Vertical or Oblique Aerial Photography

for Semantic Building Interpretation

PHILIPP MEIXNER[1] & FRANZ LEBERL[1]

Zusammenfassung: Schrägluftbilder erfahren einen enormen Siegeszug. Ausgehend aus den USA und angetrieben von der Firma Pictometry werden weltweit Schrägluftbilder erflogen und für die Betrachtung und für einfache Messungen in urbanen Bereichen eingesetzt. Uns interessiert der Vergleich dieser Schrägluftbilder mit vertikal aufgenommenen traditionellen Luftbildern in der Anwendung auf die Analyse von Hausfasssaden. Wir untersuchen, wie Schräg- gegenüber Vertikalaufnahmen bei der Zählung der Geschosse und Fenster abschneiden. Vertikalbilder zeigen Fassaden am Bildrand oft unter etwa 25° zeigen, in Schrägaufnahmen sind dies 40°. Wir zeigen, dass eine Zählung der Fenster und Stockwerke in Vertikalbildern mit einer Genauigkeit im Bereich von etwa 90% möglich ist, dass aber in Schrägaufnahmen diese Zählung durch Verdeckungen beeinträchtigt ist. Schrägbilder weisen keine wesentlichen Vorteile auf, wenn Vertikalbilder mit hoher Überlappung bestehen.

Abstract: Oblique aerial photography has become a widely used resource for urban imaging. Originating in the US and championed by Pictometry, oblique images are now being acquired world-wide. We are interested in a comparison between oblique and vertical aerial photography, especially addressing the façades in urban areas and façade details such as the number of floors and windows. Can one automate these tasks, and how do vertical aerial images compare to oblique images? One can image facades in vertical aerial imagery at the image's edge under an angle of 25 °. With new wide angle systems, this angle increases to 35 °. Oblique cameras produce larger angles at 35 to 55 °. With vertical images, high image overlaps are needed to obtain all façades at these angles. Our results show that vertical imagery is well-suited to façade analysis, and that oblique images deliver results compromised by occlusions. This indicates that the benefit of oblique images is questionable in cases were high overlap vertical images exist.

1 Introduction

In previous work we have shown that the automated count of floors and windows is feasible using vertical aerial imagery (Meixner & Leberl, 2010). The initial accuracies reach levels beyond 90%. Current efforts are directed towards an increase of that accuracy by the use of multi-images, and by the use of the 3rd dimension. Initial results from developing 3D point clouds of façades from vertical imagery and then using these in the façade analysis do offer encouragement. In the process one of course finds a relationship between the look angles used to image a façade and the accuracy of the analysis. Below 15°, the results become poor. At 20° and

[1] Institute for Computer Graphics and Vision, Graz University of Technology
{meixner, leberl}@icg.tugraz.at

beyond, the limitations of any façade analysis are more due to the occlusions by other buildings and by vegetation, or by the algorithm's ability to deal with balconies and other irregularities, than by a lack of image geometry and quality.

A new question emerged about the advantages of oblique aerial photography over vertical photography. The topic is relevant since the Internet inspires an interest in showing urban areas and modeling them in 3D from vertical and oblique aerial photography, aerial LiDAR and from street side imagery and street side LiDAR. Vertical aerial photography continues to be the workhorse for complete maps and orthophotos, whereas many dense 3D point clouds today are being produced by LiDARs (Leberl et al., in print). With the transition to digital sensing, image overlaps can increase without adding costs. This improves the accuracy of 3D data, the automation opportunities and the extent of occlusions. One can argue that the no-cost-per-image-paradigm has changed previous value systems: LiDAR may not add the value it once had over point clouds from film imagery, and highly overlapping digital vertical images may show facades in sufficient detail to eliminate the need for oblique photography.

Using a test area with about ~ 200 buildings and ~ 800 façades, we show that facades can be successfully analyzed from vertical aerial photography; oblique photography does not add value beyond that available from vertical photography in the analysis of façades. A qualitative visual inspection also raises doubts that oblique outperforms high quality, high resolution and high dynamic range vertical imagery.

2 Vertical and Oblique Aerial Cameras

2.1 Advances in Vertical Cameras

Vertical aerial cameras obviously produce centrally perspective imagery with the optical axis pointing towards the nadir. Until about 2003, such cameras operated with film and produced the minimum number of images needed for a project because there was significant variable cost associated with producing the film images, but also with manually processing these one by one. By 2003, digital aerial cameras began to get accepted and by this, the variable cost of creating an image was eliminated, the color capabilities got increased by a separate infrared channel and the radiometry improved from the film's 7 to 8 bit to beyond 12 bit. This is a 16-time increase from the film's 128 to 256 grey values to the digital system's 7000 values (Scholz & Gruber, 2009). Area array cameras initially produced images with 11.5 K by 7.5 K pixels, thereby assembled large format from multiple smaller image tiles. Today the same technology is at 17.5 K x 11.3 K pixels. A special development is the UltraCam-G with a swath width of nearly 30K pixels. However, aerial cameras are using a variety of technologies, including push broom sensing with a single image line that sweeps across the terrain, such as the Leica ADS-80.

The data quantities produced today by digital cameras exceed by two orders of magnitude what had previously been created via film. To eliminate the variable cost of processing an image has required that fully automated workflows become available. Such workflows now exist (Reitinger & Gruber, 2009).

At the border of a vertical aerial photograph, façades are visible under an angle of up to 36.5°, for example using a wide angle system such as the UltraCamXP-WA, or at 27.5 ° when a normal angle system is in use, such as the UltraCamXP. To ensure that all facades are imaged under as large an angle as possible, images must be taken at rapid intervals along a flight line, and the flight swaths must overlap as much as economically reasonable. Therefore digital imaging missions have abandoned the traditional paradigm of minimum photos per project and now oftentimes produce overlaps at 80% in-flight and 60% across the flight line. In urban cores with high rise buildings one uses 90%/80% overlaps. The benefits are an increase in accuracy because where one had previously a single stereo model from two images, one now can work with 45 such stereo pairs using all variations of pairs from a 10-image overlap. The robustness of automation gets improved by this redundancy, reducing the need for manual labor. And occlusions are being avoided because one now can look at the bottom of any street canyon using the appropriate image. Inversely, high overlaps lead to having each façade showing up in 10 or more images and there always will be one where the façade is shown under a large look angle in excess of 20º.

2.2 Oblique Cameras

Oblique photography is being acquired with tilted centrally perspective cameras and an optical axis looking away from the nadir at an angle of perhaps 35º. The off-nadir angle under which a façade might get imaged lies between 35 ° and 65 ° (Prandi, 2008). Such aerial images are more descriptive to a naïve viewer than vertical photographs. However, the photogrammetric workflow with aerial triangulation, dense point cloud generation and interactive stereo measurements is far less developed and more difficult than with vertical photography. The purpose of producing oblique photography is the viewing of raw images to please the eye ("eye candy") without any computer vision being applied to the images to produce derived data products. Therefore the oblique cameras typically have a much smaller pixel array than vertical mapping cameras. Petrie (2009) assembled a catalog of current oblique cameras and classified them into three different technologies. We add a 4th class here:

- **Fans** of digital cameras, all optical axes in one vertical plane, but looking in different directions to assemble a panoramic coverage. The application may predominantly be in defense-related surveillance. An example is the Zeiss KS-153 reconnaissance camera in Figure 1. DiMAC, IGI and RolleiMetric use a twin camera configuration. The German Space Agency (DLR) built a small format digital camera with 3 oblique cameras. The Russian NPO KSI organization has produced a camera that has, depending of the flying height, 4, 6 or 8 lenses to cover the largest possible area.
- **Block configurations** of multiple camera heads seek to create a larger field-of-view by arranging the tilted optical axes not in a plane, but spatially extended and covering a larger ground area from one single exposure station. These include also the vertically imaging cameras for mapping such as the Intergraph DMC. Other solutions are by Rolleimetric in its AICx4, IGI and American Space System Division with an assembly of 6 single oblique camera cones to survey larger areas.
- **"Maltese Cross" configurations** combine one vertical and multiple oblique cameras. This is the typical oblique systems used by Pictometry or Track-Air's Midas system in Figure 2. The technology was used in the 1930 by the US Geological Survey and the US Army Corps of Engineers for mapping applications. Apart from Pictometry and TrackAir, there exist solutions by IGI with its Penta-DigiCAM system, by DiMAC with

the DIMAC oblique camera, by GetMapping with its Azicam, and by RolleiMetric with the Aero Oblique System (AOS).

- **Scanning area arrays** represent a new class of oblique technology. A single optical imaging cone with an area array sweeps by a scanning motion across the flight line and produces multiple images with varying optical axes. An example is Vision Map's A3-system (see www.visionmap.com).

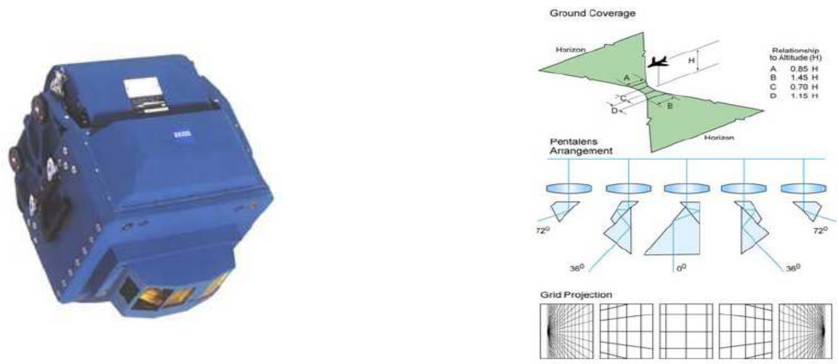

Figure 1: The ZEISS KS-153 Pentalens 57 oblique camera system using a fan arrangement Source: Petrie (2008)

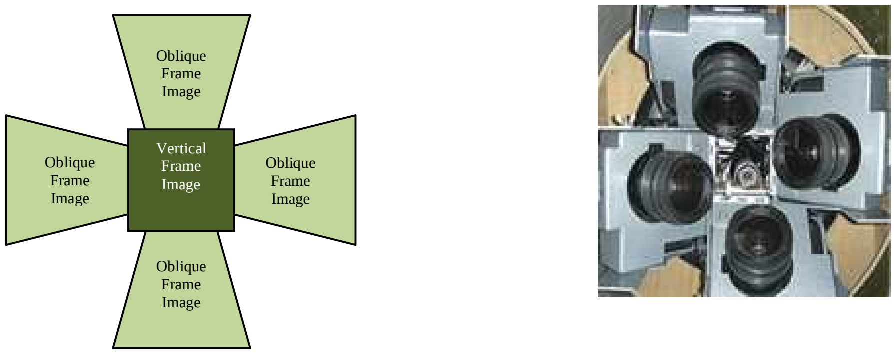

Figure 2: The Maltese Cross configuration with five camera heads (left) and an realization in the form of the TrackAir Midas System (right). Source: Petrie (2008)

3 Geometric Parameters of Oblique versus Vertical Photography

3.1 Oblique Camera Geometry

Oblique aerial imagery has been easy to come by in the form of Microsoft's BING/Maps mapping website. The images have been produced with a Maltese Cross configuration. While exact technical data of such oblique images remain unavailable, we have taken it upon ourselves to reverse engineer them.

For this purpose we can access high quality vertical aerial photography from commercial sources taken, in our case, with the UltraCam-series of aerial cameras. Figure 3 presents an example of an area in Graz (Austria) with a vertical coverage and superimposed the outlines of an oblique image. Also shown is the oblique image itself in its original geometry.

Figure 3: Detail from Microsoft Bing Maps. Left is the orthophoto and superimposed the outline of an oblique aerial image produced with the Pictometry system operated by Blom. Right is the oblique aerial image.

The oblique images have 2,672 rows and 4,000 columns representing a format defined from traditional consumer cameras at 36 mm x 24 mm.
Well-mapped terrain with large vertical structures including a church can serve as a test area to compute a resection in space with self-calibration, in which we compute the internal geometry with its focal length. The result is summarized in Table 1.

Pixel size in the image plane =	9µm	Near range off-nadir angle =	37 º
Focal length =	85.5mm	Far range off-nadir angle =	53 º
Viewing angle of the camera =	24° x 16°	Horizontal GSD at near range =	14 cm
Flying height above ground =	1,130 m	Horizontal GSD at far range =	19 cm
Distance to near range =	850 m	Distance to far range =	1530 m

Table 1: BING/Maps oblique imagery parameters reconstructed from known terrain points.

3.2 Vertical Image Geometry

Table 2 summarizes some relevant geometric parameters of a digital aerial camera in the form of the UltraCam-X and wide angle UltraCam XP.

	UltraCam X	**UltraCam XP - WA**
Image Rows x Columns	14, 430 x 9,420	17, 310 x 11,310
Image size in X and Y, in mm	103.9 x 67.8	
Pixel size in image plane (µm)	7.2	6
Focal length, mm	100	70
Max Look angle off-nadir (°)	27.5	36.5

Table 2: Some geometric data of two typical digital aerial cameras (from www.vexcel.com)

3.3 Pixel Sizes on Façades

For a vertically-looking camera, the pixel on a façade (FSD or Façade Sampling Distance) changes as a function of the look angle off-nadir α with

$$FSD = GSD / \tan(\alpha)$$

These results in Table 3 for a GSD at 10 cm, a typical value for urban aerial photography. The façade pixels are rectangular.

Angle (deg)	0.	5	10	15	20	25	30
Pixel vertical [cm]	∞	114	57	37	27	21	17

Table 3: Incidence or look angles and vertical pixel size within a façade. The horizontal pixel size is at 10 cm.

For an oblique camera, the pixel size within a vertical plane is defined by two angles. Angle β is the orientation of the optical axis off-nadir. Angle α is the angle between the optical axis and the actual imaging ray. The off-nadir angle β produces 2 GSD values, one in the direction of the inclination of the optical axis GSDr, with r being the range direction or direction between nadir and the optical axis; and the GSDa in azimuth direction

$$GSDr = p * H * \cos(\alpha) / (f * \cos^2(\beta)); \qquad GSDa = p * H * \cos(\alpha) / (f * \cos(\beta))$$

A vertical façade is resolved as a function of where in the image it is located and this is defined by the second angle α, producing a vertical pixel dimension FSDv:

$$FSDv = p * H * \cos(\alpha) / (f / \sin(\beta))$$

Table 4 presents some vertical façade pixel sizes for the oblique camera with a look angle at 53° and compares this to pixel sizes form an UltraCam X with an angle with 22°. In this example, the vertical façade pixel sizes for the oblique camera at an angle of 53° and for an UltraCam X at an angle of 22° are almost identical. The simple conclusion is permitted that the pixel size not only is a function of the look angle, but also of the flying height and the GSD. While the look angle appears a lot less attractive from a consideration of look angles in the Ultracam, on a given façade, this does not propagate into an inferior geometric resolution.

	Degrees	Azimuth (cm)	Range (cm)	Façade vertical (cm)
Oblique camera (near range)	37	14.7	20.9	27.7
Oblique camera (far range)	53	19.6	27.7	20.9
UC-X	22	8.1	8.1	20.0

Table 4: Size of a pixel on a façade in cm, as a function of the look angle, in °

3.4 Efficiency of Aerial Data Collection

A consideration of image pixel sizes ignores the efficiency of one versus another imaging approach and technology. Flying at a certain flying height to achieve small pixels, and producing images with a large format will be more efficient than to fly with small formats for a small swath width and at a low flying height. An UltraCam for example produces 17.5 K pixels in one single flight line. An oblique camera will have to match this number to be comparatively productive. At a frame size of 4,000 pixels, one will not easily match the productivity of a vertical mapping camera.

4 Experimental Results

We work with a 400m x 400m test data set in the city of Graz (Austria). The vertical images were acquired with an UltraCamX (Vexcel/Microsoft) at a GSD of 10cm and 80/60 image overlaps. The oblique images were taken from the Microsoft BING/Maps website in its “Classic” version, and have a GSD of nominally 12cm.

4.1 Visual Comparison

A visual comparison of vertical versus oblique images in Figure 4 does not result in a clear advantage of one versus the other approach. At an off-nadir angle of about 45°, the oblique images have more significant occlusions, given that the vertical images show the same facades at an angle of only 27°. Regarding the radiometric range, one would give the vertical images an advantage. This visual impression from original images is overwhelmed by the differences in geometry. To eliminate that factor, we create rectified versions in the plane of the façade, as shown in Figure 5. Again, a visual comparison does not show a clear advantage of one over the other technology.

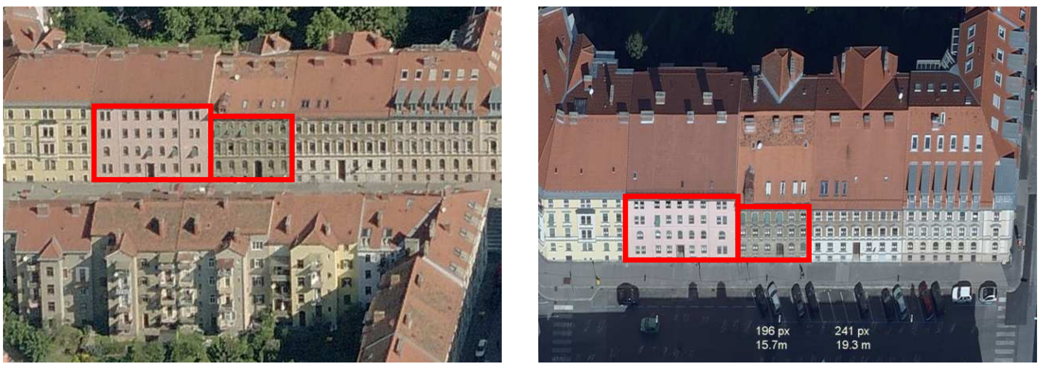

Figure 4: Oblique aerial image at 45° look angle taken from Microsoft Bing Maps (left); Vertical aerial image obtained from UltraCamX at a look angle of 27° (right).

Figure 5: The marked sections of Figure 4 have been rectified. At left are two sections from the oblique data at 45°, at right from the vertical data at 27° .

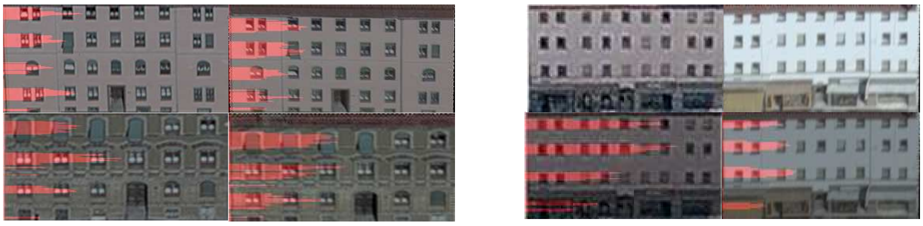

Figure 6: Two examples for floor detection using edge histograms. Left is for Fig 5. Note in the right example the advantage of image quality in the vertical data set (right).

4.2 Counting Floors

A less subjective and more quantitative comparison of oblique versus vertical is expected to result from analysing the images and extracting semantic information. A floor detection algorithm has been explained by Meixner & Leberl (2010). Figure 6 explains that a histogram is being built from horizontal Prewitt edges and local extrema of the histogram serve to get a floor count. Applying this approach to about 870 facades in the Graz test area's vertical images, and to a subset of 120 facades in the corresponding oblique images (from Bing/Maps) leads to Table 5. We find in this type of quantitative analysis that the result is seriously compromised by the occlusions which naturally are larger in the oblique images.

4.3 Counting Windows

A histogram-based count can also deliver the number and locations of windows. Figure 7 explains the principle of the approach. Of course one will want to apply various constraints on window size and distance betweens etc. to overcome the effects of data noise. Table 5 shows the accuracy achieved in the Graz test data set form the facades on vertical and oblique images. Again, occlusions are the main obstacle to a competitive result from oblique images.

Angle [deg]	**< 5**	**5 - 10**	**10 - 15**	**15-20**	**20-25**	**> 25**	**Oblique**
Floor Detection	0	7 / 21	79 / 103	191 / 221	255 / 279	228 / 246	90/120
Floors Percentage	0%	33%	77%	86%	91%	93%	75%
Window detection	0	6 / 21	69 / 103	174 / 221	233 / 279	212 / 246	79/120
Windows Percentage	0%	29%	67%	79%	83%	86%	66%

Table 5: Counting floors (above) and windows (below) from vertical images and results depending on look angles. Last column is from oblique images where floor counts are compromised by occlusions.

In Table 5 the success rate of window and floor detection is calculated by dividing the total number of facades for every angle (e.g. 5-10) by the number of facades where the floors and windows are correctly determined (e.g. 7/21). As one can see the floor and window detection results for oblique images are not as good as the results using vertical aerial images. Reasons for that are the poor resolution of the oblique aerial images and occlusions from other buildings and vegetation. Concerning the floor detection occlusions are the main reason for these results. Concerning the window detection the poor resolution of the images is one of the reasons for the outcome.

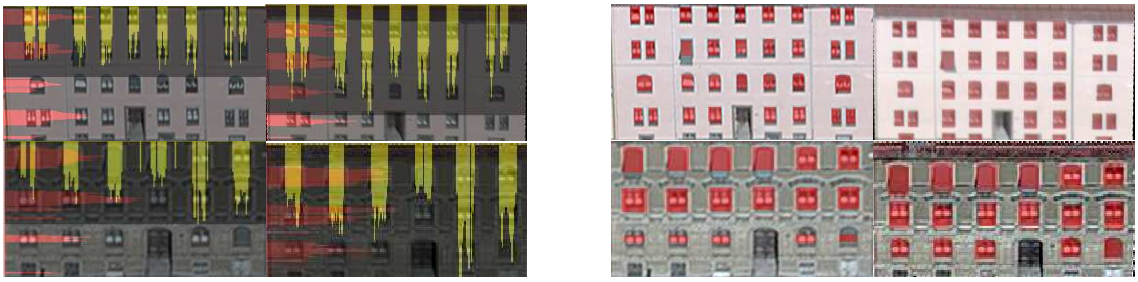

Figure 7: Window detection approach using edge histograms. To the right are the marked window locations and sizes, as detected.

5 Conclusions

We demonstrate that façades can be analyzed with a 90% success rate from vertical aerial photography. This is feasible since the images have been taken with large overlaps so as to image each façade at a sufficiently large look angle of 20° to 27°. We also show that the visual inspection of vertical versus oblique images favors the vertical data due to better radiometry at comparable pixel sizes. The major problems of oblique images are occlusions that prevent one from counting the correct number of floor and windows.

The efficiency of aerial imaging may favor vertical technologies over the oblique approach. Vertical images today produce 200 Megapixels per exposure, whereas oblique cameras still operate at the 10 Megapixel level. Even if one were to consider that in a Maltese Cross arrangement one operates with 5 such cameras, this still adds up to only 50 Megapixels.

A limitation of the current “normal angle” aerial cameras is the look angles one can achieve in the direction of flight at perhaps 17° off nadir. Solutions are either a cross flight pattern, or the use of a wide angle camera model such as the UltraCam Xp-WA with 26° in flight direction, or the use of the new single CCD-chip DMC-II, recently announced by Intergraph.

Going beyond a mere “eye candy” approach for the use of oblique images, one will quickly find that novel high-redundancy vertical aerial images offer a superior source of information about urban areas, street canyons and facades. We suggest that the benefits from vertical aerial photography have been undervalued, and that conversely benefits from oblique images have been overstated.

6 References

LEBERL F., H. BISCHOF, T. POCK, A. IRSCHARA, S. KLUCKNER (2010) *Aerial Computer Vision for a 3D Virtual Habitat.* IEEE Computer, June 2010, pp.1-8

LEBERL F., A. IRSCHARA, T. POCK, P. MEIXNER, M. GRUBER, S. SCHOLZ, A. WIECHERT (in print) *Point Clouds: LiDAR versus 3D Vision*. Photogrammetric Engineering and Remote Sensing.

MEIXNER P. LEBERL F. (2010) *From Aerial Images to a Description of Real Properties: A Framework*. Proceedings of VISAPP International Conference on Computer Vision and Theory and Applications, Angers 2010.

PETRIE G. (2009) *Systematic Oblique Aerial Photography using Multiple Digital Frame Cameras*. Photogrammetric Engineering & Remote Sensing, Vol. 75, No. 2, p. 102-107, 2009.

PRANDI F. (2008) *Lidar and Pictometry Images Integrated Use for 3D Model Generation*. Proceedings of the XXI ISPRS Congress, Beijing 2008.

SCHOLZ S., M. GRUBER (2009) *Radiometric and Geometric Quality Aspects of the Large Format Aerial Camera UltraCam Xp.* Proceedings of the ISPRS, Hannover Workshop 2009 on High-Resolution Earth Imaging for Geospatial Information, XXXVIII-1-4-7/W5, ISSN 1682-1777