Leberl F, M. Gruber . New Photogrammetry Enabled by Digital Large Format Aerial Cameras.
Proc. of SS C2005 Spatial Intelligence, Innovation and Praxis: The national biennial Conference of the Spatial Sciences Institute.
September 2005. Melbourne: Spatial Sciences Institute. ISBN 0-9581366-2-9

NEW PHOTOGRAMMETRY ENABLED BY DIGITAL LARGE FORMAT AERIAL CAMERAS

Franz W. Leberl[1] , M. Gruber[2]
1) Institute for Computer Graphics and Vision, Graz University of Technology
Inffeldgasse 16, A-8010 Graz, Austria, Leberl@icg.tu-graz.ac.at, www.icg.tu-graz.ac.at

2) Vexcel Imaging, Münzgrabenstrasse 11, A-8010 Graz, Austria

KEY WORDS: Stereo photogrammetry, multi-image computer vision, photo-realism, 3-dimensional terrain models, virtual habitat.

ABSTRACT

The advent of the large format digital aerial camera is initially seen as a straight forward replacement of two of photogrammetry's defining pieces of equipment: the film camera and the precision scanner. The transition from film to digital is to be justified simply by comparing the resulting imagery: here the scanned film image, there the digitally sensed image. This paper challenges this narrow view and argues that digital cameras are transforming the field of photogrammetry in a fundamental way by replacing the 2-image stereo-paradigm by multi-image computer vision. This permits the creation of new products using new procedures that support a high level of automation. This not only innovates the technology, but causes changes of the constituents of the field of photogrammetry.

BIOGRAPHICAL SKETCH OF DR. FRANZ LEBERL

Is professor of computer science at Graz University of Technology (since 1992). He also is founder and shareholder of Vexcel Corporation (USA, 1985) and Vexcel Imaging (Austria, 1993), and the motor behind the development of the large format digital aerial camera at Vexcel. His career is in its 37th year, starting at the ITC in the Netherlands in 1969 with some assignments in Bogotá (Colombia). His work addressed academia and business in the USA and Austria. His diploma (1967) and doctorate (1972) are in geodetic engineering/applied science, Vienna University of Technology. Most recently he was President of Commission IIII (Theory and Algorithms, 2000-2004) of the International Society for Photogrammetry & Remote Sensing (ISPRS).

1. "NEW PHOTOGRAMMETRY"?

1.1 What is "New"?

"New" is a photogrammetry that

- focuses on the creation of 3-dimensional rather than 2- or 2.5- dimensional spatial data of terrain,
- creates photo-realistic models of the terrain and the objects on its surface rather than just their symbolized interpretation, and
- achieves its results inexpensively by a high degree of automation instead of manual or machine-supported manual work.

---

[1] Also CEO of Vexcel Imaging GmbH, Graz, Austria, manufacturer of the UltraCam-D large format digital aerial camera system

Figure 1 is one type of a "new" photogrammetric product combining aerial and terrestrial source data. The output is fully 3-dimensional and photorealistic . Given the advances in computer graphics, visual information handling, computer-generated entertainment and computing, the transition from 2-D or 2.5 D to fully 3-dimensional models of the human habitat is to be expected. Full automation, however, has long been an elusive goal of photogrammetry; that it should be feasible as a result of a "new" approach to photogrammetry requires some explanation.

We argue that the new large format digital aerial cameras are a major reason that automation is now within photogrammetry's reach. For this we replace the idea of the stereo model as the fundamental source of 3-D information by the multi-image assembly. A "new" photogrammetry is thus also a result of

- abandoning the 2-image paradigm, which was dictated by human vision based on two eyes, and replacing it by the multi-eye capability of a computer vision system.

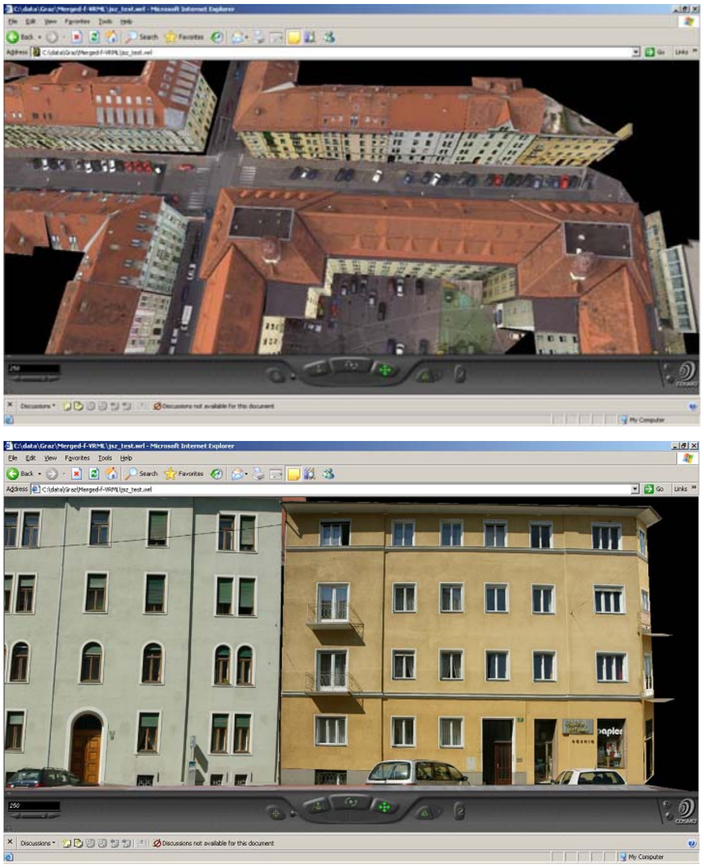


Figure 1:
Above is a 3D photo-textured model strictly from aerial UltraCam imagery of a section of Graz, 8 cm pixels. Below is a close-up from a terrestrial vantage point, and the terrestrial imagery is now being called up for the photo texture (Courtesy J. Szabo, Vexcel Corp.).

This leads us to speak about "paradigm shifts" in photogrammetry. Previous advocates of a "paradigm shift" saw the laser scanner as the core innovation justifying a "new photogrammetry" (Gruber et al., 2003). We argue that the multi-image assembly of the all-digital workflow will obviate the laser scanner and therefore is a much more fundamental justification of a new paradigm in photogrammetry.

1.2 Do we need a "new photogrammetry"?

Considering the decline of the discipline in recent years, as documented by the declining number of participants, conferences and attendance at these conferences, and the closing of academic centers, an invigorating innovation is needed to redefine “photogrammetry” vis-à-vis the emerging field of computer vision.

2. ABANDONING THE MINIMIZATION OF A PROJECT’S NUMBER OF IMAGES

2.1 Project Costs as a Linear Function of the Number of Images?

Minimizing the number of photographs for a given project was a major ambition since the inception of photogrammetry in the late 1800’s. The reason? A project’s cost was directly proportional to the number of images to be flown, film to be purchased, processed in a photo-laboratory, perhaps scanned into a digital format, and finally manually converted into information by photogrammetric procedures. The costs for an Aerotriangulation, the DEM generation and Ortho-photo processing were a linear function of the number of images in a project, and were quadratically growing with the scale number and the resulting number of flight lines. .

2.2 Digital Cameras Are Taking Over Quickly

Digital large format aerial cameras now are rapidly taken over from film cameras. A recent survey indicates that by mid-2005, the number of novel large format digital cameras in use was 75, as compared to 50 at the end of 2004, 16 at the end of 2003 and 6 in 2002. By the end of 2006 digital cameras may produce more images than all film cameras taken together. Concerned with the UltraCam-D, we are very aware of the rapid changes in the market (Leberl & Gruber, 2003 and 2005).

This rapid acceptance of digital cameras results from an increase in image quality and from the economic advantages over the continued use of film cameras, simply because there no longer exists the need to purchase film, develop it and scan it into pixel arrays. The result is a dynamic towards the transition from film-based to fully-digital workflows. However, the most important change is not the savings on film and scanning. Instead, the most significant change vis-à-vis film is the fact that the individual digital image does not produce any variable cost. Therefore the operation of a digital aerial camera is feasible with fixed annual costs only, independent of the number of images being produced, be this 1,000 or 100,000 images per year. Only the costs for the use of the aerial camera platform are variable.

2.3 Increasing the Number of Images Taken for a Project

The changed cost structure of a digital camera results in the new option to vastly increase the number of images per project by selecting a much greater overlap within each flight line, and to fly more flight lines with increased side-laps. Additionally, missions could be flown at lower flying height producing smaller pixels (larger scales) without the quadratric increase in project costs associated with film. In fact, costs may not even increase linearly. In addition, digital images can be produced with far less noise and with sharper edges than film. Therefore the prevailing goal of aerial photogrammetry’s project design could be changed from minimizing the number of images to instead minimize the costs of a project.

2.4 Automation as a Major Driver towards a New Photogrammetry

The result is an opportunity to automate photogrammetric procedures that previously could not be efficiently automated. Failures to automate had two main reasons: first, the film’s quality limitations because of grain noise and a limited dynamic range; second the film’s economy preventing the creation of sufficiently redundant image blocks to properly support automation.

Of course, producing a larger number of images only makes sense if subsequent processing efforts of those images also are independent of their number. Therefore automation is the prerequisite for a free choice of image numbers, as the free choice of image numbers is a prerequisite for successful automation.

2.5 Towards New Photogrammetry Software

The paper introduces an initiative to create a fully automated photogrammetric workflow with automated aerial triangulation using a vast redundancy of 1000 or more tie points in every image, the creation of a DEM exploiting the 80% to 90% forward overlap and 30% to 80% side-lap to look in between vertical objects, and the orthophoto in the form of a 3D photo-texture placed on top of the 3D object surface. We denote this effort as “New Photogrammetry Initiative” NPI. 3D vectors from a land use classification and from image analysis support many of the automation tasks and refine the DEM and 3D orthophoto. In addition they are the basis of machine support for the collection of 3D GIS input data. We document first results from the automated workflow to support the expectation that digital sensor-based automation will reduce labor by a factor of 2.

3. AT THE CORE: THE WEB-ENABLED PHOTOGRAMMETRIC SERVER

3.1 More Images per Year Require a “System” to Manage the Images

Quintupling the number of images per project is to be expected under the new rules. Instead of a 60% forward overlap, 80% could well become the standard. Instead of 20% sidelap, 60% could become the standard. In addition, the restrictions on scale may be dropped and flying low, under the clouds, may become acceptable if the number of images is an irrelevant factor.

When an organization were to change to these new rules, its annual “harvest” of images would grow by a factor of perhaps 10. To eliminate the labor associated with the number of images it ids necessary to automate the handling of images. The photolab, the film archive, the storage, duplication and preservation need to be taken over by a new system: the “Server”.

3.2 The Definition of the “Photogrammetric Server”

Figure 2 illustrates the example of the UltraMap Server by Vexcel. It receives all digital images from a project, most likely in their raw form before they are post-processed into their final deliverable format, equivalent to the scanned film format. The Server processes the raw images into the desired format, and in doing so it creates meta data for a catalog. The meta-data will typically consist of quick-views, small versions of each image, and the associated data from the camera position, attitude, mission parameters etc.

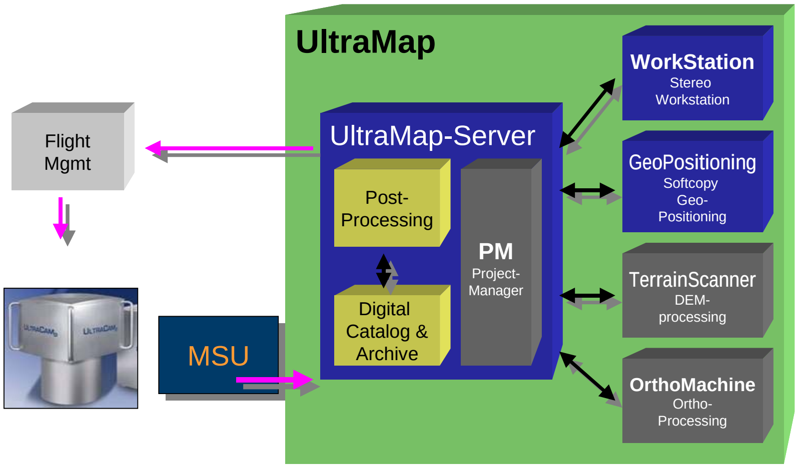


Figure 2:
The UltraMap Server concept of Vexcel Corporation (Colorado). It serves as a catalog based on meta-data of imagery, an archive using robotic storage systems, an image processing center with project management capabilities (Courtesy W. Walcher, Vexcel Corp.).

The meta data go into a “catalog” which is a data base which can be searched by project, by image block, by coverage, by flight line etc. Additionally, the images go into an “archive” at full resolution. The archive administrates all images available to the Server, be they accessible in real time on some disks, in near-real time on some robotic storage medium, or off-line in some disk or tape repository. Keeping the archive protected, maintaining the necessary redundancy and refreshing the storage media is an inherent function of the Server.

Finally, the Server is also a project manager and multi-processor computer. Photogrammetric image processing tasks can be performed on the Server's computers, and the required images are being provided by the catalog/archive structure. The Server thus replaces the traditional arrangement of office PCs, or manages these PCs to perform a task.

The Server is internet-enabled and is accessible via Internet browsers. Therefore a Server can support the worldwide access to the catalog to authorized users.

3.3 A Multi-Functional System

As described, the Server replaces:

- The film archive,
- The photo lab,
- The various photogrammetric workstations,
- The order fulfillment system,
- The IT-infrastructure of a traditional photogrammetry operation.

The Server becomes the backbone of all things happening in a photogrammetric operation. The all-digital workflow simplifies the diversity of systems and equipment, but creates the need for something new, something that was not part of traditional film-based digital photogrammetry.

4. AERIAL TRIANGULATION VERSUS DIRECT GEOPOSITIONING

4.1 Accuracies

Table 1 summarizes the σ_0 – values obtained from a series of aerial triangulations with both digitally sensed images and from scanned film. The values in the range of ± 1μm to ± 2 μm are superior to values obtained with typical aerial film. Those uncertainties in the images propagate into very high accuracies on the ground, particularly when using 80% forward overlap. Table 2 compares errors on the ground with 60% and 80% forward overlap. This high overlap "stiffens" the image-strip and -block beyond the stiffness of a film-strip at 60% overlap.

The reasons for this high accuracy is the improved image quality that is a result of the complete absence of any grain nose, and of the availability of more than 4,000 gray values.

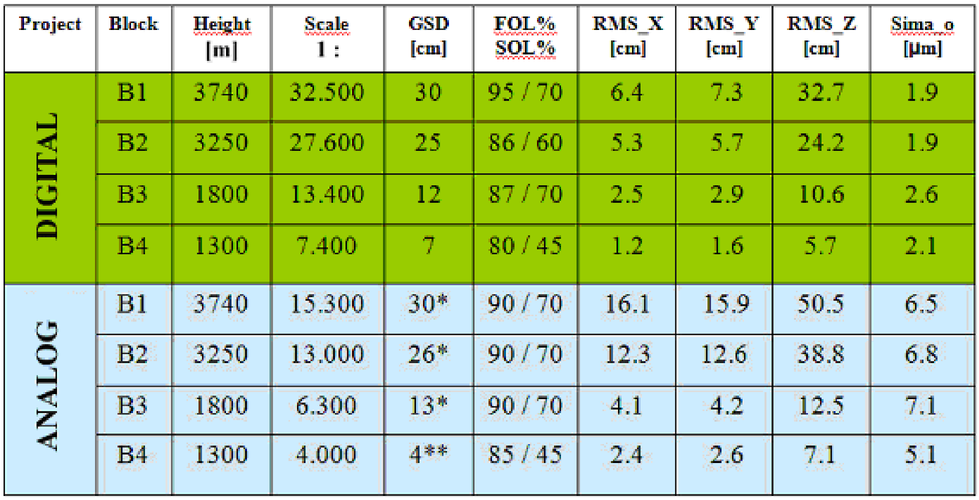
4.2 Automation

The important point is less the accuracy of the AT, but the ability to fully automate the process. The expectation is that a high forward overlap reduces instances of failed matches because the sequential images are more similar than pairs where the overlap is at 60%. A strip of images will thus be connected by matches of sequentially overlapping adjacent images. Adjacent strips get connected with greater success if the overlap is no longer 20%, but increased to 40 or 60%.

Automation is also more likely to be automated if the redundancy of the used tie points gets increased. Figure 3 illustrates a typical 80/60 image block with 1000 (new) points per image, and thus 350,000 tie points in a block with 350 images.

4.3 Support from Direct Geopositioning Systems

The benefits from the use of an inertial system for direct geopositioning include the "instant gratification" by attaching an exterior orientation to each image without much further work, and thus circumventing the aerial triangulation altogether. A second benefit accrues in the case where an aerial triangulation gets performed and the number of required ground control points gets reduced by virtue of the direct geopositioning measurements.



Project	Block	Height [m]	Scale 1 :	GSD [cm]	FOL% SOL%	RMS_X [cm]	RMS_Y [cm]	RMS_Z [cm]	Sima_o [µm]
DIGITAL	B1	3740	32.500	30	95 / 70	6.4	7.3	32.7	1.9
	B2	3250	27.600	25	86 / 60	5.3	5.7	24.2	1.9
	B3	1800	13.400	12	87 / 70	2.5	2.9	10.6	2.6
	B4	1300	7.400	7	80 / 45	1.2	1.6	5.7	2.1
ANALOG	B1	3740	15.300	30*	90 / 70	16.1	15.9	50.5	6.5
	B2	3250	13.000	26*	90 / 70	12.3	12.6	38.8	6.8
	B3	1800	6.300	13*	90 / 70	4.1	4.2	12.5	7.1
	B4	1300	4.000	4**	85 / 45	2.4	2.6	7.1	5.1

Table 1:
AT- results with UltraCam image blocks and analog film images (by Dr. Richard Ladstätter, Institute for Remote Sensing and Photogrammetry, TU Graz). The analog film camera at f=21 cm and UltraCam-D were operated simultaneously in a twin hole aircraft. Flight missions by the Federal Office of Surveying and Mapping (BEV), Vienna, Austria (Leberl & Gruber, 2005). (*) Scan at 20 µm. (**) scan at 10 µm.

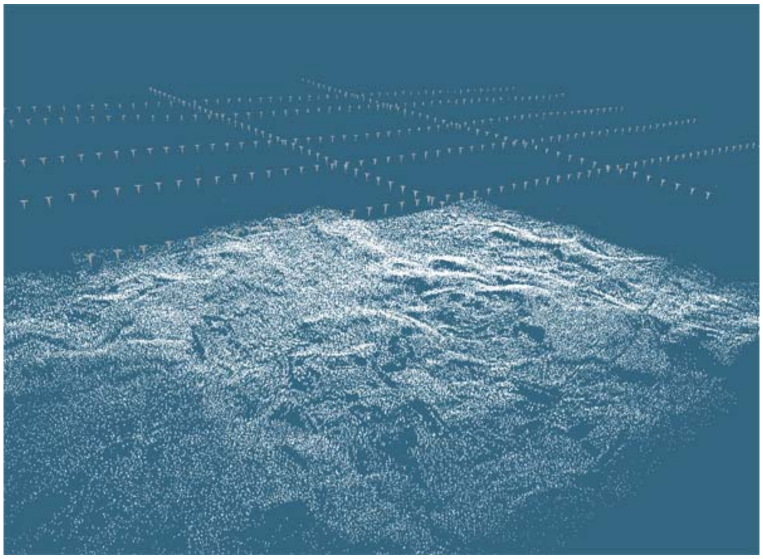


Figure 3:
Image block, 350 images, 350,000 tie points. Computed automatically by K. Karner, VRVis-Graz

However, these benefits are based on the following two restrictions: first it is assumed that the aerial triangulation cannot be performed automatically and therefore time and effort need to get invested before an exterior orientation can be obtained (film scanning, tie point measurements, manual control of tie-points, ground control point identification and surveying etc.). Second it is assumed that a fairly significant number of tie points is needed to overcome the poor rigidity of a block of images flown at 60% forward and 20% side-lap, and the resulting propagation of small errors into the resulting ground points.

It now remains to be investigated how these benefits measure up when the triangulation no longer requires much manual work, and when the rigidity of the image block is strengthened by the liberal increase of the overlaps.

5. DIGITAL ELEVATION MODELING VERSUS TERRAIN SCANNING

5.1 Accuracies

“Accuracy” of a Digital Elevation Model DEM or a terrain point cloud is being assessed first by the error of an individual surface elevation, second by the occurrence of gross errors (spikes) and third by missing areas due to occlusions.

Figure 4 illustrates by comparison between film and digital sources that successful stereo matching is more frequent in digital material. By reviewing the parallaxes in flat areas, one can also confirm that the accuracy of the surface points is better with digital sources. Residual parallaxes in film data are in the range of ±0.3 pixels, whereas in digital images, this reduces to ±0.08 pixels.

Occlusions will decrease as the overlaps increase, since the camera will see better “in-between” objects such as buildings. Maintaining a multi-image (stereo) capability in areas between large vertical objects requires a high image repeat rate. Digital cameras outperform film in this respect.

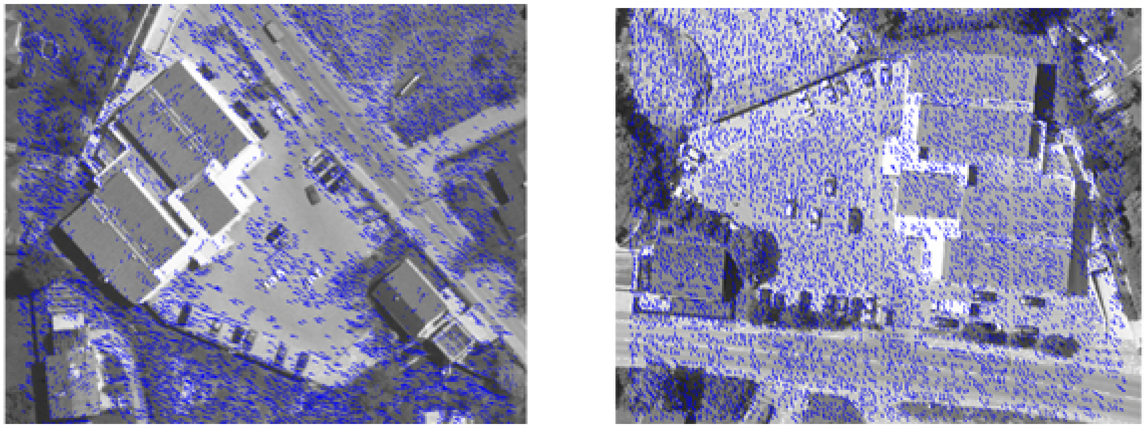


Figure 4:
Digital elevation points from film and digital images (Source: Perko, 2005)

5.2 Automation

The creation of surface point clouds from stereo images has become a routine photogrammetric capability, employing scanned film. The issue has not been the capability to automate, but the capability to automate so well that no spikes need to be edited, no water bodies needed to be excluded by hand, no hole needed to be filled, no detail needed to get filled in to properly represent subtle geo-morphological features etc. Because of these limitations of automatically prepared DEM point clouds, they have typically only been useful for the production of orthophotos, not however as a replacement for the accurate portrayal of terrain morphology.

The change of the usefulness of automatically created point clouds from digital images derives from the very dense point clouds that now can be generated using not 2 images per ground point, but 10 or so. It also benefits from land use classifications to avoid manual editing for water bodies or mismatches in certain types of vegetation. Ands it benefits from a stronger ability to extract for the image assembly a cloud of straight line-segments to support the DEM by so-called “break-lines”.

Figure 5 illustrates a dense point cloud automatically created from aerial images taken at 3 cm pixel size at an overlap of more than 70%, producing more than 3 images per object point. Note that this type of stereo imagery is not usually available from scanned film (scanning at 20 μm pixel size), since at a focal length of 15 cm, the plane would have to fly at an illegal 225 m above the ground. It would require a focal length of 21 cm to be legal, and an imaging interval of the film camera at less than 1.5 seconds.

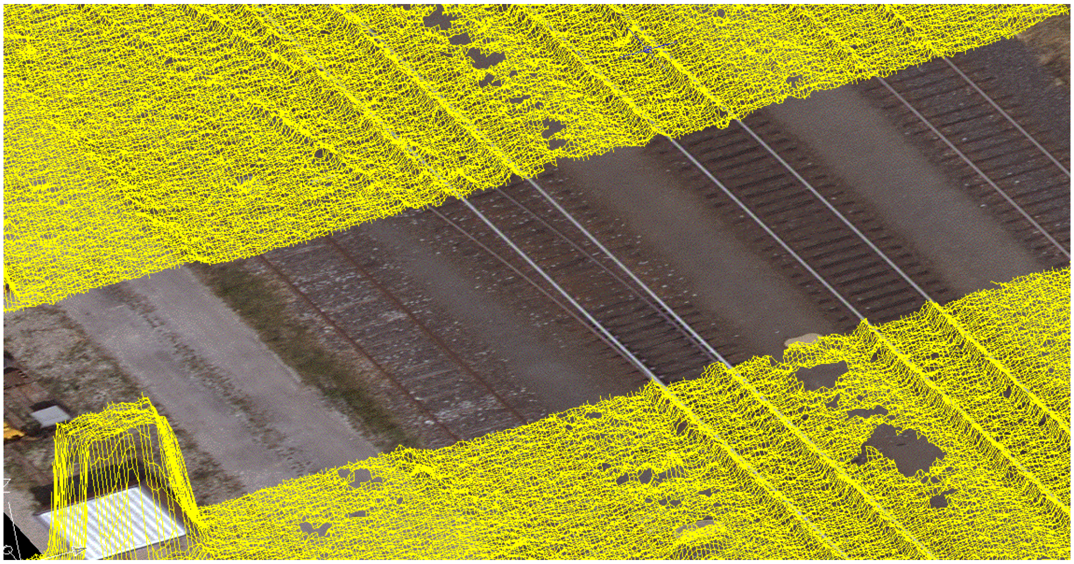


Figure 5:
Stereo-imagery at a pixel size of 3 cm, taken from a conventional survey airplane using the UltraCam-D digital aerial camera over an area at the Southern edge of Graz (Austria), and DEM extracted automatically. Note the definition of the DEM with the metal tracks elevated above the gravel bed. Point cloud from triplet of digital image set (overlap at > 70 %). Courtesy J. Jason, Vexcel Corp.

5.3 The Bald Earth and its Objects

Of great interest is not the canopy DEM, but the *bald Earth* and the mostly man-made or vegetation objects sitting on top of the bald Earth. The clouds of points and line segments need therefore to get converted into an interpreted data set. The reference terrain surface is the bald Earth. The objects on the bald Earth are in need of interpretation as buildings and their shapes, as vegetation, as complex circulation surfaces such as freeway overpasses etc. This interpretation again is in need of a land use classification to separate buildings from vegetation, to define circulation surfaces, to identify water bodies etc.

Given multiple color and texture values for each terrain point will improve the land classification.

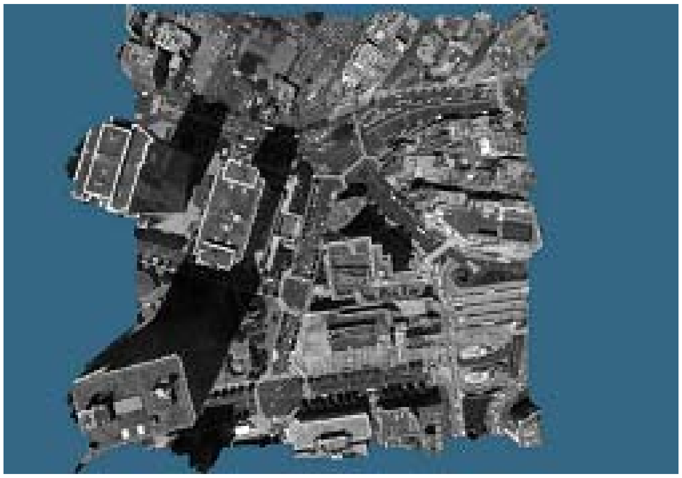


Figure 6:
Shinjuku, Tokyo. A “3-dimensional orthophoto” by draping the aerial imagery over the 3D terrain model. When the draped imagery is from a 2D orthohoto, then the facades of buildings will not have a proper texture. A data structure is needed to present both the 2D ortho-photo as well as the vertical surfaces. Imagery courtesy of PASCO, using an UltraCam-D camera. Produced by K. Karner, VRVis-Graz.

5.4 Point Clouds and Laser Scanning

The "point cloud" has become a photogrammetric data set when stereo matching was introduced. Later, point clouds have also become available from aerial laser scanners. The appeal of the laser scanner exists firstly for the reason of "instant gratification", thus a DEM obtained much sooner than by a photogrammetric process. A second benefit is the avoidance of the cost of a photogrammetric process with AT, stereo matching and manual editing. While laser scanners add cost for new equipment, savings in labor and materials can offset this. A third benefit is the reduction of occlusions since the laser scanner sweeps in a vertical plane, and therefore avoids occlusions in the flights forward direction.

If digital image photogrammetry can remove the labor and time associated with the photogrammetric process, then the first and second benefits of laser scanning disappear. High overlaps increase the look between buildings, and the third benefit of laser scanning is met.

It will have to be determined how the often quoted laser benefit of looking in between trees is really relevant and how much of this capability can also be offered by means of software actions.

6. ORTHOPHOTOS VERSUS 3-DIMENSIONAL OBJECT MODELING WITH PHOTO-TEXTURE

6.1 Product Definitions

The orthophoto is an attempt at "photo realism", but it is strictly 2-dimensional. Draping a 2-D orthophoto over a 3-dimensional terrain model creates a very detrimental and visible effect of "no pixels for vertical surfaces" (Figure 6). For 3-D object modeling, the orthophoto is insufficient.

The promotion of the "true orthophoto" is an attempt to improve the geometric accuracy of the orthophoto by correcting its geometry for the effect of vertical objects such a buildings. As soon as someone presents the true orthophoto, a detractor will comment how this product now hides valuable information about the height of vertical objects.

The solution is the 3D orthophoto, thus a data set that can be used to drape photo-texture over the detailed DEM and provides texture also in vertical areas. The "3D orthophoto" is the texture portion of a photo-textured 3D model of the terrain. It converts to a true orthophoto if the 3d model is presented in an orthogonal projection. It reveals the height of each elevated structure in any one of a variety of visualizations.

6.2 Challenges

Major challenges for the 3D orthophoto are:

- The data structure to hold the data, since there is currently no standard;
- The occlusions due to vegetation and traffic (cars, trucks, people);
- The illumination as the sun influences each surface patch;
- The un-modeled microstructure resulting in visualizations that are inconsistent with the geometric model, for example in the case of window sills, skylights, chimneys;
- The selection of each pixel's photo-texture when there is a choice to be made among10 or so values.

While there are ideas for solutions for each of these issues, it will be a hard task to make them all work in concert and thereby avoid reliance on manual interference.

7. INPUTS INTO A 3-D GEOGRAPHIC INFORMATION SYSTEM

7.1 Classification

The greatest need for labor is in the creation of vector inputs to a GIS. This has been a reserve of strictly manual work, and a human operator may spend between 2 hours to 1 month on each stereo model. Abandoning the stereo models and replacing them with multi-image computer vision leads one to hope that automated classification will produce a majority of the vectors to describe the terrain.

7.2 Integrating Vision Technologies

There exists in computer vision the concept of "*Shape-from-X*". In analogy one could postulate for the photogrammetric workflow that there be a combination of technologies all supporting the avoidance of labor hours when creating input for a 3D GIS. Color and texture can separate the scene into regions, line following and detection of straight line-segments can reinforce any vectors. Shape in 2D and in 3D should be used, for example in the form of a DEM and its gradients. Shadows need to be consistent with the solar illumination, object shapes and colors. The interpretations of various surface classes must be consistent with some rules, for example that circulation surface must lead to buildings.

8. CONCLUSION

In the evolution of the "continuum" of photogrammetric technology we are experiencing one of the most tumultuous transitions into a new world. Abandonment of film and replacing it with a completely digital workflow opens a new box of opportunities to automate and to deliver new data products that are consistent with the computer developments towards virtual and augmented reality. This in turn provides the opportunity to revitalize the field and give it new meaning and new challenges.

"Content" is an increasing need of the ever progressing systems that feed data to humans via broadband Internet. Photogrammetry can deliver such content for use by a geo-locationally "challenged" public.

Photogrammetric skillS have recently come under attack from many sides. Computer vision assumed the leadership in innovation, direct geopositioning with the GPS and inertial systems trivialized the importance of the aerial triangulation, aerial laser scanning seemed to make the photogrammetric DEM irrelevant. The orthophoto had become a trivial product at the ed of its innovation options.

The novel digital large format aerial cameras help photogrammetry to "strike back". The lack of any variable cost for a digital image changes the rule of photogrammetric imaging and mission planning. This leads to automation opportunities that bring back the need for photogrammetric innovation. Opportunities to create new types of data and thus "content" for the ever-expanding information society are a direct result of the affordability due to automation. Photogrammetry may gain in relevance by moving center stage in the world of high-bandwidth digital data creation and delivery.

REFERENCES

Gruber M., Franz Leberl, Roland Perko (2003) *Paradigmenwechsel in der Photogrammetrie durch Digitale Luftbildaufnahme?* Photogrammetrie, Fernerkundung, Geoinformation. August

Perko R. (2005) *Image Quality: Digital Pansharpening Versus Full Color Film*. Manuscript submitted for publication. Check also www.vexcel.com.

Leberl F., M. Gruber (2003) Economical Large Format Aerial Digital Camera. GIM International, June 2003, pp 12-15.

Leberl F., M. Gruber (2005) UltraCam-D: Understanding Some Noteworthy Capabilities. Photogrammetric Week 2005, University of Stuttgart, Germany.