THE ULTRACAM STORY

Franz Leberl
Graz University of Technology, A-8010 Graz [Austria]. leberl@icg.tugraz.at

Michael Gruber, Martin Ponticelli, Alexander Wiechert
Microsoft, Anzengruberstrasse 8, A-8010 Graz [Austria], [michgrub | nponti | alwieche]@microsoft.com

Commission I, WG I/2 - LIDAR, SAR and Optical Sensors for Airborne and Spaceborne Platforms

KEY WORDS: Photogrammetry, Camera, CCD, Aerial, High Resolution

ABSTRACT

The UltraCam-project created a novel Large Format Digital Aerial Camera. It was inspired by the ISPRS Congress 2000 in Amsterdam. The search for a promising imaging idea succeeded in May 2001, defining a tiling approach with multiple lenses and multiple area CCD arrays to assemble a seamless and geometrically stable monolithic photogrammetric aerial large format image. First resources were spent on the project in September 2011. The initial UltraCam-D was announced and demonstrated in May 2003. By now the imaging principle has resulted in a 4th generation UltraCam Eagle, increasing the original swath width from 11,500 pixels to beyond 20,000. Inspired by the original imaging principle, alternatives have been investigated, and the UltraCam-G carries the swath width even further, namely to a frame image with nearly 30,000 pixels, however, with a modified tiling concept and optimized for orthophoto production.

We explain the advent of digital aerial large format imaging and how it benefits from improvements in computing technology to cope with data flows at a rate of 3 Gigabits per second and a need to deal with Terabytes of imagery within a single aerial sortie. We also address the many benefits of a transition to a fully digital workflow with a paradigm shift away from minimizing a project’s number of aerial photographs and towards maximizing the automation of photogrammetric workflows by means of high redundancy imaging strategies. The instant gratification from near-real-time aerial triangulations and dense image matching has led to a reassessment of the value of photogrammetric point clouds to successfully compete with direct point cloud measurements by LiDAR.

1. 2000-2001: DECIDING ON AN AERIAL CAMERA

1.1 Why Start a Camera R&D Project?
Vexcel Imaging GmbH in Graz (Austria, founded 1993) initially manufactured, marketed and sold over 400 photogrammetric film scanning systems UltraScan5000 from 1999 to 2005. By 2000 financial resources existed to add new products. At the ISPRS Congress (Amsterdam) first digital aerial mapping cameras with a large format were announced. Scrutiny of these products and an assessment of opportunities from all-digital photogrammetry were the basis for deciding in favor of developing a digital camera.

1.2 Assessing the 2000-Technologies
A review of patent applications reveals a plethora of ideas about digital photogrammetric imaging systems. We want to consider only technologies capable of producing “large format” images to replace photogrammetric aerial film.

The Leica ADS is a tri-linear scanning system analogous to the proposal by Derenyi (1970; see also Konecny, 1970), then describing a film based strip camera, and later implemented in a digital format by the German Aerospace Agency DLR (Neukum et al., 2001). Leica had purchased that technology. The initial product description was by Sandau et al. (2000), then Fricker (2001). Limitations were (i) much increased costs vis-à-vis film cameras, (ii) abandonment of the frame photograph and thus of the familiar workflows, (iii) uncertainty about achieving a photogrammetric accuracy at 1/10⁸ of a pixel, and perhaps (iv) constraining data acquisition to a mere 3 images per terrain point and foregoing the opportunity of high overlaps and redundancy.

The Zeiss-Intergraph DMC produced a large image format by assembling 4 separate image tiles that get imaged simultaneously through 4 separate optical cones and merged by software. This preserves the frame photo paradigm and thus the traditional photogrammetry workflows. Hinz et al. (2000 and 2001) described the approach and it was quickly evident that (i) it lacked software leverage and thus needed high cost of goods; (ii) the geometry relied on mechanical stability, not on redundancy and overlaps; (iii) the radiometric performance would suffer from a limited analog-to-digital conversion and high pan-sharpening ratio.

The INPHO ISPRS-2000 Camera was presented at the ISPRS-2000 congress with little technical detail, except for a patent reference to Teuchert & Mayr (2000). This development was abandoned in 2001. From it emerged a European multi-entity CONPIE GmbH under INPHO-leadership (Mayr & Ohlhoff, 2004). A product never materialized. Thus only two vendors offered large format digital aerial systems, yet deliveries were delayed. Computer technology still was bulky for very high in-flight data-rates and for operating reliably on non-pressurized survey airplanes. The collected data sets seemed forbiddingly large for space- and resource-constrained survey operations both in the air and in small photogrammetry offices.

1.3 Inventing a Novel Technology
Vexcel found it worthwhile to search for a novel digital large format camera concept that would (i) produce images as similar to scanned aerial film as possible, (ii) offer software leverage to support an end user price not higher than the traditional film cameras [thus at half the price for competing systems] and (iii)
assure geometric and radiometric excellence by clever design rather than mechanical expenses. The search was successful, patent protection was applied for in May 2001 and a first employee started full-time on the UltraCam-project in September 2001. The basic imaging principle is illustrated in Figures 1 and 2 and various publications (Leberl et al., 2003, 2005; Gruber & Leberl, 2005, 2010; Gruber et al., 2003, 2008, 2012; Leberl & Gruber, 2003; 2005 a, b; Wiechert et al., 2011).

Figure 1: The core idea of the UltraCam-technology is the assembly of one large aerial image from 9 separate tiles imaged through 4 separate optical cones numbered 1 through 4 [at right]. At left is a view of the 4 linearly arranged optical cones for the high resolution panchromatic image segments and the 4 additional optical cones for the lower resolution color channels. The resulting large format image has 5 bands with pan-infrared-red-green-blue. The pixel numbers at right concern the UltraCam-Xp model with 17.3K x 11.3K.

Figure 2: Stitching image segments into one large image is a core skill being accomplished by one of two approaches. Above is the initial assembly based on the overlaps within the 4 separate panchromatic image segments. Below is a simultaneous adjustment of all image tiles using their overlaps as well as the green color-channel. A complex consideration concerns the effects of temperature on the deformation of the individual cones. Key publications about stitching are by Gruber & Ladstädter (2006, 2012) and Ladstädter et al. (2010).

2. ULTRACAM-REQUIREMENTS

2.1 Users Continue with Existing Workflows
To ease the user from film into the all-digital workflows, images were to be as similar to scanned aerial film as possible. The then-established digital processing workflows were to get continued, avoiding add-on training and new photogrammetry software.

2.2 Geometric Accuracy at ±1 μm
In analogy to film, the geometric accuracy needed to be high and stable over time at ~ ±1 μm. Two considerations were important. First, assembling tiles into a large image needed to rely on overlaps to support a “triangulation” with good redundancy. Secondly, the individual image tiles needed to be taken with identical perspectives. The so-called “syntopic” imaging system [“syn” = same, “topos” = place] ensured that these conditions were met. The idea is illustrated in Figure 3.

2.3 Superior Radiometry
Digital radiometry can outperform film. At issue is one’s ability to maintain a 72 dB signal-to-noise-ratio. This will result in nearly 14 bit of radiometric range and a need to assign to each color channel a 16-bit number. And there is no grain noise.

2.4 Simultaneous Infrared for True and False Color
The film-to-digital transition substitutes the 3 film-colors red-green-blue by 4 digital colors red-green-blue- near infrared. Considerable automated image analysis work is enabled by the additional infrared channel, especially as concerns vegetation.

2.5 Digital Systems at the Film Camera Cost Level
The photogrammetric services industry operated on business plans that include the depreciation of the film camera and the film processing costs, all to be earned via aerial flight projects. Digital cameras already can be expected to become rapidly obsolete along with all things digital. Therefore the depreciation and maintenance of a digital aerial camera must be financed by the savings on film and film processing.

A reliance on software can produce a so-called “software-leverage” to reduce the cost-of-goods in the new UltraCam-technology. Functions previously achieved by precision mechanics and optics must be taken over by software and elaborate calibration techniques [Krüpfel et al., 2004]. This can reduce manufacturing and material costs.

3. 2003: ULTRACAM-D INTRODUCTION

3.1 Sufficiently Advanced Computer Technology
Large image formats and high repeat rates result in data rates at 3 Gigabits per second and image data at initially 350 Mbytes per photo. Desirable image overlaps produced thousands of photos where previously only hundreds were taken. By 2003 the TByte had become manageable in a modern photogrammetric office and this matched the needs of a digital aerial camera. It may well have been for this reason that 2003 became the kick-off year for all digital aerial camera installations.
3.2 The “Film-is-Dead” Syndrome
The initial camera introductions in 2000 were timid. It was still unclear why one should abandon film. But by 2003 we could see and argue the end of aerial film. This was greatly doubted, even considered overly brash. However, the many advantages of digital over film were increasingly recognized. As seen from today’s perspective, was a 2003-message of “film is dead” the right signal?

3.3 The R&D Project and Schedule
The UltraCam was developed by a team that grew from 1 person in September 2001 to ~10 persons by 2003. As much work was contracted out as possible, however, with a strong concern that all intellectual property on electronics, optics, mechanics and software would remain in-house. The core in-house teams were in electronics and software.

Manufacturing was entirely contracted out to Wild in Austria. Optics developments were initially accomplished by Schneider and later by the Rodenstock-spinout LINOS. A flying product and full operation were achieved by April 2003, thus within 18 months and after an investment of about 12 person-years.

4. MARKET ACCEPTANCE & FOLLOW-ON PRODUCTS

4.1 Number of Cameras Sold
Immediately upon the product announcement in May 2003, Vexcel Imaging received camera orders. The market was eager for a product meeting the requirements specified above. Customers trusted that the technology would be valid and that the company would come through on the product specifications. Deliveries started in January 2004. The growth of the install base is shown in Figure 4. The market share today is at 45% of all digital mapping cameras sold globally [see Figure 5].

![Figure 4: Global UltraCam installations](image)

As shown in Figure 6, the growth of the image size from initially 11,500 pixels across the flight path to 20,000 pixels in the most recent product UltraCam Eagle is taking advantage of CCD technology which advanced from a pitch at 9 μm to 5.2 μm.

Additionally, computer technology also helped to miniaturize the product by going to solid state image memories and integrating the GPS and IMU-functions into the camera. Much of this is a result of faster and smaller electronics. Total weight of the UltraCam-Eagle at 75 kg is nearly half of the UltraCam-D, and so is the size, going from 2 separate boxes to just one [Was: one camera, one image storage. Now: single integrated device].

4.2 From UltraCam-D to UltraCam-Eagle
Figures 6 and 7 are illustrations of the product and technology evolution since 2003. The basic assembly of large images from tiles is maintained. A core skill is to understand how to achieve sub-micrometer accuracies in a large image fused from tiles, invalidating the need to work with large CCD arrays.

![Figure 6: The image format increased from 2003 [UltraCam-D, 11.5 K pixels across the swath] to 2011 [UltraCam-Eagle, 20K pixels across the swath]. The UltraCam-Eagle is the first solution with exchangeable lenses so that focal lengths can be varied in the field.](image)

![Figure 7: The UltraCam-family of products and their evolution over time.](image)

4.3 The UltraCam G and an “Orthophoto-Robot”
Microsoft has a need for orthophotos inside its BING-Maps location-aware Internet search engine. The continental USA with its 7.5 million km² should be mapped in ~3 years at 30 cm pixel-orthos and at an operating cost preferably at ~USD 1 per km².
This global orthophoto may be the first transnational aerial mapping at that resolution (Gruber et al., 2012). Part of this project was the development of the UltraCam-G with a swath size of nearly 30,000 pixels. The Microsoft-internal workflow is automated to such a high level that it may well be called an “orthophoto robot”.

4.3 Software and Workflow-Products
At Vexcel, there existed in 2005/6 a development rush towards a full photogrammetric workflow software system. This was halted by Microsoft when it stepped in as new owner. Instead, the software development was refocused towards an in-house workflow from the raw imagery to the finished 3D urban model and orthophotos in support of the BING location needs.

With time, Microsoft-policy ultimately changed and novel photogrammetry workflows became deliverable products. Fully automated aerial triangulation of highly overlapping blocks of thousands of aerial photos became available in the form of the UltraMap-AT software (Gruber & Reitinger, 2008). The capability to interact visually with large blocks of images was taking advantage of the Seadragon software independently available through other Microsoft-channels (Reitinger et al., 2008).

The release of the existing in-house dense matching system as a product has been announced by Microsoft in 2011, so that finally, the pre-2006 Vexcel-ambitions do become a reality. Figure 8 illustrates the roadmap as it stands in 2012.

![Software Roadmap](image_url)

**Figure 8:**
The original ground based post-processing software was denoted as “OPC” [Office Processing Center]. With the advent of Microsoft’s Seadragon capability, this morphed into the UltraMap system. The augmentation by a dense matching product is on its way.

5. EFFECTS ON PHOTOGRAMMETRIC WORKFLOWS

5.1 Increasing Image Overlaps
An essential paradigm shift in aerial and terrestrial mapping results from vanishing variable costs per image. Along a flight line, increasing the overlap from traditional 60% to novel 90% is absolutely cost-free, but increases the number of images by a factor 4. Across the flight direction, additional flight lines are needed to go from 20% to 60% side-lap, but the associated cost is only for air time, not for the imagery itself. Going from 20% to 60% increases the number of images again by a factor of 2.

Transiting from 60/20 to 90/60 provides an increase in redundancy by nearly an order of magnitude. To use this increase the workflows must be sufficiently automated to eliminate any variable labor per image.

5.2 Smaller Ground Sampling Distances
Since the number of images no longer affects the project cost, one can think differently about pixel size. The choice of a ground sampling distance GSD is driven by the need to map certain detail, and also to ensure that automation succeeds. In urban mapping one might have relied on scanned film with a GSD at perhaps 20 cm. With digital cameras one might want to go to 8 cm to obtain flawless dense digital surface models with clearly delineated roof lines, gutters, chimneys, skylights, sidewalks and street details.

A film project for a 150 km² urban area at 20 cm pixels had previously been done with 150 aerial film images at 60/20 overlaps. It now may be done at 90/60 overlaps at a pixel size of 10 cm, resulting in 3,000 digital images.

5.3 Full Automation
The age-old dream of fully automated aerial triangulation, orthophoto- and DSM-production has become a reality due to higher overlaps and smaller pixel sizes. High overlaps support a “super-resolution” since every terrain point is imaged 10 times and this redundancy removes the noise in the resulting measurements. A DSM no longer is created at postings at 10 times the pixel size, but now at a single pixel spacing. The idea of “dense matching” has arrived (Klaus, 2007).

5.4 Color and False Color Infrared
A stark difference between traditional film and current digital imaging is the ubiquitous availability at no variable or fixed extra cost of an infrared image channel. This has led to strong advances in 3D mapping, which has a perennial issue with occlusions by vegetation. Vegetation can be mapped with infrared color.

5.5 Urban 3D Modeling
The crux of urban 3D-modeling is the cost per building, and the level of detail. Dense matching at each pixel and semantic interpretation of scene contents make 3D building models affordable and cause a surge in interest and applications. As stated before, full automation of façade- and roof-maps relies on high image overlaps, use of simultaneous 4 or 5-color bands and small pixel sizes (Zebedin et al., 2006; Leberl et al., 2012; Bimber, 2010).

6. PHOTOGRAMMETRY AND MICROSOFT

The UltraCam story would be incomplete without a view of the effect of Microsoft on photogrammetry and vice-versa.

6.1 BING Maps
Microsoft’s mapping history started in ~ 1995. Initial products were Streets & Trips, later Map-Point, based on street vector data. Mapping-related projects included the TerraServer (undated), Encarta (undated) and flight simulation. The advent of the Internet has broadened the applicability of Map-Point to travel planning and directions, and GPS positioning resulted in the applicability of mapping data to navigation. However, Microsoft’s personnel would not associate itself with photogrammetry. The association was strictly with computer vision and computer graphics.
By 2005, images started to augment the street vectors, and 3D information began to support search. This marked the transition from Map-Point to Virtual Earth with the idea of searching the Internet in a location-aware manner and possibly in an AR [Augmented-Reality] environment. Virtual Earth was later renamed BING-Maps.

Microsoft’s approach to address this new requirement for geo-imagery was, in analogy to Google’s, to acquire an existing company so that the skill sets missing at Microsoft in the image-based mapping arena would be added externally. Microsoft’s choice was Vexcel Corporation and Vexcel Imaging GmbH. Today, the Colorado-location of Microsoft in Boulder has morphed into BING Imagery Technologies and is the main site for BING-Maps and the Global Ortho Project.

High quality aerial photography is essential to Microsoft’s automated workflows for 2D and 3D mapping data. It is for this reason that – contrary to pundit’s opinions [see Petrie, 2006] – Microsoft has maintained its interest in aerial cameras for over 5 years and continues to support the evolution of digital camera technology and products at its Austrian site. Additionally, this Austrian site also delivers computer vision components and algorithms.

6.2 Photosynth, Neo-Photogrammetry and MAV-Imagery
Photosynth had been a computer vision development via the cooperation between the University of Washington computer vision team and Microsoft Research (Snively et al., 2006). It has been developed into an operational service available globally (Agüera y Arcas, 2007). The value of Photosynth has been recognized, and organizationally it has been moved into the Colorado Bing-Imagery Technologies team.

Photosynth is at the heart of a development towards urban mapping by means of amateur photography. The designation “neo-photogrammetry” describes the novel phenomenon of citizen’s amateur digital photography, accompanied by GPS-based positioning, uploaded to an Internet website and serving as the basis for 3D mapping (Leberl, 2010a, b).

The UltraMap-development into fully automated aerial triangulation, dense matching and orthophoto production is applicable to neo-photogrammetric approaches. Additionally, the same high-overlaps-mapping paradigm is applicable to novel Micro-Aviation Vehicles MAVs as an additional path into the future of image-based mapping (Wendel et al., 2012).

6.3 Microsoft Research
The acquisition of Vexcel by Microsoft has resulted in the awareness at Microsoft how photogrammetry skills differ from computer science and computer vision. At the same time, access to Microsoft’s research and development resources can greatly inspire progress and innovation. First examples have already become visible in the form of Seadragon and Photosynth (Reitinger et al., 2008). The evolution of neo-photogrammetric approaches in an Internet-context hold great promise for a much enhanced role for photogrammetry.

7. CONCLUSIONS -- INNOVATION NEEDS

The UltraCam digital aerial camera project has developed into a great technology and commercial success. It has added Microsoft as a new industrial player to the photogrammetric industry. The achievement depended on mastering the art of tiling a large image from smaller segments collected with multiple lens cones. A second factor is the reliance on software in lieu of precision mechanics, best described as software-leveraged hardware. This basic idea, as simple as it may seem, required an innovative approach to geometric calibration at sub-micrometer accuracies. We have listed numerous other innovations that we consider a prerequisite for the UltraCam success, ranging from high signal-to-noise performance to real-time on-board software.

What had started out as an innovation in cameras now represents innovations in photogrammetric procedures from full automation via a vast increase in image overlaps and redundancies. One thereby abandons “two-eye-stereo” by “multi-ray” geometries. Arbitrarily configured sets of images will be triangulated without any manual labor. Dense matching at every pixel produces a super-resolved surface model not only fully automatically, but also at an increased quality. Resulting data products can be traditional orthophotos, but also more interesting data sets such as interpreted computer models of the human habitat.

At issue is continued progress towards more automation, towards broadening the image sources to include neo-photogrammetric amateur imagery, augmented reality Internet-based applications and miniaturization towards the ubiquitous use of handheld devices and smart phones. Image blocks will increase to millions of images per project and with instant updates of urban models, along the “Rome-in-a-day” paradigm (Frahm et al., 2010).

The Internet in combination with the smart telephone creates the new infrastructure for photogrammetric work, sensing, data and applications. We are at the verge for photogrammetry to deliver information about our habitat to everyone at all times and in all places.

ACKNOWLEDGEMENT
We are grateful for Figures 4 to 8 – they have been made available by Engelbert Breg from Microsoft-Graz.

REFERENCES


Terraserver (undated) http://en.wikipedia.org/wiki/Terraserver.com

