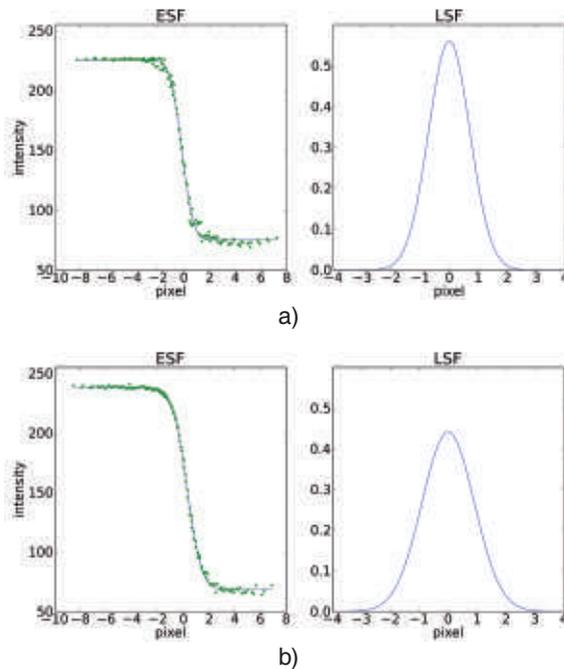


a)

b)

**Fig. 10:** a) Edge spread and line spread function (σ = 0.73 pixels) of image with shutter speed 1/2000 s, f-number 4.0, ISO 640, b) Edge spread and line spread function (σ = 0.92 pixels) of image with shutter speed 1/8000 s, f-number 2.8, ISO 1250.
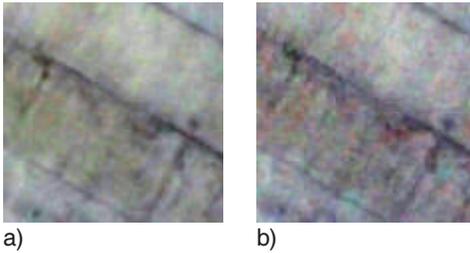
**Fig. 11:** Visualization of 3K sensor noise (stretched images); a) ISO 250, 1/2000 s, f2.8, b) ISO 1600, 1/2000 s, f5.6.

the best camera settings with highest effective GSD. The range of settings for the shutter speed was 1/1000 s to 1/8000 s, for the f-number 2.8 to 5.6, and for the ISO value 250 to 1600. The campaign showed that the best results are obtained by taking fixed settings for the exposure time and the f-number, while the ISO setting is variable according to the illumination conditions. The settings vary from case to case, for instance areas made from concrete have a higher reflectance than forests.

The resolution (effective GSD) of the 3K/3K+ camera system was determined by a Siemens star (Fig. 12) with a diameter of five metres. According to that experiment we obtain an effective GSD of 18.6 cm for the 3K and 13.2 cm for the 3K+ camera compared to the theoretical GSD of 15 cm respectively

13 cm from 1000 m above ground. Also the signalized edge is sharper in the 3K+ image than in the 3K image. The standard deviation of the LSF is $\sigma = 1.07$ pixels for the 3K image and $\sigma = 0.69$ pixels for the 3K+ one.

### 4.3 *Real-time Performance*

This section focuses on the processing time of the data acquisition and orthorectification processes. Switching to newer and faster hardware can accelerate the on-board processing, but changing hardware of aircraft equipment always involves a complex and protracting certification procedure. Therefore, it makes sense to concentrate on methodic improvements and run tests with a given hardware specification (Tab. 5).

The test setups in Tab. 6 use the same input data but each of the tests runs a different program version. Test A runs in a single thread on one CPU core and takes 13600 ms including memory transfers. It serves as a reference for the following GPU tests. Test B takes 2251 ms for the test image, where 63 ms and 137 ms are spent on the host-to-GPU-memory transfer and on the GPU-to-host-memory transfer, respectively. Here, the real-time constraint of 24 Mbit/s demanded for the burst mode can be held, as the processing speed achieves about 31 Mbit/s.
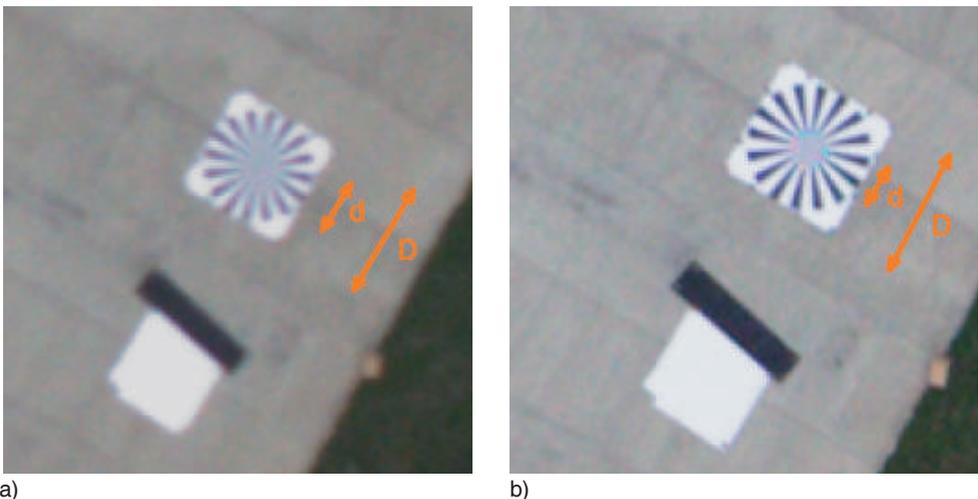


**Fig. 12:** Siemens star and signalized edge in a) 3K image, b) 3K+ image.

**Tab. 5:** On-board hardware and software configuration.

| Hardware | | Software | | Test data | |
|---|---|---|---|---|---|
| CPU | Intel Core 2 Duo E8400 @ 3GHz | Compiler | nvcc release 4.0 | Size input image | 4992 x 3328 pixel |
| RAM | 2.0 GB | Display drivers | NVIDIA Developer Drivers for Linux 270.41.19 | Size output image | 7257 x 5182 pixel |
| Video | Nvidia GeForce 9800 GTX,512 MB, Compute Capability 1.1 | Debugging tools | NVIDIA Compute Visual Profiler 4.0.12, Cuda-gdb release 4.0 | GSD | 0.2 m |
| Main Board | Intel Desktop Board DG35EC | | | DEM resolution | 25 m |

**Tab. 6:** Test cases of different versions of the orthorectification process.

| | Host-to-GPU transfer time (ms) | Kernel execution time (ms) | GPU-to-Host transfer time (ms) | Total time (ms) |
|---|---|---|---|---|
| Test A: CPU | --- | --- | --- | 13600.0 |
| Test B: GPU | 63.3 | 2051.1 | 137.5 | 2251.9 |
| Test C: GPU | 41.5 | 81.4 | 136.6 | 259.5 |
| Test D: GPU | 41.5 | 79.2 | 134.8 | 255.6 |

For further optimization, the kernel execution time is reduced in Tests C and D by avoiding data transfers between host and GPU and by moving constant parameters from the uncached global GPU memory to cached memory areas. For this purpose, the whole input image and the DEM are transferred to the GPU memory before the orthorectification starts. After complete calculation of the output image, it will be transferred back to the host. There are no further host-GPU transfers in between. Test configuration C uses nearest neighbour interpolation, whereas configuration D uses bilinear interpolation. Both run a version which reduces the memory data transfers to minimum and puts all constant parameters required for orthorectification in the cached constant memory area. The usage of a cached memory results in a significant speedup of the kernel execution time by a factor of 25 compared to Test B. When comparing tests C and D, the use of bilinear interpolation has no effect on the performance, which is remarkable given that the processing time

of CPU-based orthorectification programs increases significantly with more complex interpolation methods.

In conclusion, the implementation of Test D is 53 times faster than the CPU implementation and meets the timing constraints with a data rate of 282 Mbit/s. In other words, the cameras can operate with a continuous shooting speed of 3 to 4 frames per second without risking a delay by the orthorectification process.

## 5    Conclusions

In this paper, the performance of low-cost camera systems for real-time mapping applications are exemplarily evaluated based on already existing sensor systems operated at DLR. These sensors are designed for traffic and disaster monitoring as well as for monitoring during mass events and should fulfill certain requirements in terms of cost, weight, and volume. The sensor systems are evaluated

in terms of FOV, resolution, processing time, geometrical accuracy, and radiometric properties.

The sensors are designed to cover wide areas with high resolution. Thus, three cameras with different viewing directions are arranged on a sensor platform. The real-time capacity of the established processing chain is evaluated and tested. Based on this chain, images acquired with a frame rate of 4 frames per second can be read out, orthorectified, and selective images can be sent down to the ground station without delay (i.e. the first image of each camera "burst"). Even traffic parameters can be extracted automatically in real-time on-board (ROSENBAUM et al. 2008). The real-time implementation is based on a concept by (THOMAS et al. 2008), GPU-based orthorectification, and fast image processing libraries.

At two test sites, the sensors were geometrically calibrated by a bundle adjustment with self calibration in the years 2008 and 2010. The results showed that focal length and lens distortion parameters are stable over the years whereas the principal point is weakly determined and, thus, deformations of the camera lens connection can neither be ruled out nor confirmed. The self calibration with additional parameters describes the camera geometry well; the remaining residuals are random with a maximum of 1.5 pixel. The accuracy achieved at the object points were calculated based on GPS-measured check points. The RMSEs range is from 0.14 m to 0.38 m in the XY- and from 0.45 m to 0.76 m in the Z-coordinates for all camera systems. Higher errors in the XY-coordinates of 0.65 m in the 2008 campaign are caused by focusing problems of the 3K camera system.

Besides, the performance of the direct georeferencing based on GPS/Inertial measurements is evaluated. Most crucial is the accuracy of the image attitude determination which highly dominates the accuracy that can be achieved. Differences of reference image attitudes with IMU data show deviations of about 0.05°, which exceeds the basic IMU accuracy. Reasons for this may be deformations between cameras and IMU positions due to shocks, vibrations, and the long distance between sensor and IMU. Finally, based on the maximum deviations encountered the

obtained accuracy is 1.45 m in the nadir and around 3 m at the FOV edges at a flying height of 1000 m above ground. Nevertheless, the geometrical requirements coming from the envisaged real-time applications are fulfilled, except for real-time 3D applications where the relative accuracy has to be improved e.g. by a bundle adjustment.

The radiometric properties of the sensors in terms of resolution and luminosity are improved with the use of Zeiss lenses. Tests are made to find out the best camera settings for highest resolution. Due to the forward motion blurring, short exposure times are recommended, which forces the f-number and film speed to more inappropriate values. Best results were obtained with fixed f-number e.g. 4.0 and shutter speed e.g. 1/2000 s in combination with variable ISO values adjusted to the illumination conditions. Images acquired with higher ISO values during bad weather conditions and during flights below clouds are noisier but still acceptable e.g. for traffic parameter extraction. Even in total darkness, vehicle headlights are visible in airborne images with high ISO values, low f-number, and shutter speed around 1/1000 s.

Further investigations aim to replace the cost-intensive IMU with a real-time software resolution. Besides, additional processors like a DSM generator and a tool for qualitative crowd monitoring as well as people tracking will be developed and included in the real-time processor chain.

# 6 Acknowledgments

# References

COPPINGER, R., 2010: Fuel cell motor-glider basis for endurance UAV. – Flight International **177** (5233): 25.

CRAMER, M., STALLMANN, D. & HAALA, N., 2000: Direct Georeferencing using GPS/Inertial Exterior Orientations for Photogrammetric Applications. – IAPRS, **XXXIII**, Part B3: 198–205, Amsterdam, Netherlands.

CRAMER, M., 2010: The DGPF-Test on Digital Airborne Camera Evaluation – Overview and Test Design. – PFG **2/2010:** 73–82.

ERNST, I., SUJEW, S., THIESSENHUSEN, K., HETSCHER, M., RASSMANN, S. & RUHÉ, M., 2003: LUMOS – airborne traffic monitoring system. – Intelligent Transportation Systems **1:** 753–759.

FRASER, C.S., 1997: Digital camera self-calibration. – ISPRS Journal of Photogrammetry and Remote Sensing **52** (4): 149–159.

HAARBRINK, H. & KOERS, E., 2006: Helicopter UAV for photogrammetry and rapid response. – [online] Proceedings of the 2nd International Workshop: The Future of Remote Sensing, VITO and ISPRS Intercommission Working Group I/V Autonomous Navigation, Antwerp, Belgium.

IGI, 2011: Ingenieur Gesellschaft für Interfaces mbH. – http://www.igi.eu (22.12.2011).

KOZEMPEL, K. & REULKE, R., 2009: Camera Orientation Based on Matching Road Networks. – Image and Vision Computing New Zealand, IVCNZ **'09**: 237–242, 24th International Conference.

KURZ, F., 2009: Accuracy assessment of the DLR 3K camera system. – DGPF Tagungsband **18**, Jahrestagung.

KURZ, F., ROSENBAUM, D., LEITLOFF, J., MEYNBERG, O. & REINARTZ, P., 2011: Real-time camera system for disaster and traffic monitoring. – Proceedings of the International Conference on SMPR (Sensors and Models in Photogrammetry and Remote Sensing) **2011:** Tehran, Iran.

LEITLOFF, J., HINZ, S. & STILLA, U., 2010: Vehicle Detection in Very High Resolution Satellite Images of City Areas. – IEEE Transactions on Geoscience and Remote Sensing **48** (7): 2795–2806.

LEHMANN, F., BERGER, R., BRAUCHLE, J., HEIN, D., MEISSNER, H. & PLESS, S., 2011: MACS – Modular airborne camera system for generating photogrammetric high-resolution products. – PFG **6/2011:** 435–446.

MÜLLER, R., LEHNER, M., MÜLLER, R., REINARTZ, P., SCHROEDER, M. & VOLLMER, B., 2002: A program for direct georeferencing of airborne and spaceborne line scanner images. – Proceedings of the ISPRS Commission I Symposium, Integrating Remote Sensing at the Global, Regional and Local Scale **1:** 148–153, Denver.

NAGAI, M., CHEN, T., SHIBASAKI, R., KUMAGAI, H. & AHMED, A., 2009: UAV-Borne 3-D Mapping System by Multisensor Integration. – IEEE Transactions on Geoscience and Remote Sensing **47** (3): 701–708.

NVIDIA, 2011: CUDA C Programming Guide, version 4.0.

PETRIE, G., 2003: Airborne digital frame cameras. The technology is really improving. – GEOInformatics **2003** (Oct/Nov): 18–27.

ROSENBAUM, D., KURZ, F., THOMAS, U., SURI, S. & REINARTZ, P., 2008: Towards automatic near real-time traffic monitoring with an airborne wide angle camera system. – European Transport Research Review **1** (1): 11–12.

SASA, S., MATSUDA, Y., NAKADATE, M. & ISHIKAWA, K., 2008: Ongoing research on disaster monitoring UAV at JAXA's Aviation Program Group. – SICE Annual Conference **2008:** 978–981.

SCHARDT, M., RAGGAM, H., WACK, R., OFNER, M., GUTJAHR, K. & KOUDELKA, O., 2010: Luftgestützte Low-Cost-Aufnahmeplattform zur Unterstützung von Katastropheneinsätzen. – Dreiländertagung **2010:** Wien, Austria.

SHASTRY, A.C. & SCHOWENGERDT, R.A., 2005: Airborne video registration and traffic-flow parameter estimation. – Intelligent Transportation Systems, IEEE Transactions **6** (4): 391–405.

SIRMACEK, B. & REINARTZ, P., 2011: Automatic crowd density and motion analysis in airborne image sequences based on a probabilistic framework. – Proceedings of the 2nd IEEE ICCV Workshop on Analysis and Retrieval and Tracked Events and Motion in Imagery Streams (ARTEMIS'11) **2011:** Barcelona, Spain.

SRINIVASAN, S. & LATCHMAN, H., 2004: Airborne traffic surveillance systems: video surveillance of highway traffic. – The 2nd ACM international workshop on Video surveillance & sensor networks, 131–135.

SRS, 2011: Schulze Radio Systems. – http://www.srsw.de (22.12.2011).

THOMAS, U., ROSENBAUM, D., KURZ, F., SURI, S. & REINARTZ, P., 2009: A new Software/Hardware Architecture for Real-Time Image Processing of Wide Area Airborne Camera Images. – Real-Time Image Processing Journal **4:** 229–244.

ZONGJIAN, L., 2008: UAV for mapping – low altitude photogrammetric survey. – IAPRS **XXXVII** (B1), Beijing 2008.

ZHU, K., D'ANGELO, P. & BUTENUTH, M., 2010: Comparison of Dense Stereo using CUDA. – 11th European Conference on Computer Vision (ECCV), Kreta, Griechenland.

Addresses of the authors:

Dr.-Ing. FRANZ KURZ, Dipl.-Ing. SEBASTIAN TÜRMER, Dipl.-Ing. OLIVER MEYNBERG, Dr. rer. nat. DOMINIK ROSENBAUM, Dipl.-Ing. HARTMUT RUNGE, Prof. Dr.-Ing. PETER REINARTZ, German Aerospace Center,

Remote Sensing Technology, Münchner Straße 20, 82234 Wessling, Germany, Tel.: +49-8153-28-2764, -3600, -3534, -2678, -1364, -2757, Fax: +49-8153-281444, e-mail: firstname.lastname@dlr.de (ü = ue)

Dr.-Ing. JENS LEITLOFF, Karlsruhe Institute of Technology, Institute of Photogrammetry and Remote Sensing, Englerstraße 7, 76131 Karlsruhe, Germany, Tel.: +49-721-608-47410, e-mail: jens.leitloff@kit.edu