A Permanent Test Field for Digital Photogrammetric Systems

Eija Honkavaara, Jouni Peltoniemi, Eero Ahokas, Risto Kuittinen, Juha Hyyppä, Juha Jaakkola, Harri Kaartinen, Lauri Markelin, Kimmo Nurminen, and Juha Suomalainen

Abstract

Comprehensive field-testing and calibration of digital photogrammetric systems are essential to characterize their performance, to improve them, and to be able to use them for optimal results. The radiometric, spectral, spatial, and geometric properties of digital systems require calibration and testing.

The Finnish Geodetic Institute has maintained a permanent test field for geometric, radiometric, and spatial resolution calibration and testing of high-resolution airborne and satellite imaging systems in Sjökulla since 1994. The special features of this test field are permanent resolution and reflectance targets made of gravel. The Sjökulla test field with some supplementary targets is a prototype for a future photogrammetric field calibration site.

This article describes the Sjökulla test field and its construction and spectral properties. It goes on to discuss targets and methods for system testing and calibration, and highlights the calibration and testing of digital photogrammetric systems.

Introduction

Photogrammetric imaging techniques are being converted from analog to digital. The improved image quality will mean new potential for automation for image measurement and interpretation. Digital imaging, coupled with modern direct georeferencing techniques, makes up-to-date georeferenced imagery available for users within hours of image collection. In order to exploit these advancements to optimum effect, the performance of the entire photogrammetric system must be studied, and the overall system calibration must be accurately determined.

Calibration of analog cameras relies on laboratory calibration. Laboratory calibration maintains its position as the core of the calibration process for digital sensors (Pacey *et al.*, 1999; Schuster and Braunecker, 2000; Diener *et al.*, 2000; Heier *et al.*, 2002; Kröpfl *et al.*, 2004; Cramer, 2004), but the role of field calibration methods will also become substantial (Cramer, 2005). The wide use of new photogrammetric sensors in mapping and remote sensing applications will necessitate calibration of geometric, spatial resolution, radiometric, and spectral properties. Spatial response is a fundamental quality indicator for imaging systems (e.g., Leachtenauer *et al.*, 1997; Ryan *et al.*, 2003). Radiometric and spectral calibration and testing are crucial for utilization of the advanced radiometric and spectral properties of digital photogrammetric instruments, and it will also simplify the tonal processing of the images for visual applications.

Geometric stability and calibration are elementary properties of metric photogrammetric cameras (e.g., Cramer, 2004). For instance, experience with both analog and digital systems has shown that geometric laboratory calibration is not necessarily valid in airborne conditions; the use of accurate system parameters is advantageous especially in direct georeferencing applications and in applications with accurate image center position information (Heipke *et al.*, 2002; Honkavaara *et al.*, 2003, 2006a, and 2006c). The advantages of field calibration are that the entire imaging system and production line is calibrated and that the parameters determined are valid in airborne conditions.

Various approaches can be used in field calibration and testing. They can be performed in a permanent or temporary test field or in the mapping area itself. The reference targets may be permanent or transportable artificial targets or objects from the imaging area. It is to be expected that in the end, a sort of calibration will be performed on each flight. Although at the moment, there are no generally accepted methods for field calibration and testing; several international working groups are working to develop uniform practices and standards, e.g., the CEOS/ISPRS calibration and validation task force (Morain and Zanoni, 2004), the USGS digital camera characterization initiative (USGS, 2004), and the EuroSDR network on digital camera calibration (Cramer, 2004). Empirical results are needed to optimize the methods. This article emphasizes methods based on permanent test fields and permanent or transportable artificial targets.

Some permanent test fields are available for photogrammetric instruments around the world. Typically, these test sites contain targets for geometric calibration and testing only. An exception in terms of versatility is the NASA test field at the Stennis Space Center (USGS, 2004), where calibration and testing of all central image properties is possible. Well-known geometric test fields are the Vaihningen test field in Germany (Cramer, 2005), the Fredrikstad test field in Norway (Nilsen, 2002), the Pavia test field in Italy (Casella and Franzini, 2005), and the USGS/OSU Madison test and calibration range (Merchant *et al.*, 2004). Photogrammetric camera manufacturers also maintain geometric test fields: e.g., the Elchingen test field (Dörstel, 2003) and the Herbrugg test field (Tempelman *et al.*, 2003). Due to its versatility and openness, the Sjökulla test field of the Finnish Geodetic Institute (FGI) is special compared with any other permanent test field.

Finnish Geodetic Institute, Geodeetinrinne 2, FIN-02430 Masala, Finland (eija.honkavaara@fgi.fi).

The FGI established a permanent test field at Sjökulla in 1994 for high-resolution airborne and satellite imaging systems (Kuittinen *et al.*, 1994; 1996; Ahokas *et al.*, 2000). The test field contained targets for geometric, spatial resolution, and radiometric testing. The key ideas of the permanent test field were that it should be available year-round, its maintenance operations should be minimal, it should exist over a long period of time, it should be continuously monitored, and it should be freely and easily accessible by image producers. Because environmental conditions in Finland are severe, a unique approach was taken to construct permanent radiometric and resolution targets of gravel. Over the years, the test field has been improved and testing methods have been developed. Transportable targets have been developed to supplement the permanent ones. The camera testing carried out at the Sjökulla test field is part of the quality control systems used by Finnish mapping companies and the National Land Survey; the whole photogrammetric production line is evaluated regularly with the test field. Recently, test flights at Sjökulla have been increased substantially. In 2004 and 2005, several digital photogrammetric systems were tested at Sjökulla to study the performance and calibration aspects of digital sensors.

Calibration and testing are discussed in this article. In principle, the same steps are taken in both processes. The major distinction is that the objective of testing is to study the performance of the system, while calibration has two objectives: system testing and determination of new system parameters.

The objective of this article is to describe a prototype test field for digital photogrammetric systems. The Sjökulla test field with some supplementary portable targets is used as the basis of the prototype. The article depicts the construction and spectral properties of the gravel features at Sjökulla and discusses the targets and calibration and testing methods. The article concentrates on reference targets; methods are reviewed only on a general level.

The Permanent Sjökulla Test Field

The Sjökulla test field is located in a rural area surrounded by lakes, fields, and forests at Sjökulla, Kirkkonummi, near the Metsähovi Research station, and a moderate distance from the head office of the FGI and airports. Parts of the test field are an image quality test field for radiometric and spatial resolution calibration and testing (Plate 1), and networks of targeted benchmarks for geometric calibration and testing (Figure 1).

The Permanent Image Quality Test Field
Construction
An aerial photograph of the image quality test field is shown in Plate 1. The size of the image quality test field is approximately 60 m × 100 m. The original structure, described by Kuittinen *et al.* (1994), has been preserved. In order to ensure stability and resistance against frost, water, and vegetation, the test field was constructed in several layers. From bottom to top, these are sand (100 to 200 mm layer), fiber material, gravel (diameter 25 to 55 mm; 100 mm layer), fiber material, and gravel (diameter 8 to 16 mm or 4 to 8 mm; 50 mm layer). The base material is clay. The uppermost layer has various test patterns made of four gravel types: dark gabbro (black) from Hyvinkää, gray granite from Kuru, red granite from Ridasjärvi, and white limestone from Sipoo (Plate 2).

Alternatives to gravel targets are painted targets, for example, on tarpaulin, plywood, or concrete. The problem with painted targets is that the radiometric and spectral properties

change rapidly due to fading and dirt (Salamonowicz, 1982; Hakkarainen, 1991; Moran *et al.*, 2001). Further problems are experienced with many materials because of morning dew, puddles after rain, and dirt, which necessitate visits and wiping and cleaning of the targets before test flights. The well-constructed gravel targets have several advantages over painted targets. Gravel is weather resistant; on the gravel field the rainwater washes the dust and dirt away and the surface dries quickly; rain does not damage it; snow and frost do not distort the figures; and several gravel types have attractive reflectance properties (see below).

Spectral Properties of Gravel
The brightness, color, and spectral shape of all materials depend on the illumination conditions and the direction of observation. The bidirectional properties depend on the structural and optical properties of the material.

The detailed bidirectional reflectance functions (BRF) of the gravel samples have been measured using the Joint Research Centre (JRC) and FGI spectrogoniometers (Peltoniemi *et al.*, 2005; Peltoniemi, unpublished data, 2006). BRFs of gray and black gravel are shown in Figure 2; the behavior of red and white gravel was similar. The results showed that the anisotropy of the scattering was significant; all samples were strong backscatterers.

Reflection spectra taken at nadir with a solar zenith angle of 50° are given in Figure 3. The measurements were performed on four occasions; at JRC undisturbed samples were measured in 1997 in the laboratory using artificial light, and the FGI measurements were performed in the field in 2001, 2004, and 2005 (the gravel had been in open-air conditions for 8 to 12 years). Differences in measurement conditions can slightly affect the scale of the spectra. The spectral behavior was fairly consistent over the 500 to 2,000 nm range; black was always darker than gray, and gray darker than white. Red granite had a slightly more varied spectrum. Below 500 nm there was some darkening with all samples, and above 2,000 nm there were also some other features. The white limestone had changed dramatically in visual wavelengths, but all other samples had withstood time rather well. Some lichen, moss or algae has started growing on the surface of the limestone and that darkened the substance up to 60 percent, but in a very inhomogeneous way. The approximate reflectances in the wavelength range of 500 to 1,000 nm, derived from Figure 3, are given in Table 1. The reflectance range of the undisturbed samples was 5 to 50 percent; the reflectance range of the 11 to 12 year old samples was 4 to 30 percent.
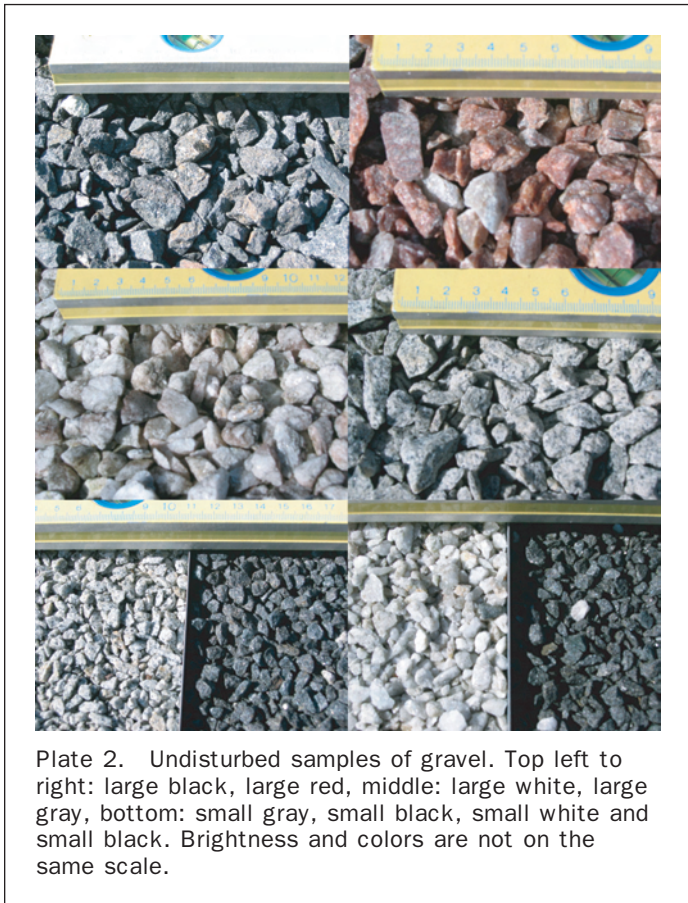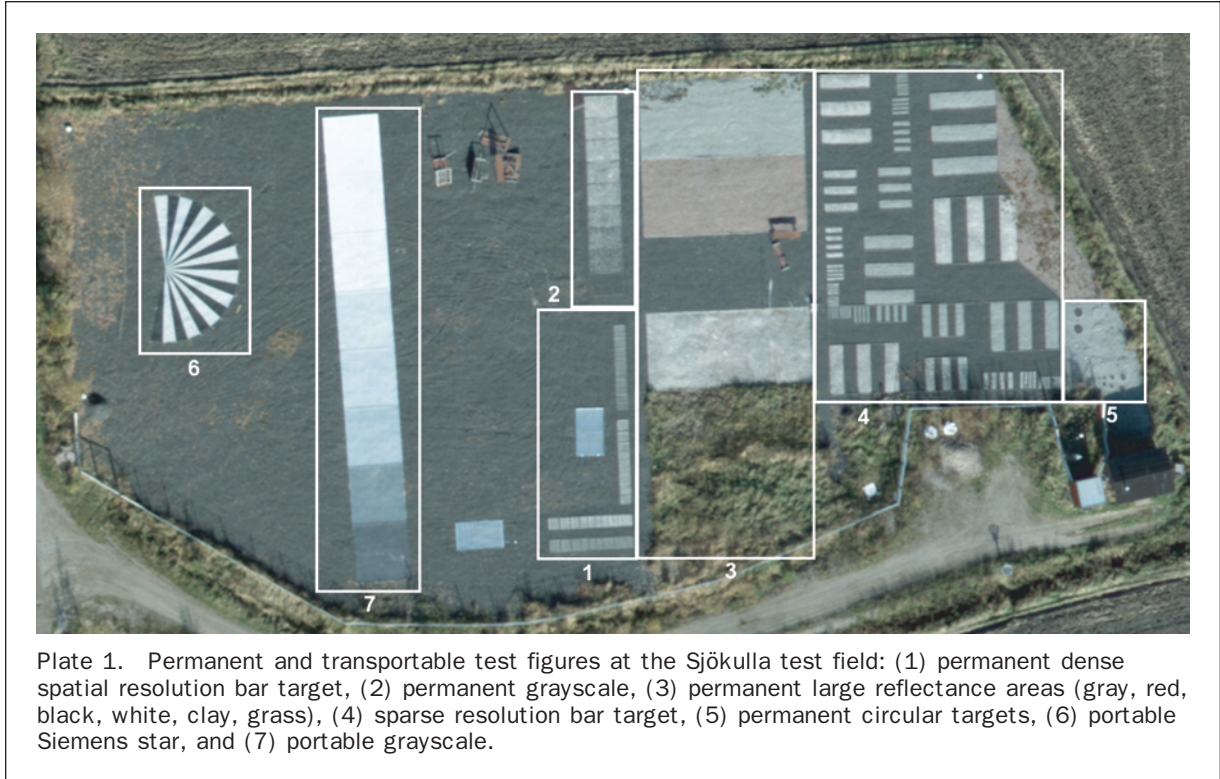
BRFs of the materials should be measured annually, and the gravel should be renewed as soon as significant radiometric/spectral changes occur. The lifetime of the white limestone is less than five years; the other gravel types have maintained their spectral and radiometric properties for 10 years or even more.

Test Targets
The image quality test field includes the following permanent test figures made of gravel (numbering refers to Plate 1):

1. Dense resolution bar target (Figure 6).
2. A gray scale with nine steps (Figure 11).
3. Large rectangular areas made from black, gray, red and white gravel, sand and clay/grass (Figure 10). The sand and grass targets have become overgrown with time and are no longer used.
4. Sparse resolution bar target.
5. Circular targets for spatial resolution studies.

The rest of the test field is covered with black gravel for the installation of temporal test targets.

Plate 1. Permanent and transportable test figures at the Sjökulla test field: (1) permanent dense spatial resolution bar target, (2) permanent grayscale, (3) permanent large reflectance areas (gray, red, black, white, clay, grass), (4) sparse resolution bar target, (5) permanent circular targets, (6) portable Siemens star, and (7) portable grayscale.

Plate 2. Undisturbed samples of gravel. Top left to right: large black, large red, middle: large white, large gray, bottom: small gray, small black, small white and small black. Brightness and colors are not on the same scale.
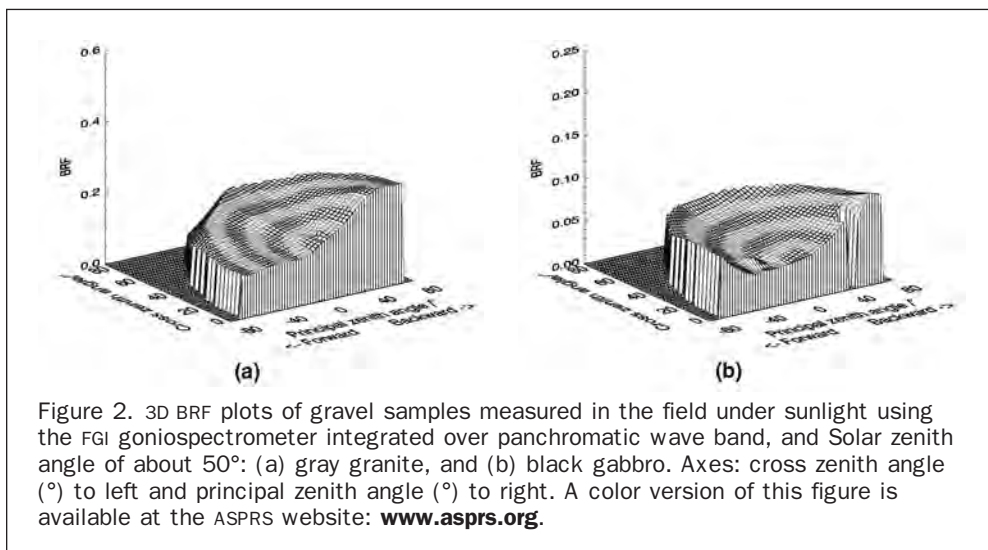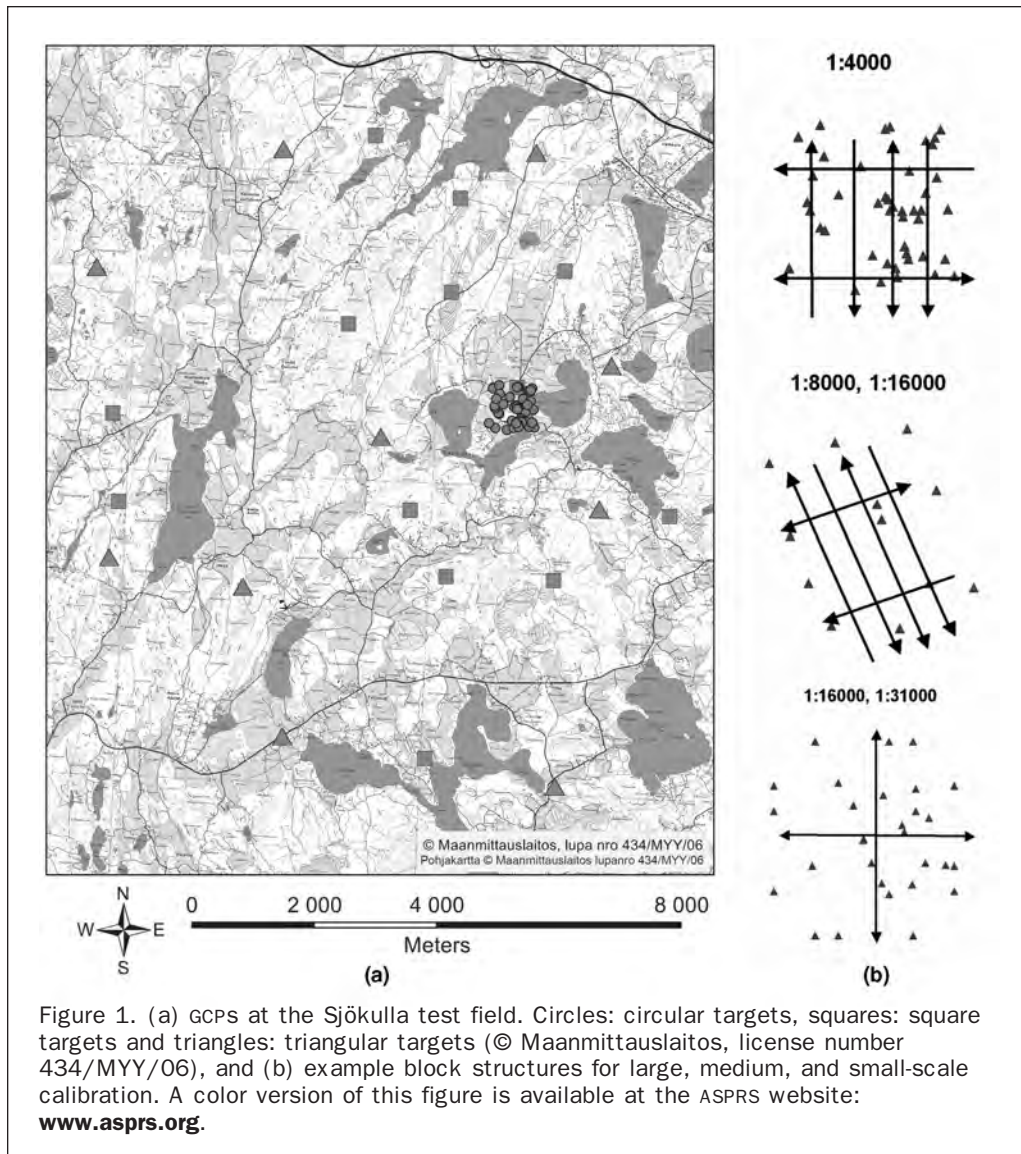
Geometric Test Field

The geometric test field contains targeted benchmarks for calibration at large, medium, and small scales (Figure 1). The benchmarks have been measured with the static GPS method. For medium-scale benchmarks, a minimum of six-hour sessions were measured, and for the smallest scale benchmarks, a minimum of one-hour sessions were measured. In the large-scale test field, 20 minutes or longer sessions have been used. In the large-scale field, 18 benchmarks are attached to bedrock or large stones with bolts, and 26 benchmarks are attached to soil and constructions; for the smaller scales, all the benchmarks are attached to the bedrock with bolts. The points fixed on soil and constructions are re-measured annually. The estimated accuracy (RMSE) of the benchmarks is 5 mm in horizontal coordinates and 10 mm in height. The final accuracy of the ground control point (GCP) is dependent on the target type as shown in Table 2.
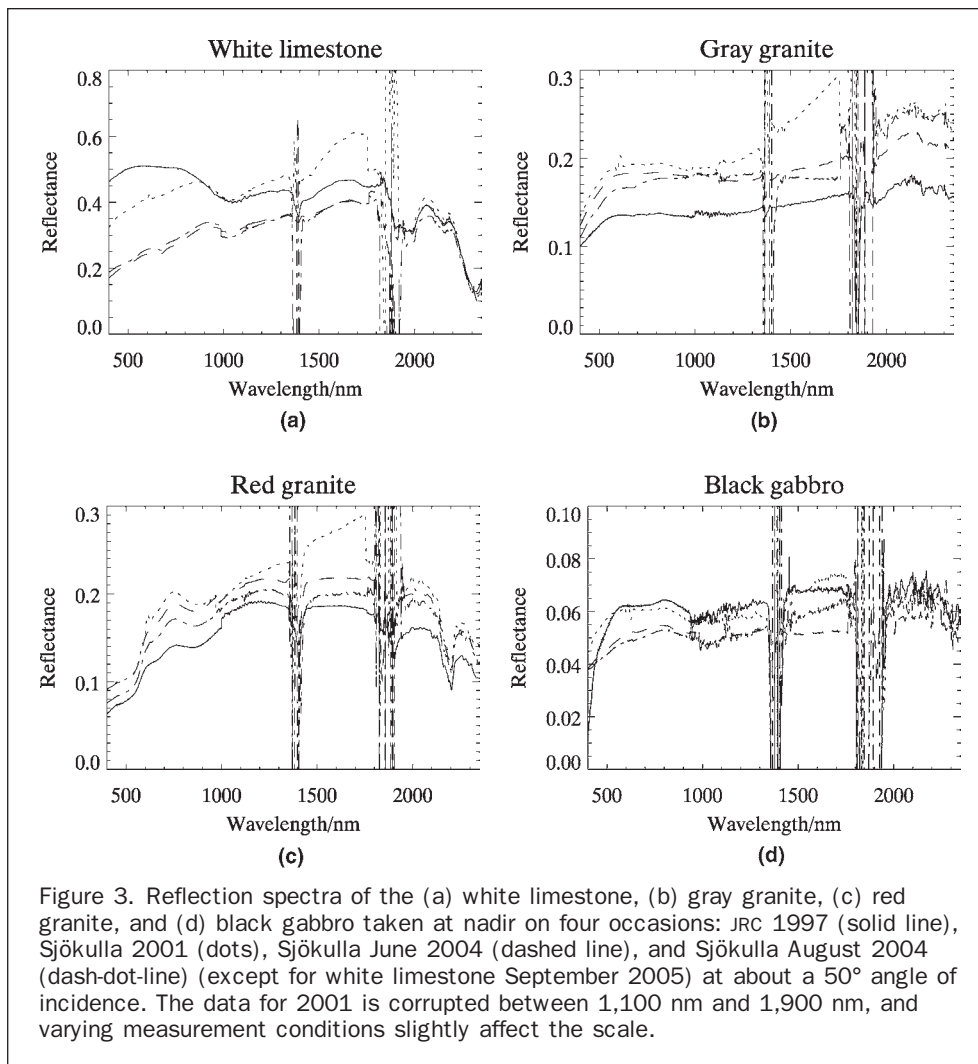
Three types of targets are used (Figure 4):

- Circular white targets of 0.3 m and 0.4 m in diameter painted on plywood. The background is black.
- Square white targets of 1 m × 1 m constructed from plywood.
- Triangular targets *slit frames* of timber with a side length of 2.4 m.

The triangular targets are permanent constructions, while the circular and square targets are installed in the field only during the imaging season. The distribution of the benchmarks and the target type determine feasible ground sample distance (GSD) and imaging scale (Table 2, Figure 1).

A possible improvement to the Sjökulla geometric test field is to increase the number of benchmarks; a minimum of 50 GCPs at each scale is desirable. A specific difficulty is to

Figure 1. (a) GCPs at the Sjökulla test field. Circles: circular targets, squares: square targets and triangles: triangular targets (© Maanmittauslaitos, license number 434/MYY/06), and (b) example block structures for large, medium, and small-scale calibration. A color version of this figure is available at the ASPRS website: **www.asprs.org**.

Figure 2. 3D BRF plots of gravel samples measured in the field under sunlight using the FGI goniospectrometer integrated over panchromatic wave band, and Solar zenith angle of about 50°: (a) gray granite, and (b) black gabbro. Axes: cross zenith angle (°) to left and principal zenith angle (°) to right. A color version of this figure is available at the ASPRS website: **www.asprs.org**.

Figure 3. Reflection spectra of the (a) white limestone, (b) gray granite, (c) red granite, and (d) black gabbro taken at nadir on four occasions: JRC 1997 (solid line), Sjökulla 2001 (dots), Sjökulla June 2004 (dashed line), and Sjökulla August 2004 (dash-dot-line) (except for white limestone September 2005) at about a 50° angle of incidence. The data for 2001 is corrupted between 1,100 nm and 1,900 nm, and varying measurement conditions slightly affect the scale.

build and measure control points for the large scales accurately enough. In the large scales (e.g., 1:4 000), the accuracy of the circular points in the image space is 2.5 μm in the horizontal coordinates and 5 μm in height. This accuracy can be considered sufficient for the GCPs for calibration purposes. In the test flights at Sjökulla, the best point determination accuracy results have been at the level of "scale factor $\times$ 1.5 μm" in horizontal coordinates and 0.02 percent of object distance in the height component; the accuracy of checkpoints should be one-third of that or better. Thus, for the largest scales (1:4 000; object distance 400 m), the checkpoint accuracy should be 2 mm in horizontal

coordinates and 3 mm in height. In practice, it is extremely difficult to attain such high accuracy. The limited accuracy of the checkpoints must be taken into account in evaluation of the results.

Field Calibration and Testing Methodology

General

The result of calibration and testing is dependent on many factors (Hakkarainen, 1991; Read and Graham, 2002). Of these factors, some are permanent system factors, some are systematic factors related to certain conditions, and the rest are random components. In the case of the digital sensors these factors can be placed in the following six groups:

1. Camera: e.g., lens, CCD, aperture, filters, shutter, exposure time, FMC, in-flight data processing (e.g., compression),
2. System: e.g., camera mount, camera port glass, navigation system, GPS, GPS/IMU,
3. Flight: e.g., flight altitude, flight velocity, airplane vibrations, airplane swing, temperature, pressure, humidity,
4. Atmosphere: e.g., air turbulence, visibility, sun height and direction, spectral distribution of light,
5. Object: e.g., contrast, shape, and
6. Data postprocessing: e.g., geometric and radiometric correction, and resampling.

TABLE 1. APPROXIMATE REFLECTANCES OF THE SJÖKULLA GRAVELS IN THE WAVELENGTH RANGE OF 500 TO 1,000 NM (DERIVED FROM FIGURE 3). JRC: LABORATORY MEASUREMENT OF UNDISTURBED SAMPLES USING JRC GONIOMETER. FGI: FIELD MEASUREMENT OF 11 TO 12 YEAR OLD SAMPLES USING FGI GONIOMETER (AUGUST 2004 AND SEPTEMBER 2005)

Method	White	Gray	Red	Black
JRC	40–50%	13%	8–16%	5–6%
FGI	20–33%	15–18%	9–18%	4–5%

TABLE 2. GEOMETRIC TEST FIELDS AT SJÖKULLA

Scale	Area (km²)	Analog Scale Number	GSD (m)	GCPs Number	Target Type	Accuracy (cm)
Large	1 × 1	3000 to 4000	<0.1	44	42 circular 2 square	$\sigma_X = \sigma_Y = 1.0$, $\sigma_Z = 2.0$ $\sigma_X = \sigma_Y = 1.0$, $\sigma_Z = 2.0$
Medium	4 × 5	8000 to 16000	<0.3	12	10 square 2 triangle	$\sigma_X = \sigma_Y = 1.0$, $\sigma_Z = 2.0$ $\sigma_X = \sigma_Y = \sigma_Z = 5.0$
Small	10 × 10	16000 to 40000	<0.5	23	14 square 9 triangle	$\sigma_X = \sigma_Y = 1.0$, $\sigma_Z = 2.0$ $\sigma_X = \sigma_Y = \sigma_Z = 5.0$

It is important to test and calibrate the system over the complete GSD range of applications, which is typically 3 to 50 cm. This is necessary, because environmental conditions are different at different flying altitudes and different components limit the performance on different scales, e.g., image motion limits the image quality especially at large scales. The influence of environmental conditions can be different for different systems; they should be factored into the evaluation of each system.

System Testing at Sjökulla

The Sjökulla test field has been part of the quality systems of Finnish image producers since 1994. In the analog era, the test field was used for geometry and spatial resolution testing. Recommendations for the use of the original test field and the first results were published by Ahokas *et al.* (2000).

Recently, use of the test field has increased significantly. In 2002, the National Land Survey of Finland (NLS) started calibrations with integrated camera/GPS/IMU systems. The results of test flights with the NLS systems were reported by Honkavaara *et al.* (2003).

The first test flights with digital cameras were performed at Sjökulla in 2004, and since then, several systems have been tested (Table 3). In many of these tests, an analog photogrammetric camera was operated simultaneously. In some cases, reflectance measurements were performed in the

Figure 4. Target types at the Sjökulla test field. From top to bottom: circle (terrestrial, GSD 5 cm, GSD 8 cm), square (terrestrial, GSD 25 cm, GSD 50 cm), and triangle (terrestrial, GSD 25 cm, GSD 50 cm). Processing may have affected image quality.

field simultaneously with the test flights. The tests were performed in a comprehensive way, carefully considering the factors affecting the image quality. The first results of these tests were published by Honkavaara *et al.* (2005, 2006a, 2006b, 2006c, and 2006d) and Markelin *et al.* (2005).

An overview of the complete system calibration and testing process is given in Figure 5. The process begins with mission planning. In the preparation phase, for instance, portable test targets are carried to the test field. During the flight mission, the weather, radiation, and reflectance measurements can be performed, and GPS reference data is collected. After the mission, the collected images are postprocessed, and aerial triangulation is performed. After the aerial triangulation, calibration and testing steps can be performed automatically if the positions of the targets are known. Finally, the results of testing and calibration are analyzed. In principle, camera operators can perform this procedure independently, but if portable targets are used, the assistance of the FGI is needed then. Honkavaara *et al.* (2006b) recently demonstrated the complete calibration concept using the Intergraph DMC.

Spatial Resolution Testing and Calibration
Parameters
The optical transfer function (OTF) and the resolving power (RP) of the lens are evaluated in the laboratory calibrations of photogrammetric cameras (Boland *et al.*, 2004). It is reasonable to include similar parameters in the field calibration and testing as well; the RP, modulation transfer function (MTF), point-spread function and edge response should be evaluated. Various quality indicators can be derived from the above, e.g., full width of half maximum, 10 percent MTF, area-weighted average resolution and modulation (AWAR, AWAM), and the general image quality equation (GIQE). (Leachtenauer *et al.*, 1997; Ryan *et al.*, 2003; USGS, 2004; Becker *et al.*, 2005; Honkavaara *et al.*, 2006d).

Resolution Test Targets
At the Sjökulla test field, the resolution is determined with the permanent dense and sparse bar targets and the portable Siemens star. Details of the figures are summarized in Table 4 and discussed below.

Dense and sparse permanent bar targets are frames made of timber (bar width >12 cm) or steel (bar width ≤12 cm) filled with gravel. Figures are either 3- or 4-bar square-wave targets. Low (black/gray: ≈1:2) and high (black/white: ≈1:8) contrasts are available. Two perpendicular sets of test figures exist, so the spatial resolution can be analyzed in the flying direction and in the cross-flight direction. The dense figures (Figure 6) were designed for precise resolution determination for large-scale imagery (GSD 3 to 8 cm). The width of the bars varies from 3 cm to 12 cm; the bar width increment is $\sqrt[6]{2}$ (≈12 percent), which is generally considered appropriate (e.g., ISO 3334; MIL-STD-150). The length of each bar is 1 m, which makes automation of the measurement process easier. The gravel size is 4 to 8 mm. The sparse figures were
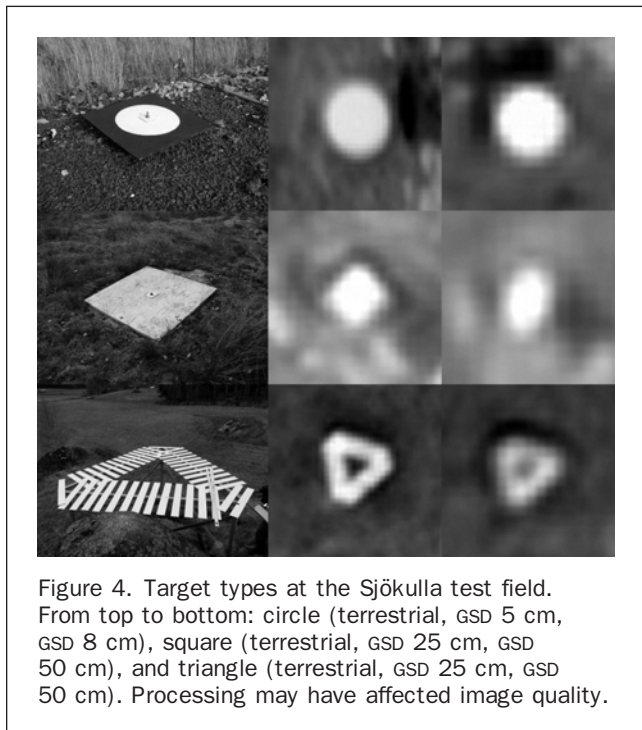
TABLE 3. DIGITAL SYSTEMS TESTED AT SJÖKULLA IN 2004–2005

System Camera	Mount	GPS/IMU	Date	GSD (cm)
UltraCamD	Non-stab.	–	11.10.2004	4
UltraCamD + RC20	Non-stab.	–	14–15.10.2004	4, 8, 25, 50
UltraCamD	Non-stab.	GPS	14.5.2005	4
Emerge DSS + laser	Non-stab.	GPS/IMU	12–14.7.2005	16, 50
DMC + RC20 + goniometer	Stabilized	–	1–2.9.2005	5, 8, 25, 50
ADS40 + RC20 + goniometer	Stabilized	GPS/IMU	26–27.9.2005	15, 25

designed for coarse resolution validation purposes for a wide scale range (GSD 3 cm to 1 m). The widths of the bars are 1.5 m, 1.25 m, 0.75 m, 0.38 m, 0.19 m, 0.13 m, 0.07 m, and 0.03 m. The gravel size is 8 to 16 mm.

The portable Siemens star (a semicircle) has 10° sectors and a 6.8 m radius, leading to a maximum width of a sector of 1 m (Figure 7). The Siemens star was sewn from black and gray poly-acrylic fabric. The materials have steady reflectance in the visible and NIR areas; the contrast is 1:6 to 1:11 in the wavelength range 400 to 1,000 nm (Figure 8). The star is attached to the ground with elastic bands and steel poles. The weight of the fabric is approximately 290 g/m^2; the total weight of the Siemens star is 21 kg. The Siemens star has been used for evaluation of the MTF of images with a GSD of 0.25 m or less. The use of the semicircular Siemens star has been a functional approach.

Figure 5. Process flow of field calibration and testing at Sjökulla.

The bar targets and Siemens star are viable tools for evaluating the resolution. The Siemens star has several advantages over the bar targets. If the sector size is small enough (e.g., 5° or 10°), the star target is multi-directional, allowing calculation of the resolution in all interesting directions. Furthermore, the Siemens star provides continuous resolution, while the bar targets are discrete. Automation of the measurement of the Siemens star is easier than with bar target.

The dimensions and scale range of the test figures are important factors, which affect on the one hand the accuracy of the resolution measurement and on the other hand the cost, size, and maintenance of the test field. The minimum line width in the bar figure and the Siemens star should correspond to the smallest GSD being tested. Empirical studies at Sjökulla have shown that an appropriate maximum line width for the RP evaluation is 2 to 3 times the largest GSD that is being tested. For the MTF determination, the required maximum line width is dependent on the system MTF, which in the case of digital systems is still a research issue. With analog systems, to obtain the complete MTF, the maximum line width should correspond to the RP of 2 to 5 line pairs/mm on the smallest scale being tested (Read and Graham, 2002).

Possible improvements at the Sjökulla test field are to construct a permanent Siemens star with 10° sectors and a maximum sector width of 2 m or more, and to improve the reflectance targets (Figure 10) so that they can be used as edge targets. Furthermore, the maximum bar width of the permanent dense bar target figure (Figure 6) should be increased to 1 or 1.5 m. Based on the results given in Spectral Properties of Gravel section, the best materials for these targets are gray and black gravel. With the above extensions, the Sjökulla test field could be considered as an airborne realization of the ISO standard 12233:2000.

Test Flights
A block of images is needed to evaluate the resolution. The flight lines are designed so that the resolution can be determined in flight and cross-flight directions. Overlaps and locations of the flight lines are selected so that the test targets are located in various parts of the focal plane; various flying heights are evaluated (Table 4; Honkavaara *et al.*, 2006d).

Methods for Spatial Resolution Evaluation
Due to the disadvantages of the visual measurement of RP from bar targets (Hakkarainen, 1991), an automatic measurement method has been developed (Kuittinen *et al.*, 1996; Ahokas *et al.*, 2000; Honkavaara *et al.*, 2006d). Several profiles are created over the bar figure. By calculating the differences between dark and light bars, standard deviations, and the geometric correctness of the figure, the software determines the thinnest line that can be detected.

To determine the MTF, the contrast transfer function is obtained either from the dense bar targets or from the Siemens star as the ratio of the image and the object modulations, and converted to MTF (Coltman, 1954; Hakkarainen
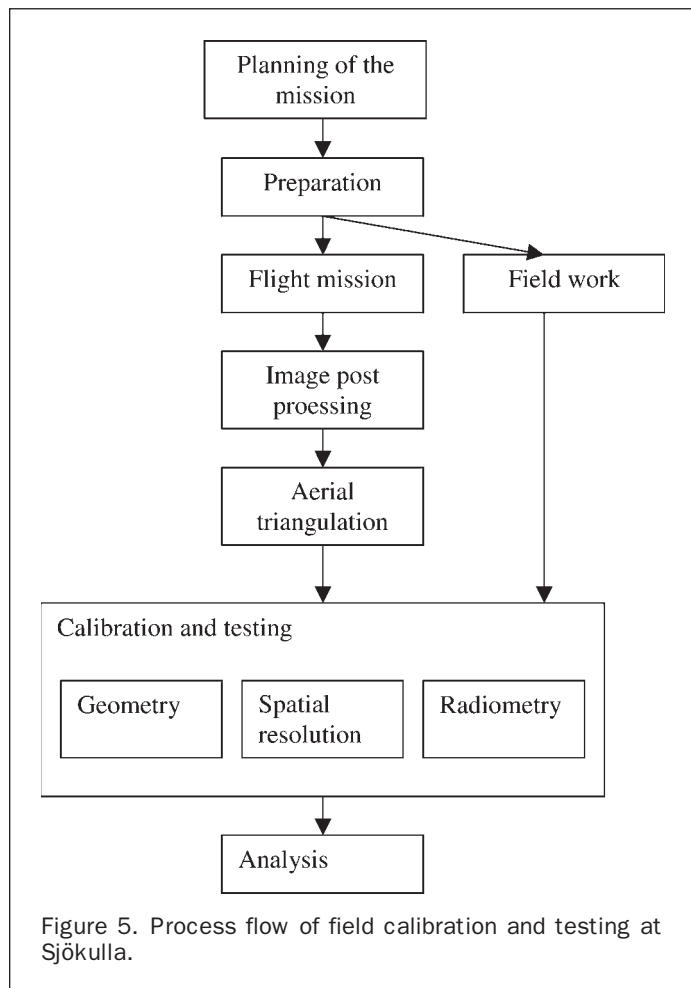
TABLE 4. CHARACTERISTICS OF THE SJÖKULLA RESOLUTION TARGETS

Target	Gravel Diameter (mm)	Contrast Ratio	Bar Width (m)	Bar Length (m)	Increment	RP Range (l/mm)	Scale Range***	GSD Range (m)
Dense	4 to 8	1:2	0.03–0.12	1	$\sqrt[6]{2}$	25–100*	<1:6 000	0.03 to 0.06
Dense	4 to 8	1:8	0.03–0.12	1	$\sqrt[6]{2}$	25–100*	<1:6 000	0.03 to 0.06
Sparse	8 to 16	1:2	0.03–1.5	1–6	irregular	40–80**	<1:60 000	0.03 to 1
Sparse	8 to 16	1:8	0.03–1.5	1–6	irregular	40–80**	<1:60 000	0.03 to 1
Siemens	–	1:6 to 1:11	0–1	–	continuous		<1:25 000	0.01 to 0.25

*scale 1:3 000
**scale 1:60 000
***Conventional analog systems
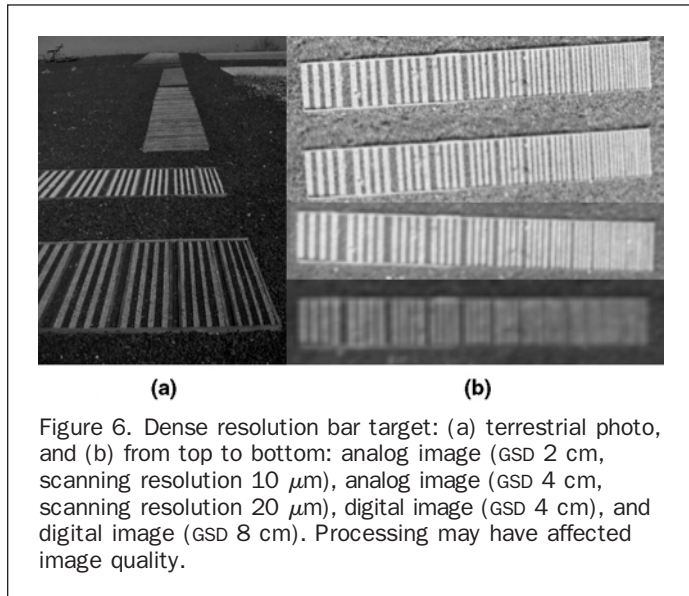
Figure 6. Dense resolution bar target: (a) terrestrial photo, and (b) from top to bottom: analog image (GSD 2 cm, scanning resolution 10 μm), analog image (GSD 4 cm, scanning resolution 20 μm), digital image (GSD 4 cm), and digital image (GSD 8 cm). Processing may have affected image quality.

Figure 8. Reflectance of some polyacrylic fabrics at nadir. Fabrics 34249 and 34205 were selected for the Siemens star.

Figure 7. Portable Siemens star. Left to right: terrestrial photo, air photo with 4 cm, 8 cm, 25 cm, and 50 cm GSD. Processing may have affected image quality.

and Rosenbruch, 1982). The MTF curve can be calculated for all directions or in various directions (e.g., flight, cross-flight). When calculating the MTF for all directions, in the case of a semicircle, only a quarter circle is used to avoid weighting any direction too much (Honkavaara et al., 2006d).

Radiometric Testing and Calibration
Parameters
The laboratory calibration of analog cameras does not include the radiometric component. The situation is different with digital sensors; radiometric calibration is an inherent part of the calibration of CCD sensors. The general objective of the radiometric calibration is to obtain the functional relationship between the incoming radiation and the instrument output digital number (DN). In practice, the radiometric laboratory calibration includes the determination of the
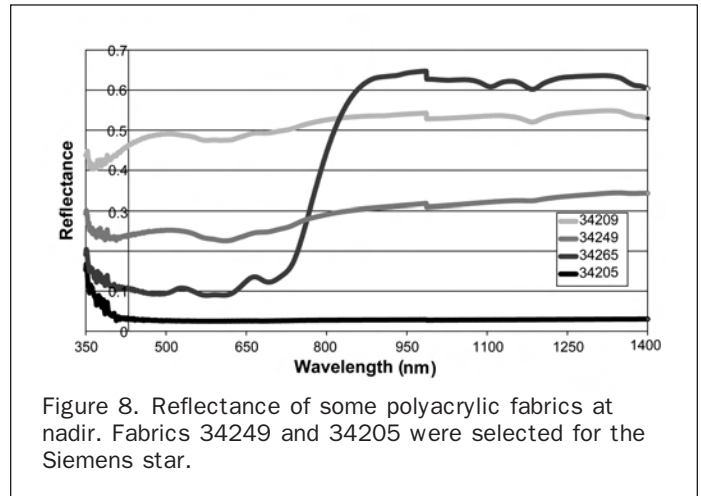
corrections for the sensitivity of each CCD element, defect pixels, aperture, and vignetting, and the determination of spectral response and absolute calibration parameters (Schuster and Braunecker, 2000; Diener et al., 2000; Heier et al., 2002). In the field calibration and testing, it is feasible to evaluate linearity, dynamic range, noise, uniformity, and stability of the system and to determine the absolute calibration parameters. Field calibration of individual CCD elements of high-resolution airborne sensors is difficult.

Reflectance Reference Targets
Reflectance reference targets at the Sjökulla test field are portable reflectance targets (gray scale), permanent reflectance targets, and a permanent gray scale.

The portable reflectance targets (Figure 9) were originally designed for testing analog aerial cameras for the wavelength area of 400 to 800 nm. Altogether, eight targets were manufactured with nominal reflectances of 5 percent, 10 percent, 20 percent, 25 percent, 30 percent, 45 percent, 50 percent, and 70 percent. The targets were made of polyester 1100 decitex (DTEX) with polyvinyl chloride (PVC) coating weighing 600 g/m². The targets were coated with titanium dioxide and carbon-black paint mixing pigments. A de-lustering agent was added to the paint to obtain a matte surface and to decrease the non-Lambertian reflectance effects. The size of one target is 5 m × 5 m. The tarpaulins can be attached together in a line; the number and combination of targets can be varied depending on the available space. BRF laboratory measurements of the targets showed that reflectance was clearly directionally dependent, with a gentle bowl shape, i.e., darkest at the zenith and brightening at larger zenith angles (Markelin, unpublished data, 2006). There was also some wavelength dependence, especially in the forward
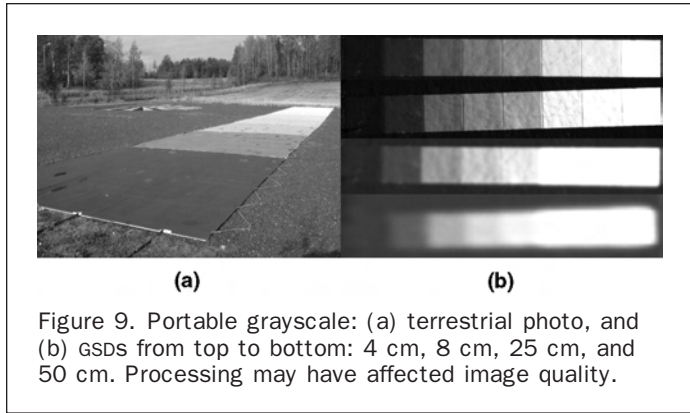
Figure 9. Portable grayscale: (a) terrestrial photo, and (b) GSDs from top to bottom: 4 cm, 8 cm, 25 cm, and 50 cm. Processing may have affected image quality.
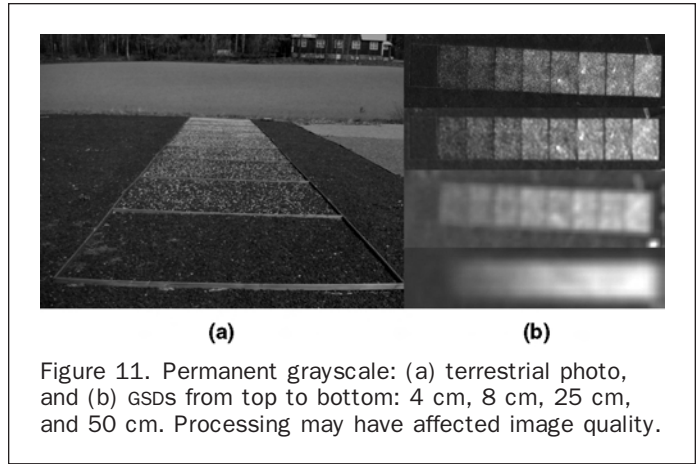

Figure 11. Permanent grayscale: (a) terrestrial photo, and (b) GSDs from top to bottom: 4 cm, 8 cm, 25 cm, and 50 cm. Processing may have affected image quality.

direction, and outside the design wavelength range of 400 to 800 nm. The reflectances for various color channels integrated with the spectral sensitivities of the Intergraph DMC digital sensor, taken at nadir with a solar zenith angle of 56° from vertical, are shown in Figure 12. The reflectances were fairly similar to the nominal values. The reflectance of the tarps slightly decreased as the wavelength increased; as a consequence, for instance, the tarps are the brightest at the blue channel and the darkest at the NIR channel. The same grayscale has recently been used for laser calibration (Kaasalainen *et al.*, 2005). The transportable grayscale has a wide reflectance range and a high radiometric resolution; it is suitable for comprehensive testing and calibration tasks for GSDs of 0.5 m or less.

The large rectangular reflectance targets of gray, black, red, and white gravel are 15 m × 7 m in size (Figure 10). The spectral and radiometric properties of these materials were previously described. The reflectance targets are suitable for radiometric calibration and testing for GSDs of 0.7 m or less.

The permanent grayscale, consisting of nine steps 2 m × 3 m in size, was constructed of wooden frames filled by mixing white and black gravel (Figure 11). Special problems of the permanent grayscale are that the reflectance of the steps are not uniform due to non-uniform mixing of the gravel, the topography of the figure, non-uniform changes of gravel, and the size of the gravel; accurate radiometric calibration of this kind of target is difficult.

Based on the FGI experience, the best concept for constructing permanent gravel reflectance targets is to use a single gravel type for each step rather than to mix various gravels. It appears that it is possible to obtain

a larger reflectance range with the painted figures; so, painted figures probably are needed for the most demanding applications. The desirable reflectance area is 3 to 90 percent, or even larger.

Test Flights
The test flights should be designed so that the targets are located in various parts of the focal plane in order to evaluate the uniformity of the system; for the stability evaluation, several images are needed.

Method for Radiometric Calibration and Testing
The measured BRF must be used for all calibration and validation related to image radiometry, spectrometry/color properties, and/or contrast. For utilizing the BRF information, the direction of the sun, the accurate direction of observation for each pixel, and the amount of diffuse radiation must be known. Simultaneous *in situ* reflectance measurements are often preferable. For the most accurate analysis, detailed atmospheric properties must also be determined. To make use of the test field more self-sufficient, equipment for automatic *in situ* radiation measurement should be permanently installed at the test field.
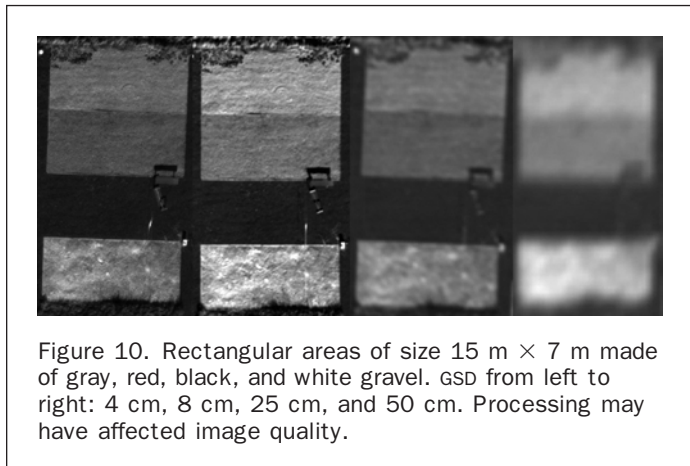

Figure 10. Rectangular areas of size 15 m × 7 m made of gray, red, black, and white gravel. GSD from left to right: 4 cm, 8 cm, 25 cm, and 50 cm. Processing may have affected image quality.
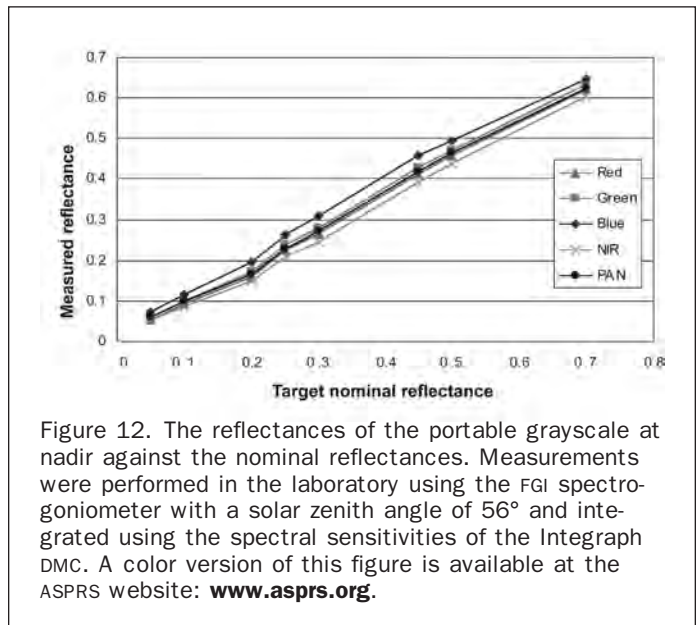

Figure 12. The reflectances of the portable grayscale at nadir against the nominal reflectances. Measurements were performed in the laboratory using the FGI spectrogoniometer with a solar zenith angle of 56° and integrated using the spectral sensitivities of the Integraph DMC. A color version of this figure is available at the ASPRS website: **www.asprs.org**.

At Sjökulla the most versatile radiometric evaluation can be performed using the portable grayscale (Figure 9). Statistics (average and standard deviation of DNs) are calculated at each reflectance step of the gray scale. To evaluate the dynamic range and linearity, the average DNs are plotted against the reference reflectance values. By evaluating standard deviation in each tarpaulin, the noise of the system is obtained; special methods are needed to eliminate the topographic effects of flexible tarpaulins (Figure 9). For the absolute calibration and the most comprehensive analysis, the surface reflectances should be converted to at-sensor radiances by utilizing the atmospheric information. Absolute calibration parameters can be determined by solving, e.g., the linear parameters between DNs and radiances. This calibration procedure is similar to the methods used to calibrate satellite and video sensors (Dianguirard and Slater, 1999; Biggar et al., 2003; Pagnutti et al., 2003; Edirisinghe et al., 2004) and to method used by USGS (2004) to characterize airborne sensors. Preliminary results from use of the FGI reflectance targets for the testing and calibration of an UltraCamD photogrammetric camera were reported by Markelin et al. (2005).

Geometric Calibration and Testing
Parameters
Geometric laboratory calibration of analog cameras includes determination of the lens parameters, i.e., focal length, principal point, and radial and tangential lens distortions (Boland et al., 2004; Cramer, 2004). Calibration of the digital sensors has required some changes in laboratory calibration methods (Pacey et al., 1999; Schuster and Braunecker, 2000; Heier et al., 2002; Cramer, 2004; Kröpfl et al., 2004).

In the field calibration and testing, the lens parameters should be determined. A complication with the new multi-head digital systems, such as Vexcel UltraCamD and Intergraph DMC, is that the traditional self-calibration parameters do not model their systematic errors to a sufficient extent; for these systems new parameterization must be developed. If the photogrammetric system includes GPS and IMU, the misalignments (boresight angles and lever arms) must also be calibrated in airborne conditions. An inherent part of geometric calibration and testing is accuracy evaluation. Results of geometric accuracy evaluation are statistics of point determination and back projection accuracy and statistics of the accuracy of external orientation observations. Honkavaara et al. (2003, 2006a, and 2006c).

Test Flights
For geometric calibration and testing, an image block, or preferably several blocks taken from different flying heights, is needed. At Sjökulla, the geometric calibration is possible at large, medium, and small scales (Table 2). With the analog cameras, the possible calibration scales are from 1:3 000 to 1:40 000. The feasible GSD range for digital cameras is 3 to 50 cm. At large and medium scales, the recommended approach is to use a block with two to four flight lines and two to four cross flight lines with overlap percentages of 60 to 80 percent. The small-scale test field was optimized for a cross-block with two perpendicular bi-directional flight lines, but a block with parallel and cross flight lines can also be collected. The block with high side overlap is ideal for the calibration and evaluation of point determination accuracy in the most accurate tasks. To evaluate performance in normal mapping applications, blocks with lower side overlaps should be used. (Figure 1; Honkavaara, 2003).

Method for Geometric Calibration and Testing
Geometric calibration and testing is performed with a self-calibrating bundle block adjustment (Kraus, 1997; Cramer, 2004; Förstner et al., 2004). As discussed above, the new

types of sensors require careful consideration (Heier et al., 2002; Kröpfl et al., 2004, Honkavaara et al., 2006a and 2006c). Details of the GCPs at Sjökulla are given in the Geometric Test Field section. At Sjökulla, if system calibration is performed, all the benchmarks are used as GCPs. If the test field is used for accuracy evaluation, some of the benchmarks are used as GCPs, and the rest are used as checkpoints. Only the large-scale test field has a sufficient number of GCPs for accuracy evaluation. In the medium and small-scale fields, special error estimation techniques, such as the leave-one-out method (Fukunaga, 1990), should be used to obtain good GCP distribution and the highest possible number of checkpoints. Existing laser data can be used as an additional height control.

Conclusions
In this article, the design principles, characteristics, and use of the FGI permanent Sjökulla test field were described. This test field with some supplementary targets is a prototype of a future field calibration and testing site for geometry, spatial resolution, and radiometry of digital photogrammetric systems. The principles presented can be used as the basis for new permanent test fields and field calibration methods. The same methodology can also be used for calibrating and testing of other airborne and satellite instruments with appropriate resolution (GSD 0.03 to 0.5 m) and spectral properties (wavelength preferably 400 to 1,000 nm).

Ten years of experience in Sjökulla has shown that gravel, combined with a proper substructure, form a weather-resistant solution for constructing spectrally, radiometrically, and geometrically stable permanent targets for spatial resolution and reflectance testing. The gravel targets are durable even in extreme northern conditions with frost, rain, and snow. The particularly difficult task is to construct a high-resolution permanent grayscale with a wide reflectance range (3 to 90 percent); for the time being, portable painted gray scale will be used at Sjökulla.

The field calibration and testing of photogrammetric systems is going to be crucial in the future. It is to be expected that numerous versatile permanent photogrammetric test fields will be established. A permanent test field is justified especially if there large numbers of systems are to be tested (e.g., by the camera manufacturers and national mapping authorities). To make use of these test fields easier from the image producers' and users' point of view, standardized test targets and methods should be established. The targets and methods used at Sjökulla have a lot of potential, but empirical results from digital systems are needed to specify the methodology in detail. One of the most fundamental issues in the development of the calibration practices is the stability of the systems.

Our scenario for the future calibration process of digital photogrammetric sensors/systems is the following. After the sensor is constructed, it is calibrated in the laboratory and tested in the field by the camera manufacturer. After installation, the entire system is field-tested and calibrated. Calibration and testing in a permanent test field is performed annually. Finally, each collected image is self-calibrated. The properties of the sensor/system are documented over time, and if significant changes are detected, corrective actions are taken.

Recently, the Sjökulla test field has been used for calibrating and testing several digital photogrammetric sensors, and the complete calibration concept has been demonstrated in practice. The evaluation of the collected material is in progress and the detailed results will be published in the near future. Results have already shown the importance of field-testing.

Acknowledgments

Author Contributions

E. Honkavaara has conducted and participated in the recent calibration and testing method development and analysis, and compiled this article. J. Peltoniemi has supervised the BRF measurements and processed and analyzed the BRF data. E. Ahokas is one of the major builders and maintainers of the Sjökulla test field, and he has developed methods for testing of analog systems and for radiometric testing of digital sensors. R. Kuittinen invented the Sjökulla test field and the gravel targets, and he has supervised the method development and analysis, particularly during the analog era. J. Hyyppä has supervised the test field and method development. J. Jaakkola is the author of the software for the spatial resolution measurement and he has participated in the spatial resolution analysis. H. Kaartinen has maintained and improved the Sjökulla test field. L. Markelin has performed BRF field measurements, developed methods for radiometric and spatial resolution testing, and performed radiometric analysis. K. Nurminen has investigated the geometric calibration of oblique images. J. Suomalainen has performed BRF field measurements.

References

Ahokas E., R. Kuittinen, and J. Jaakkola, 2000. A system to control the spatial quality of analogue and digital aerial images, *International Archives of Photogrammetry and Remote Sensing*, 33(4):45–52.

Becker, S., N. Haala, and R. Reulke, 2005. Determination and improvement of spatial resolution for digital aerial images, *Proceedings of ISPRS Hannover Workshop 2005, High-Resolution Earth Imaging for Geospatial Information*, unpaginated CD-ROM.

Biggar, S.F., K.J. Thome, and W. Wisniewski, 2003. Vicarous radiometric calibration of EO sensors by reference to high-reflectance ground targets, *IEEE Transactions on Geoscience and Remote Sensing*, 41(6):1174–1179.

Boland, J., T. Ager, E. Edwards, E. Frey, P. Jones, R.K. Jungquiet, A.G. Lareau, J. Lebarron, C.S. King, K. Komazaki, C. Toth, S. Walker, E. Whittaker, P. Zavattero, and H. Zuegge, 2004. Cameras and sensing systems, *ASPRS Manual of Photogrammetry*, Fifth Edition, (J.C. McGlone, E. Mikhail, and J. Bethel, editors), American Society for Photogrammetry and Remote Sensing, pp. 581–676.

Casella, V., and M. Franzini, 2005. Experiences in GPS/IMU calibration: Rigorous and independent cross-validation of results, *Proceedings of ISPRS Hannover Workshop 2005, High-Resolution Earth Imaging for Geospatial Information*, unpaginated CD-ROM.

Coltman, J.W., 1954. The specification of image properties by response to sine wave input, *Journal of the Optical Society of America*, 44(6):468–471.

Cramer, M., 2004. EuroSDR network on digital camera calibration – Report Phase I (Status 26 October 2004), University of Stuttgart, Institute for Photogrammetry, URL: http://www.ifp.uni-stuttgart.de/eurosdr/index.en.html (last date accessed: 21 October 2007).

Cramer, M., 2005. 10 Years of ifp test site Vaihingen/Enz: An independent performance study, *Proceedings of Photogrammetric Week '05*, Herbert Wichmann Verlag, pp. 79–92.

Dianguirard, M., and P.N. Slater, 1999. Calibration of space-multispectral imaging sensors: A Review, *Remote Sensing of Environment*, 68(3):194–205.

Diener, S., M. Kiefner, and C. Dörstel, 2000. Radiometric normalisation and colour composite generation of the DMC, *International Archives of Photogrammetry and Remote Sensing*, 33(1):82–88.

Dörstel, C., 2003. DMC – Practical experiences and photogrammetric system performance, *Proceedings of Photogrammetric Week '03*, Herbert Wichmann Verlag, pp. 59–65.

Edirisinghe, A., J.P. Louis, and G.E. Chapman, 2004. Potential for calibrating airborne video imagery using preflight calibration coefficient, *Photogrammetric Engineering & Remote Sensing*, 70(5):573–580.

Förstner, W., B. Wrobel, F. Paderes, R. Craig, C. Fraser, and J. Dolloff, 2004. Analytical photogrammetric operations, *ASPRS Manual of Photogrammetry*, Fifth Edition, (J.C. McGlone, E. Mikhail, and J. Bethel, editors), American Society for Photogrammetry and Remote Sensing, pp. 763–936.

Fukunaga, L., 1990. Parameter estimation, *Introduction to Statistical Pattern Recognition*, Second edition, Academic Press, Inc., Boston, pp. 181–249.

Hakkarainen, J., and K.J. Rosenbruch, 1982. Image quality and lens aberrations of an aerial camera, *Photogrammetria*, 38(3):87–102.

Hakkarainen, J., 1991. RP Comparison between black-and-white negative and colour diapositive films, *Surveying Science in Finland*, 9(1):3–15.

Heier, H., M. Kiefner, and W. Zeitler, 2002. Calibration of Digital Modular Camera, *Proceedings of FIG XXII International Congress*, Washington, D.C., 19–26 April, unpaginated CD-ROM.

Heipke, C., K. Jacobsen, and H. Wegmann, 2002. Analysis of the results of the OEEPE test - Integrated sensor orientation, *OEEPE Official Publication* (C. Heipke, K. Jacobsen, and H. Wegmann, editors), 43:31–49.

Honkavaara, E., R. Ilves, and J. Jaakkola, 2003. Practical results of GPS/IMU/camera system calibration, *Proceedings of Workshop: Theory, Technology and Realities of Inertial/GPS/Sensor Orientation*, ISPRS WG I/5, Barcelona, unpaginated CD-ROM.

Honkavaara, E., 2003. Calibration field structures for GPS/IMU/camera-system calibration, *The Photogrammetric Journal of Finland*, 18(2):3–15.

Honkavaara, E., L. Markelin, R. Ilves, P. Savolainen, J. Vilhomaa, E. Ahokas, J. Jaakkola, and H. Kaartinen, 2005. In-flight performance evaluation of digital photogrammetric sensors, *Proceedings of ISPRS Hannover Workshop 2005: High-Resolution Earth Imaging for Geospatial Information*, unpaginated CD-ROM.

Honkavaara, E., E. Ahokas, J. Hyyppä, J. Jaakkola, H. Kaartinen, R. Kuittinen, L. Markelin, and K. Nurminen, 2006a. Geometric test field calibration of digital photogrammetric sensors, *ISPRS Journal of Photogrammetry and Remote Sensing*, Special Issue on Digital Photogrammetric Cameras, 60(6):387–399.

Honkavaara, E., J. Jaakkola, L. Markelin, J. Peltoniemi, E. Ahokas, and S. Becker, 2006b. Complete photogrammetric system calibration and evaluation in the Sjökulla test field-Case study with DMC, *Proceedings of EuroSDR Commission I and ISPRS Working Group 1/3 Workshop EuroCOW 2006*, unpaginated CD-ROM.

Honkavaara, E., J. Jaakkola, L. Markelin, K. Nurminen, and E. Ahokas, 2006c. Theoretical and empirical evaluation of geometric performance of multi-head large format photogrammetric sensors, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(A1).

Honkavaara, E., J. Jaakkola, L. Markelin, and S. Becker, 2006d. Evaluation of resolving power and MTF of DMC, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36(A1).

ISO 3334:1989. *Micrographics – ISO Resolution Test Chart No. 2 – Description and Use.*

ISO 12233:1999. *Photography. Electronic Still-picture Cameras. Resolution Measurements.*

Kaasalainen, S., E. Ahokas, J. Hyyppä, and J. Suomalainen, 2005. Study of surface brightness from backscattered laser intensity: Calibration of laser data, *IEEE Transactions on Geoscience and Remote Sensing Letters*, 2(3):255–259.

Kraus, K., 1997. *Photogrammetry, Volume 2, Advanced Methods and Applications*, 1997 Ferd-Dümmlers Verlag.

Kröpfl, M., E. Kruck, and M. Grüber, 2004. Geometric calibration of the digital large format aerial camera UltraCamD, *International Archives of Photogrammetry and Remote Sensing*, 35(1):42–44.

Kuittinen R., E. Ahokas, A. Högholen, and J. Laaksonen, 1994. Test field for aerial photography, *The Photogrammetric Journal of Finland*, 14(1):53–62.

Kuittinen R., E. Ahokas, and P. Järvelin, 1996. Transportable test-bar targets and microdensitometer measurements: A method to control the quality of aerial imagery, *International Archives of Photogrammetry and Remote Sensing*, 31(1):99–104.

Leachtenauer, J.C., W. Malila, J.M. Irvine, L.P. Colburn, and N.L. Salvaggio, 1997. General image-quality equation: GIQE, *Applied Optics*, 36(32):8322–8328.

Markelin, L., E. Ahokas, E. Honkavaara, A. Kukko, and J. Peltoniemi, 2005. Radiometric quality comparison of UltraCamD and analog camera, *Proceedings of ISPRS Hannover Workshop 2005: High-Resolution Earth Imaging for Geospatial Information*, unpaginated CD-ROM.

Merchant, D.C, A. Schenk, A. Habib, and T. Yoon, 2004. USGS/OSU progress with digital camera in situ calibration methods, *International Archives of Photogrammetry and Remote Sensing*, 35(2):19–24.

MIL-STD-150A, 1959. *Military Standard – Photographic Lenses*, United States Government Printing Office, Washington, D.C.

Morain, A.S., and M.V. Zanoni, 2004. Joint ISPRS/CEOS-WGCV task force on radiometric and geometric calibration, *International Archives of Photogrammetry and Remote Sensing*, 35(1):354–360.

Moran, M.S., B.B Ross, T.R. Clarke, and J. Qi, 2001. Deployment and calibration of reference reflectance tarps for use with airborne imaging sensors, *Photogrammetric Engineering & Remote Sensing*, 67(3):273–286.

Nielsen, B. Jr., 2002. Test field Fredrikstad and data acquisition for the OEEPE test – Integrated sensor orientation, Integrated sensor orientation – Test report and workshop proceedings, *OEEPE Official Publication* (C. Heipke, C. Jacobsen, and H. Wegmann, editors), 43:19–30.

Pacey, R.E., M. Scheidt, and A.S Walker, 1999. Calibration of analogue and digital airborne sensors at LH Systems, *Proceedings of the 1999 ASPRS Annual Conference*, Portland, Oregon, pp. 950–956.

Pagnutti, M., R.E. Ryan, M. Kelly, K. Holekamp, V. Zanoni, K. Thome, and S. Schiller, 2003. Radiometric characterization of IKONOS multispectral imagery, *Remote Sensing of Environment*, 88(1):53–68.

Peltoniemi, J., S. Kaasalainen, J. Näränen, M. Rautiainen, P. Stenberg, H. Smolander, S. Smolander, and P. Voipio, 2005. BRDF measurement of understory vegetation in pine forests: Dwarf shrubs, lichen and moss, *Remote Sensing of Environment*, 94(3):343–354.

Read, R.E., and R.W. Graham, 2002. *Manual of Air Survey: Primary Data Acquisition*, Whittles Publishing, Caithness, 408 p.

Ryan, R., B. Baldridge, R.A. Schowengerdt, T. Choi, D.L. Helder, and S. Blonski, 2003. IKONOS spatial resolution and image interpretability characterization, *Remote Sensing of Environment*, 88(1):37–52.

Salamonowicz, 1982. USGS aerial resolution targets, *Photogrammetric Engineering & Remote Sensing*, 48(9):1469–1473.

Schuster, R., and B. Braunecker, 2000. Calibration of the LH Systems ADS40 airborne digital sensor, *International Archives of Photogrammetry and Remote Sensing*, 33(1):288–294.

Tempelmann, U., L. Hinsken, and U. Recke, 2003. ADS40 calibration and verification process, *Proceedings of International Workshop: Theory, Technology and Realities of Inertial/GPS Sensor Orientation*, Castelldefels, Spain, 22–23 September, unpaginated CD-ROM.

USGS, 2004. Remote sensing characterization project digital camera product characterization in situ test plan for aircraft flyover at NASA Stennis Space Center, Version 0.6, Department of the Interior, U.S. Geological Survey, EROS Data Center URL: *http://calval.cr.usgs.gov/documents/* (last date accessed: 22 October 2007).