PFG

# Photogrammetrie Fernerkundung Geoinformation

Organ der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF) e. V.

Jahrgang 2012, Heft 3




E. Schweizerbart'sche Verlagsbuchhandlung
(Nägele u. Obermiller) Stuttgart 2012

Deutsche Gesellschaft für Photogrammetrie, Fernerkundung
und Geoinformation (DGPF) e. V.
Gegründet 1909

Die *Deutsche Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation* (DGPF) e. V. unterstützt als Mitglieds- bzw. Trägergesellschaft die folgenden Dachverbände:

International Society
for Photogrammetry
and Remote Sensing

Deutsche Arbeits-
gemeinschaft für
Mustererkennung e.V.

GeoUnion
Alfred-Wegener-Stiftung

Herausgeber:

Präsident: Prof. Dr. Thomas Kolbe, Technische Universität Berlin, Institut für Geodäsie und Geoinformationstechnik, Straße des 17. Juni 135, 10623 Berlin, Germany, Tel.: +49-30-314-23274
Geschäftsstelle: Dr. Klaus-Ulrich Komp, c/o EFTAS Fernerkundung Technologietransfer GmbH, Oststraße 2–18, 48145 Münster, Germany, e-mail: klaus.komp@eftas.com
Published by: E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), Johannesstraße 3A, 70176 Stuttgart, Germany, Tel.: +49-711 351456-0, Fax: +49-711 351456-99, e-mail: mail@schweizerbart.de
Internet: http://www.schweizerbart.de
♾ Gedruckt auf alterungsbeständigem Papier nach ISO 9706-1994

ISSN 1432-8364
Science Citation Index Expanded (also known as SciSearch®) Journal Citation Reports/Science Edition

Erscheinungsweise: 6 Hefte pro Jahrgang.
Bezugspreis im Abonnement: € 219,– pro Jahrgang. Mitglieder der DGPF erhalten die Zeitschrift kostenlos. Der Online-Zugang ist im regulären Subskriptionspreis enthalten.
Anzeigenverwaltung: E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), Johannesstraße 3A, 70176 Stuttgart, Germany, Tel.: +49-711 351456-0; Fax: +49-711 351456-99.
e-mail: mail@schweizerbart.de, Internet: http://www.schweizerbart.de
Bernhard Harzer Verlag GmbH, Westmarkstraße 59/59a, 76227 Karlsruhe, Germany, Tel.: +49-721 944020, Fax: +49-721 9440230, e-mail: Info@harzer.de, Internet: www.harzer.de
Printed in Germany by Tutte Druckerei GmbH, 94121 Salzweg bei Passau, Germany.

# PFG – Jahrgang 2012, Heft 3
# Inhaltsverzeichnis

## Originalbeiträge



## Mitteilungen



Zusammenfassungen der „Originalbeiträge“ und der „Beiträge aus Wissenschaft und Praxis“ (deutsch und englisch) sind auch verfügbar unter www.dgpf.de/neu/pfg/ausgaben.htm

PFG 2012 / 3, 0209–0220
Stuttgart, June 2012

Article

# Minimal Representations for Testing and Estimation in Projective Spaces

Wolfgang Förstner, Bonn

**Keywords:** estimation, minimal representation, projective space

**Summary:** Testing and estimation using homogeneous coordinates and matrices has to cope with obstacles such as singularities of covariance matrices and redundant parametrisations. The paper proposes a representation of the uncertainty of all types of geometric entities which (1) only requires the minimum number of parameters, (2) is free of singularities, (3) enables to exploit the simplicity of homogeneous coordinates to represent geometric constraints and (4) allows to handle geometric entities which are at infinity or at least very far away. We develop the concept, discuss its usefulness for bundle adjustment and demonstrate its applicability for determining 3D lines from observed image line segments in a multi view setup.

**Zusammenfassung:** Beim Testen und Schätzen mit homogenen Koordinaten und Matrizen treten wegen der Redundanz der Repräsentationen und der daraus folgenden Singularität der Kovarianzmatrizen Schwierigkeiten auf. Der Beitrag schlägt eine Repräsentation für die Unsicherheit geometrischer Elemente vor, die (1) eine minimale Zahl von Parametern benötigt, (2) frei von Singularitäten ist, (3) die Einfachheit homogenener Koordinaten bei der Formulierung geometrischer Bedingungen belässt und (4) uneigentliche Elemente, d.h. Elemente im Unendlichen bzw. sehr weit entfernte Elemente behandeln kann. Wir stellen das Konzept vor, diskutieren seine Nützlichkeit bei der Bündelausgleichung und zeigen seine Anwendbarkeit für die Schätzung von 3D Geraden aus mehreren Bildern.

## 1 Introduction

Estimation of entities in projective spaces, such as points or transformations, has to cope with the *scale ambiguity* of the redundant representations of these entities, and, as a consequence, with the definition of *proper metrics* which results from the singularity of the covariance matrices. As an unwanted side effect the number of parameters heavily increases in large estimation problems. The paper shows how to consistently perform statistical testing and maximum likelihood (ML) estimation for geometric entities and transformations in projective spaces including elements at infinity while only handling the minimum of required parameters.

The *scale ambiguity* of homogeneous entities results from the redundant representation, where two elements, say 2D points, $x$ and $y$ or two rotations $\mathcal{R}$ and $\mathcal{R}'$ are identical, in case their representations with homogeneous coordinates, here with $\mathbf{x}$ and $\mathbf{y}$ or with quaternions, here $\mathbf{q}$ and $\mathbf{q}'$, are proportional. This ambiguity regularly is avoided by proper normalization of the homogeneous entities. Mostly one applies either Euclidean normalization of homogeneous coordinates, say $\mathbf{x}^e = \mathbf{x}/x_3$, see Kanatani (1996), then accepting that no elements at infinity can be represented, or spherical normalization, say $\mathbf{x}^s = \mathbf{x}/|\mathbf{x}|$ or $\mathbf{q}^s = \mathbf{q}/|\mathbf{q}|$, then accepting that the parameters to be estimated sit on a non-linear manifold, here the unit spheres $S^2$ or $S^3$, see Collins (1993), Heuel (2004). The sign ambiguity usually does not cause difficulties, as the homogeneous constraints used for reasoning are independent on the chosen sign.

The *uncertainty* of an observed geometric entity in many practical cases, can be repre-


www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0112
1432-8364/12/0112 $ 3.00

sented sufficiently well by a Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_{xx})$. The distribution of derived entities, $\underline{\mathbf{y}} = \boldsymbol{f}(\underline{\mathbf{x}})$, resulting from a non-linear transformation can also be approximated by a Gaussian distribution, using Taylor expansion at the mean $\boldsymbol{\mu}_x$ and omitting higher order terms, possibly requiring truncation of the given distribution, see Hartley & Zisserman (2000, App.3). The degree of approximation depends on the relative accuracy and has been shown to be negligible in many cases, see Criminisi (2001, p. 55).

The invariance of estimates w.r.t. the choice of the normalization of the estimated entities usually is achieved by minimizing a function in the Euclidean space of observations, in the context of bundle adjustment being the reprojection error, leading to the objective function $\Omega = \sum_i (\underline{\boldsymbol{x}}_i - \hat{\underline{\boldsymbol{x}}}_i)^\mathsf{T} \Sigma_{x_i x_i}^{-1} (\underline{\boldsymbol{x}}_i - \hat{\underline{\boldsymbol{x}}}_i)$. This at the same time is the Mahalanobis distance between the observed and estimated entities and can be used to evaluate whether the model fits the data. A similar reasoning is used when estimating transformations, such as rotations based on quaternions $\mathbf{q} = (q, \boldsymbol{q})$ or projective transformations, e.g. 2D homographies $\mathsf{H}$, where one of the redundant elements, say the scalar part $q$ or the last element $H_{33}$, is Euclideanly normalized to 1 in order to arrive at a minimal representation.

This situation becomes difficult, in case one wants to handle elements at infinity, thus homogeneous coordinate 3-vectors with $x_3 = 0$, rotations with $q \approx 0$, thus around 180°, or homographies with $H_{33} \approx 0$, and therefore one wants to use spherically normalized homogeneous vectors or matrices, or at least normalized direction vectors when using omnidirectional cameras, as their covariance matrices are or become close to singular.

Therefore, in case we want to use these normalized vectors or matrices as observed quantities, already the formulation of the objective function based on homogeneous entities is not possible and requires a careful discussion about estimable quantities (Meidow et al. 2009). Also the redundant representation requires additional constraints, which lead to Lagrangian parameters in the estimation process. As an example, one would need four parameters to estimate a 2D point, three for the homogeneous coordinates and one Lagrangian for the constraint, two parameters more than the degrees of freedom of a 2D point.

## Related work

This problem of representing uncertain transformations has been addressed successfully for geometric transformations. Common to these approaches is to represent the uncertain transformations, say a rotation $\underline{R}$ as multiplicative deviations $R(\underline{\boldsymbol{r}})$ from the mean transformation $\mathbb{E}(\underline{R})$ and to represent the small deviations as the exponential of a matrix, say $\exp(S(\underline{\boldsymbol{r}}))$, which allows simple estimation and rigorous update, a property resulting from the group properties of these transformations. Based on the work of Bregler & Malik (1998), Rosenhahn et al. (2002) used the exponential map for modelling spatial Euclidean motions, composed of rotations and translations in $\mathbb{R}^3$. Bartoli & Sturm (2004) used the idea to estimate the fundamental matrix with a minimal representation $\mathsf{F} = R_1 \,\mathrm{Diag}(\exp(\lambda), \exp(-\lambda), 0)\, R_2^\mathsf{T}$, twice using the rotation group and once the multiplication group $\mathbb{R}^+$. Begelfor & Werman (2005) showed how to estimate a general 2D homography with a minimal representation statistically rigorously, namely using the special linear group of $3 \times 3$-matrices with determinant 1, represented as $\mathsf{H} = \exp(K)$ with matrices $K$ having trace zero, correctly reflecting the number of degrees of freedom, see the application in Meidow (2011).

To my knowledge the only attempts to use minimal representations for geometric entities other than transformations have been given by Sturm & Gargallo (2007) and Åstrom (1998), however, both are not able to represent elements at infinity, namely conics and points at infinity respectively.

This paper presents a concept for statistical testing and estimation with all types of geometric entities in projective spaces using minimal representations which are free of singularities and allow to handle entities at infinity (Förstner 2010 a,b).

## Notation

We name objects with calligraphic letters, say a point $\mathcal{X}$. We denote Euclidean coordinates

with a slanted letter $\boldsymbol{x}$, homogeneous coordinates with an upright letter $\mathbf{x}$. Matrices are denoted with sans serif capital letters, say $\mathsf{R}$, or upright in case of homogeneous matrices, e.g. H. The operator $\mathsf{N}(\cdot)$ normalizes a vector to unit length. We adopt the Matlab syntax to denote the stack of two vectors or matrices, e.g. $\boldsymbol{z} = [\boldsymbol{x};\boldsymbol{y}] = [\boldsymbol{x}^\mathsf{T},\boldsymbol{y}^\mathsf{T}]^\mathsf{T}$. The vec-operator stacks the columns of a matrix to obtain a vector: $\mathrm{vec}(\mathsf{A}) = \mathrm{vec}([\boldsymbol{a}_1,\ldots,\boldsymbol{a}_n]) = [\boldsymbol{a}_1;\ldots;\boldsymbol{a}_n]$. Stochastic variables are underscored, e.g. $\underline{\boldsymbol{x}}$. We use the skew symmetric matrix $\mathsf{S}(\boldsymbol{a})$ of a 3-vector, inducing the cross product $\boldsymbol{a}\times\boldsymbol{b} = \mathsf{S}(\boldsymbol{a})\,\boldsymbol{b}$. As an exception, we denote three-dimensional rotation vectors – the product of the rotation angle with the normalized rotation axis – with $\boldsymbol{r}$ and four dimensional quaternion vectors with $\mathbf{q}$, instead with capital letters.

## 2 Minimal Representation of Uncertainty

The natural spaces of homogeneous entities are the unit spheres $S^n$, possibly constrained to a subspace. Spherically normalized homogeneous coordinates of 2D points ($\mathbf{x}^s$) and lines ($\mathbf{l}^s$) live on the sphere $S^2$ in $\mathbb{R}^3$, those of 3D points ($\mathbf{X}^s$) and planes ($\mathbf{A}^s$) on the 3-sphere $S^3$ in $\mathbb{R}^4$ respectively. Also unit quaternions, allowing to represent all rotations $\mathcal{R}$ without singularities, live on the unit sphere $S^3$. Lines in 3D, represented by Plücker coordinates ($\mathbf{L}^s$), live on the so-called Klein quadric $Q$ which is the subspace of the unit sphere $S^5$ in $\mathbb{R}^6$ restricted by the Plücker constraint. Planar homographies, represented by $3\times3$-matrices may be normalized either enforcing their Frobenius norm $\|\mathsf{H}\|^2 = \sum_{ij} H_{ij}^2$ or their determinant $|\mathsf{H}|$ to be 1, then their vector $\mathbf{h} = \mathrm{vec}(\mathsf{H})$ also live on a unit sphere, namely $S^8$ or in another non-linear space respectively. Other transformations, such as the singular correlation matrix $\mathsf{E}$ of the relative orientation, called essential matrix, can be represented by products of basic transformations, e. g. $\mathsf{E} = \mathsf{S}(\mathbf{b})\,\mathsf{R}^\mathsf{T}$, as a function of a homogeneous 3-vector, the base direction $\mathbf{b}$, and the rotation $\mathsf{R}$.

How to represent uncertain elements on these curved manifolds is the topic of the next section.

### 2.1 *Representation of Uncertainty in non-linear Subspaces*

The principle of testing and estimation of entities can be easily visualized, if they live on a one-dimensional manifold in 2D, see Fig. 1. In both cases we assume the uncertainty of the entities is small enough compared to the curvature of the manifold to represent it as a covariance matrix in the tangent space. The approximate value may deviate from the mean value. In order to realize this procedure we need to define (1) the tangent space, which is used for representing the uncertainty and performing the testing and the estimation, and (2) the forward and (3) the backward projection taking the uncertainty of the entities into account. The definition of the tangent space in all cases is realized by linearisation, while the forward and backward projections differ for transformations and geometric entities: While the manifolds of transformations are represented explicitly, the manifolds for geometric entities are represented implicitly, requiring special care.

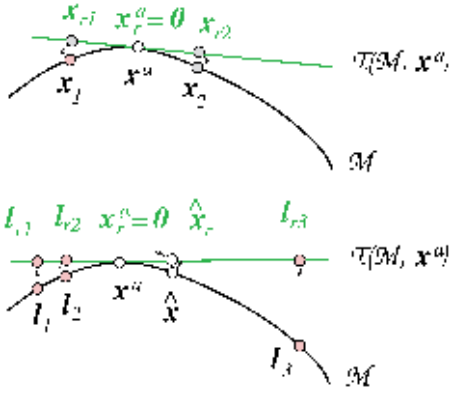



**Fig. 1:** Above: Testing the identity of two points $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$: the points are first projected into the tangent space $\mathcal{T}(\mathcal{M},\boldsymbol{x}^a)$ of the manifold $\mathcal{M}$ at some approximate value $\boldsymbol{x}^a$, leading to a substitute test of $\boldsymbol{x}_{r1}$ and $\boldsymbol{x}_{r2}$. Below: Estimating the mean of three observations $\boldsymbol{l}_i$, $i = 1,2,3$ is performed in the tangent space $\mathcal{T}(\mathcal{M},\boldsymbol{x}^a)$ at the approximate value $\boldsymbol{x}^a$ after projecting them to $\boldsymbol{l}_{ri}$ leading to an updated estimate $\hat{\boldsymbol{x}}_r$, and, after back projecting to the manifold, to the estimate $\hat{\boldsymbol{x}}$.

## 2.2 *Minimal Representation for Uncertain Transformations*

We explain the principle of representing uncertain transformations using rotations in 3D.

### Uncertain rotations

Take the space $S^3$ of rotations, represented by rotation matrices $R \in \mathbb{R}^{3\times3}$. Starting from the mean rotation $\mathbb{E}(R)$ a neighbouring uncertain rotation $\underline{R}$ can be represented by

$$\underline{R} = R(\underline{\boldsymbol{r}})\ \mathbb{E}(R)\,, \tag{1}$$

where $R(\underline{\boldsymbol{r}})$ represents a small rotation, close to the unit matrix depending on the stochastic 3-vector $\underline{\boldsymbol{r}}$. Its $3\times3$-covariance matrix $\Sigma_{rr}$ minimally and uniquely represents the uncertainty of the rotation matrix $\underline{R}$. Obviously, the vector $\boldsymbol{r}$ lies in the three-dimensional tangent space $\mathbb{R}^3$ of the rotations, evaluated at the zero-rotation $\mathcal{R}(\mathbf{0})$, represented by the unit matrix $I_3$. The function

$$\begin{aligned} R(\boldsymbol{r}) &= \exp\big(S(\boldsymbol{r})\big) \\ &= I_3 + S(\boldsymbol{r}) + \frac{1}{2!}S^2(\boldsymbol{r}) + \dots \\ &= I_3 + \frac{\sin(|\boldsymbol{r}|)}{|\boldsymbol{r}|}S(\boldsymbol{r}) + \frac{1-\cos(|\boldsymbol{r}|)}{|\boldsymbol{r}|^2}S^2(\boldsymbol{r}) \end{aligned} \tag{2}$$

maps the three-dimensional linear tangent space $\mathbb{R}^3$ to the spherical space $S^3$ of rotations. The transition from the second to the third line of (2), uses the easy to be proved fact: $S^3(\boldsymbol{r}) = -|\boldsymbol{r}|^2\, S(\boldsymbol{r})$ for collecting higher order terms of $S(\boldsymbol{r})$, see Hartley & Zisserman (2000, A4.9). This function is called the exponential map. In general, the exponential of a skew symmetric matrix is a rotation matrix. The exponential map can only represent rotations with angles $\neq \pm 180°$, which is no restriction in our context.

The mapping from a rotation $R$ to the tangent space at an approximate rotation can easily be achieved from the linearised version of (1) using some approximate rotation $R^a$ being some estimate for the mean rotation, thus

$$R \approx (I_3 + S(\boldsymbol{r}))\,R^a \quad \text{or} \quad S(\boldsymbol{r}) \approx RR^{a\mathsf{T}} - I_3 \tag{3}$$

yielding $\boldsymbol{r} = [S_{23}; S_{31}; S_{12}]$. As the elements of $\boldsymbol{r}$ occur linearly in (3), left, it also is used for linearisation within an estimation procedure.

The non-linear manifold, here of rotations, obviously can be *explicitly* represented, here by using the exponential map (2), which in the context of estimation using the Gauß-Markov model, is just a special case of a non-linear function of the unknown parameters.

### Uncertain homographies

This principle of representing an uncertain transfomation can be generalized to uncertain 2D or 3D homographies and their specialisations, namely 2D and 3D motions and similarities. An uncertain 2D homography, for example, can be represented as a left-product of the mean homography $\mathbb{E}(\mathsf{H})$ and a small uncertain homography $\mathsf{H}(\underline{\boldsymbol{k}})$

$$\underline{\mathsf{H}} = \mathsf{H}(\underline{\boldsymbol{k}})\ \mathbb{E}(\mathsf{H})\,. \tag{4}$$

The small homography is close to the unity, as can be seen, when analysing the linearisation of the transformation of the Cartesian coordinates. It depends on a stochastic 8-vector $\underline{\boldsymbol{k}}$. We assume the matrices to have determinant one, thus being spectrally normalized $|\mathsf{H}| = \prod_i \lambda_i = 1$, using the eigenvalues $\lambda_i$ of $\mathsf{H}$. Homographies close to the unit matrix having determinant one can again be represented by the exponential map

$$\mathsf{H}(\boldsymbol{k}) = \exp\big(K(\boldsymbol{k})\big) = I_3 + K(\boldsymbol{k}) + \frac{1}{2!}K^2(\boldsymbol{k}) + \dots \tag{5}$$

with the zero-trace matrix

$$K(\boldsymbol{k}) = \begin{bmatrix} k_1 & k_4 & k_7 \\ k_2 & k_5 & k_8 \\ k_3 & k_6 & -k_1-k_5 \end{bmatrix}. \tag{6}$$

Again the $8\times8$-covariance matrix $\Sigma_{kk}$ minimally and uniquely, up to the convention of the trace-less matrix $K$ (Begelfor & Werman 2005), represents the uncertainty of the spectrally normalized homography $\underline{\mathsf{H}}$. Linearisation of (4) at a given approximate homography $\mathsf{H}^a$ therefore leads to

$$\mathsf{H} \approx (I_3 + K(\boldsymbol{k}))\,\mathsf{H}^a, \tag{7}$$

which can be used within an iterative estimation procedure and allows to determine $\boldsymbol{k}$ for a homography $\mathsf{H}$ close to an approximate one $\mathsf{H}^a$, taking the first eight values of $\mathrm{vec}\,(\mathsf{H}\,(\mathsf{H}^a)^{-1} - I_3)$.

Finally, it might be useful for some applications to represent both, the homography and its inverse, linearly in 8 parameters, which can easily be achieved using $\underline{\mathsf{H}}^{-1} = (\mathbb{E}(\mathsf{H})^{-1}\,\mathsf{H}\,(-\underline{\boldsymbol{k}})) \approx (\mathbb{E}(\mathsf{H}))^{-1}\,(I_3 - K(\underline{\boldsymbol{k}}))$ together with (4).

## 2.3 *Minimal Representation for Uncertain Points in 2D and 3D*

We will now transfer the concept to uncertain unit vectors on the unit sphere $S^2$, representing 2D points and lines, and generalize it to other geometric entities.

Let an uncertain 2D point $\mathcal{X}$ be represented with its mean, the 2-vector $\boldsymbol{\mu}_x$ and its $2 \times 2$-covariance matrix $\Sigma_{xx}$. It can be visualized by the standard ellipse $(\boldsymbol{x} - \boldsymbol{\mu}_x)^\mathsf{T} \Sigma_{xx}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}_x) = 1$ . Spherically normalizing the homogeneous vector $\mathbf{x} = [\boldsymbol{x}; 1] = [u, v, w]^\mathsf{T}$ yields

$$\mathbf{x}^s = \frac{\mathbf{x}}{|\mathbf{x}|}, \quad \Sigma_{\mathrm{x}^s\,\mathrm{x}^s} = J \Sigma_{\mathrm{x}\,\mathrm{x}} J^\mathsf{T} \tag{8}$$

with the $3 \times 3$-matrix

$$\Sigma_{\mathrm{x}\,\mathrm{x}} = \begin{bmatrix} \Sigma_{xx} & \mathbf{0} \\ \mathbf{0}^\mathsf{T} & 0 \end{bmatrix} \tag{9}$$

using the Jacobian

$$J = \frac{\partial \mathrm{x}^s}{\partial \mathrm{x}} = \frac{1}{|\mathbf{x}|}(I_3 - \mathbf{x}^s \mathbf{x}^{s\mathsf{T}}) \tag{10}$$

with $\mathrm{rank}\,(\Sigma_{\mathrm{xx}}) = 2$ and $\mathrm{null}\,(\Sigma_{\mathrm{xx}}) = \mathbf{\mu}_x^s$. As the smallest eigenvalue is zero, the standard ellipsoid is flat and lies in the tangent space of $\mathbf{x}$ at $S^2$.

In the following we assume all point vectors $\mathbf{x}$ to be spherically normalized and omit the superscript $^s$ for simplicity of notation.

We now want to choose a coordinate system $[\boldsymbol{s}, \boldsymbol{t}]$ in the tangent space $\perp \mathbf{\mu}_x$, and represent the uncertainty by a $2 \times 2$-matrix in that coordinate system (Fig. 2).

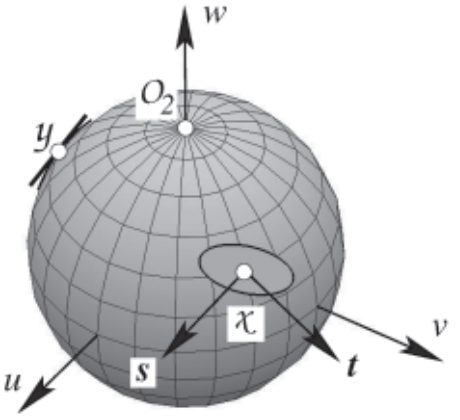



**Fig. 2:** Minimal representation for an uncertain point $\mathcal{X}(\underline{\mathbf{x}})$ on the unit sphere $S^2$ representing the projective plane $\mathbb{P}^2$ by a flat ellipsoid in the tangent plane at the mean $\mathbf{\mu}_x$. The uncertainty has only two degrees of freedom in the tangent space spanned by two basis vectors $\boldsymbol{s}$ and $\boldsymbol{t}$ of the tangent space, being the null space of $\mathbf{\mu}_x^\mathsf{T}$. The uncertainty should not be too large, such that the deviation of the distribution on the sphere and on the tangent plane do not differ too much, as at point $y$.

This is easily achieved by using the orthonormal matrix collecting the two base vectors $\boldsymbol{s}$ and $\boldsymbol{t}$.

$$J_r(\mathbf{\mu}_x) = \mathrm{null}(\mathbf{\mu}_x^\mathsf{T}) = [\boldsymbol{s}, \boldsymbol{t}], \quad [\mathbf{\mu}_x^\mathsf{T}\boldsymbol{s}, \mathbf{\mu}_x^\mathsf{T}\boldsymbol{t}] = \mathbf{0}^\mathsf{T}, \tag{11}$$

fulfilling $J_r^\mathsf{T}(\mathbf{\mu}_x)\, J_r(\mathbf{\mu}_x) = I_2$. The subscript $r$ in $J_r$ stands for the reduced (tangent) space.

The normal of the tangent space is $\mathbf{\mu}_x$. This results from linearising the constraint $h = \underline{\mathbf{x}}^\mathsf{T}\underline{\mathbf{x}} - 1 = 0$ w.r.t. $\underline{\mathbf{x}}$ leading to the linear constraint $\mathbf{\mu}_x^\mathsf{T}\,\underline{\Delta\mathbf{x}} = 0$, expressing the fact, that any deviation of a vector from the mean $\mathbf{\mu}_x$ is perpendicular to $\mathbf{\mu}_x$.

We now represent the stochastic 2-vector $\underline{\boldsymbol{x}}_r$ with mean $\boldsymbol{\mu}_{x_r} = 0$ and covariance $\Sigma_{x_r x_r}$ in the tangent space at $\mathbf{\mu}_x$. In order to arrive at a spherically normalized random vector $\underline{\mathbf{x}}$ with mean $\mathbf{\mu}_\mathrm{x}$ we need to spherically normalize the vector $\underline{\mathbf{x}}^\tau = \mathbf{\mu}_\mathrm{x} + J_r(\mathbf{\mu}_\mathrm{x})\,\boldsymbol{x}_r = \mathbf{\mu}_\mathrm{x} + \underline{x}_{r,1}\,\boldsymbol{s} + \underline{x}_{r,2}\,\boldsymbol{t}$ in the tangent space and obtain

$$\underline{\mathbf{x}}(\mathbf{\mu}_\mathrm{x}, \underline{\boldsymbol{x}}_r) = \mathsf{N}\big(\mathbf{\mu}_\mathrm{x} + J_r(\mathbf{\mu}_\mathrm{x})\,\underline{\boldsymbol{x}}_r\big) \tag{12}$$

$$J_r(\boldsymbol{\mu}_x) = \left.\frac{\partial \mathbf{x}}{\partial \boldsymbol{x}_r}\right|_{x=\mu_x} . \tag{13}$$

We thus can identify $J_r(\boldsymbol{\mu}_x)$ with the Jacobian of this transformation evaluated at $\mathbf{x} = \boldsymbol{\mu}_x$. The inverse transformation is the reduction of the homogeneous vector to its reduced counterpart

$$\underline{\boldsymbol{x}}_r^\tau = J_r^\mathsf{T}(\boldsymbol{\mu}_x)\underline{\mathbf{x}}. \tag{14}$$

omitting the superscript $\tau$ in case no confusion is to be expected. Since $J_r^\mathsf{T}(\boldsymbol{\mu}_x)\,\boldsymbol{\mu}_x = \mathbf{0}$ the mean of $\underline{\boldsymbol{x}}_r$ is the zero vector, $\boldsymbol{\mu}_{x_r} = \mathbf{0}$, as specified.

This allows to establish the one-to-one correspondence between the *reduced covariance matrix* $\Sigma_{x_r x_r}$ of the reduced coordinates $\underline{\boldsymbol{x}}_r$ and the covariance matrix $\Sigma_{xx}$ of $\underline{\mathbf{x}}$:

$$\boldsymbol{\Sigma}_{xx} = J_r(\boldsymbol{\mu}_x)\boldsymbol{\Sigma}_{x_r x_r} J_r^\mathsf{T}(\boldsymbol{\mu}_x) \tag{15}$$

$$\boldsymbol{\Sigma}_{x_r x_r} = J_r^\mathsf{T}(\boldsymbol{\mu}_x)\boldsymbol{\Sigma}_{xx} J_r(\boldsymbol{\mu}_x). \tag{16}$$

We use (14) to derive *reduced* observations and parameters and after estimating corrections $\widehat{\Delta \boldsymbol{x}}_r$ then apply (12) to find corrected estimates $\hat{\mathbf{x}} = \hat{\mathbf{x}}\,(\mathbf{x}^a, \widehat{\Delta \boldsymbol{x}}_r)$.

The non-linear manifold, here of the 3-vectors of homogeneous coordinates, is *implicitly* represented, here by the unit norm constraint, just as in estimation models using constraints onto the observations or the unknown parameters.

A similar reasoning leads to the representation of 3D points. Again, the Jacobian, $J_r(\boldsymbol{\mu}_X)$ overloading the function $J_r(\cdot)$, is the null space of $\mathbf{X}^\mathsf{T}$ and spans the three-dimensional tangent space of $S^3$ at $\boldsymbol{\mu}_X$. The relations between the singular $4 \times 4$-covariance matrix of the spherically normalized vector $\mathbf{X}$ and the reduced $3 \times 3$-covariance matrix $\Sigma_{X_r X_r}$ are equivalent to (15) and (16).

Homogeneous 3-vectors $\mathbf{l}$ representing 2D lines and homogeneous 4-vectors $\mathbf{A}$ representing planes can be handled in the same way.

## 2.4 *Minimal Representation for Straight 3D Lines*

We now generalize the concept for 3D lines. Lines $\mathcal{L}$ in 3D are represented by their normalized Plücker coordinates $\mathbf{L} = [\boldsymbol{L}_h; \boldsymbol{L}_0] = \mathsf{N}([\boldsymbol{Y}-\boldsymbol{X}, \boldsymbol{X}\times\boldsymbol{Y}])$ in case they are derived by joining two points $\mathcal{X}(\boldsymbol{X})$ and $\mathcal{Y}(\boldsymbol{Y})$. Line vectors need to fulfill the quadratic Plücker constraint $\boldsymbol{L}_h^\mathsf{T}\boldsymbol{L}_0 = 0$ and span the Klein quadric $Q$ consisting of all homogeneous 6-vectors fulfilling the Plücker constraint. The dual line $\overline{\mathcal{L}}(\overline{\mathbf{L}})$ has Plücker coordinates $\overline{\mathbf{L}} = [\boldsymbol{L}_0; \boldsymbol{L}_h]$, exchanging its first and second 3-vector. As in addition a 6-vector needs to fulfill the Plücker constraint in order to represent a 3D line, the space of 3D lines is four-dimensional.

The transfer of the minimal representation of points to 3D lines requires some care. The four-dimensional tangent space is perpendicular to $\mathbf{L}$, as $\mathbf{L}^\mathsf{T}\mathbf{L} - 1 = 0$ holds and perpendicular to $\overline{\mathbf{L}}$, as $\overline{\mathbf{L}}^\mathsf{T}\mathbf{L} = 0$ holds. Therefore, the tangent space is given by the four columns of the $6 \times 4$ matrix

$$J_r(\mathbf{L}) = \mathrm{null}\left(\left[\mathbf{L}, \overline{\mathbf{L}}\right]^\mathsf{T}\right) \tag{17}$$

again assuming this matrix to be orthonormal. Therefore for random perturbations $\underline{\boldsymbol{L}}_r$ we have the general 6-vector

$$\underline{\mathbf{L}}^\tau(\boldsymbol{\mu}_\mathrm{L}, \underline{\boldsymbol{L}}_r) = \boldsymbol{\mu}_\mathrm{L} + J_r(\boldsymbol{\mu}_\mathrm{L})\underline{\boldsymbol{L}}_r \tag{18}$$

in the tangent space. In order to arrive at a random 6-vector, which is both spherically normalized and fulfills the Plücker constraint also for finite random perturbations we need to normalize $\mathbf{L}^\tau = [\boldsymbol{L}_h^\tau; \boldsymbol{L}_0^\tau]$ accordingly. The two 3-vectors $\boldsymbol{L}_h^\tau$ and $\boldsymbol{L}_0^\tau$ in general are not orthogonal. Following the idea of Bartoli & Sturm (2005) we therefore rotate these vectors in their common plane such that they become orthogonal. We use a simplified modification, as the normalization within an iteration sequence will have decreasing effect. We use linear interpolation of the directions $\boldsymbol{D}_h^\tau = \mathsf{N}(\boldsymbol{L}_h^\tau)$ and $\boldsymbol{D}_0^\tau = \mathsf{N}(\boldsymbol{L}_0^\tau)$ (Fig. 3). With the distance $d = |\boldsymbol{D}_h^\tau - \boldsymbol{D}_0^\tau|$ and the shortest distance $r = \sqrt{1 - d^2/4}$ of the origin to the line joining $\boldsymbol{D}_h^\tau$ and $\boldsymbol{D}_0^\tau$ we have

**Fig. 3:** Enforcing the Plücker constraint onto a 6-vector $\mathbf{L}^\tau = [\boldsymbol{L}_h^\tau; \boldsymbol{L}_0^\tau]$. Starting from vectors $\boldsymbol{D}_h^\tau = \mathsf{N}(\boldsymbol{L}_h^\tau)$ and $\boldsymbol{D}_0^\tau = \mathsf{N}(\boldsymbol{L}_0^\tau)$, which in general are not perpendicular, we easily can enforce the perpendicularity by symmetric linear interpolation, leading to $\boldsymbol{D}_h$ and $\boldsymbol{D}_0$, which are perpendicular. The vector $\mathbf{M} = \mathsf{N}([|\boldsymbol{L}_h|\boldsymbol{D}_h; |\boldsymbol{L}_0|\boldsymbol{D}_0])$ yields a valid Plücker vector close to $\mathbf{L}^\tau$.

$$\boldsymbol{D}_{h,0} = (1/2 \pm r/d)\boldsymbol{D}_h^\tau + (1/2 \mp r/d)\boldsymbol{D}_0^\tau. \tag{19}$$

The 6-vector $\mathbf{M} = [|\boldsymbol{L}_h|\boldsymbol{D}_h; |\boldsymbol{L}_0|\boldsymbol{D}_0]$ now fulfills the Plücker constraint and preserves the length ratio of its two parts, as the vectors $\boldsymbol{D}_h$ and $\boldsymbol{D}_0$ have the same length by construction. The vector $\mathbf{M}$ finally needs to be spherically normalized. This leads to the normalized stochastic 3D line coordinates

$$\underline{\mathbf{L}} = \mathsf{N}(\underline{\mathbf{L}}^\tau(\boldsymbol{\mu}_\mathrm{L}, \underline{\boldsymbol{L}}_r)) \doteq \underline{\mathbf{M}}/|\underline{\mathbf{M}}| \tag{20}$$

which guarantees $\underline{\mathbf{L}}$ to sit on the Klein quadric, thus to fulfill the Plücker constraint.

The inverse relation to (20) is

$$\underline{\boldsymbol{L}}_r = \boldsymbol{J}_r^\mathsf{T}(\boldsymbol{\mu}_\mathrm{L})\,\underline{\mathbf{L}} \tag{21}$$

as $\boldsymbol{J}_r(\boldsymbol{\mu}_\mathrm{L})$ is an orthonormal matrix. The relations between the covariances of $\mathbf{L}$ and $\boldsymbol{L}_r$ therefore are

$$\boldsymbol{\Sigma}_{\mathrm{LL}} = \boldsymbol{J}_r(\boldsymbol{\mu}_\mathrm{L})\boldsymbol{\Sigma}_{\boldsymbol{L}_r\boldsymbol{L}_r}\boldsymbol{J}_r^\mathsf{T}(\boldsymbol{\mu}_\mathrm{L}) \tag{22}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{L}_r\boldsymbol{L}_r} = \boldsymbol{J}_r^\mathsf{T}(\boldsymbol{\mu}_\mathrm{L})\boldsymbol{\Sigma}_{\mathrm{LL}}\boldsymbol{J}_r(\boldsymbol{\mu}_\mathrm{L}). \tag{23}$$

# 3 Estimation and Testing with minimal Representations

## 3.1 *Estimation*

Using the minimal representations introduced in the last section, we are able to perform ML estimation for all entities. We restrict the following discussion to the model containing constraints between observed and unknown entities only, known as Gauß-Helmert model. Generalizations to include constraints between the parameters only are possible. We start the derivation with the model where the observations have regular covariance matrices and then reduce the model, such that also observations with singular covariance matrices can be handled.

### The optimization problem

We want to solve the following optimization problem

$$\text{minimize} \quad \Omega(\tilde{\boldsymbol{v}}) = \tilde{\boldsymbol{v}}^\mathsf{T}\boldsymbol{\Sigma}_{ll}^{-1}\,\tilde{\boldsymbol{v}} \tag{24}$$

$$\text{subject to} \quad \boldsymbol{g}(\boldsymbol{l} + \tilde{\boldsymbol{v}}, \tilde{\boldsymbol{x}}) = \mathbf{0} \tag{25}$$

where the $N$ observations $\boldsymbol{l}$, their $N \times N$ covariance matrix $\Sigma_{ll}$ and the $G$ constraint functions $\boldsymbol{g}$ are given, and the $N$ true corrections $\tilde{\boldsymbol{v}}$ to the observations and the $U$ true parameters $\tilde{\boldsymbol{x}}$ are unknown. The number $G$ of constraints needs to be larger than the number of parameters $U$. Also it is assumed that the constraints are functionally independent. The solution yields the ML estimates, namely the fitted observations $\hat{\boldsymbol{l}} = \boldsymbol{l} + \hat{\boldsymbol{v}}$ via $\hat{\boldsymbol{v}}$ and parameters $\hat{\boldsymbol{x}}$, under the assumption that the observations are normally distributed with covariance matrix $\Sigma_{ll} = D(\underline{\boldsymbol{l}}) = D(\underline{\boldsymbol{v}})$, and the true observations $\tilde{\boldsymbol{l}}$ fulfill the constraints given the true parameters $\tilde{\boldsymbol{x}}$.

### Example: Bundle adjustment

Bundle adjustment is based on the projection relation $\mathbb{E}(\mathbf{x}'_{ij}) = \lambda_{ij}\mathsf{P}_j\mathbf{X}_i$ between the scene points $\mathcal{X}_i$, the projection matrices $\mathsf{P}_j$ and the image points $\mathcal{X}_{ij}$ of point $\mathcal{X}_i$ observed in camera $j$. The classical approach eliminates the indi-

vidual scale factors $\lambda_{ij}$ by using Euclidean coordinates for the image points. Also the scene points are represented by Euclidean coordinates. This does not allow for scene or image points at infinity. This may occur when using omnidirectional cameras, where a representation of the image points in a projection plane is not possible for all points or in case scene points are very far away compared to the length of the motion path of a camera, e.g. at the horizon. There are two ways to eliminate the scale factor while being able to handle points at infinity.

1. The easiest way is to rewrite the model using spherical normalisation: $\mathbb{E}(\mathbf{x}'^{s}_{ij}) = \mathsf{N}(\mathsf{P}_j \mathbf{X}_i)$ and then multiplying each equation with $J_x(\mathbf{x}'^{a\mathsf{T}}_{ij})$ leading to $\boldsymbol{x}'_{r,ij} = J_x(\mathbf{x}'^{a\mathsf{T}}_{ij})\,\mathsf{N}(\mathsf{P}_j \mathbf{X}_i)$. The reduced coordinates $\boldsymbol{x}'_{r,ij}$ of the camera rays $\mathbf{x}'^{s}_{r,ij}$ have a regular covariance matrix, allowing to use the classical Gauß-Markov model for estimation.
2. The scale factor also can be eliminated by expressing the collinearity as $\mathbb{E}(\mathbf{x}'_{ij}) \times \mathsf{P}_j \mathbf{X}_i = \mathbf{0}$, thus requiring the two homogeneous coordinate vectors $\mathbf{x}'_{ij}$ and $\mathsf{P}_j \mathbf{X}_i$ to be parallel. These constraints between observations and unknown parameters thus require to use the Gauß-Helmert model from (24) (Schneider et al. 2011). However, we also can handle constraints of image lines $\ell'_{ij}$ passing though the projected point $\mathsf{P}_j \mathbf{X}_i$, reading as $\mathbb{E}(\mathbf{l}'^{\mathsf{T}}_{ij})\,\mathsf{P}_j \mathbf{X}_i = 0$. Thus both types of constraints can be represented using (25).

The singularity of the covariance matrix of the spherically normalized image points and the necessity to represent the scene points also with spherically normalized homogeneous vectors, requires to use the corresponding reduced coordinates.

For solving the generally non-linear problem, we assume approximate values $\hat{\boldsymbol{x}}^a$ and $\hat{\boldsymbol{l}}^a$ for the fitted parameters and observations to be available. We thus search for corrections $\widehat{\Delta \boldsymbol{l}}$ and $\widehat{\Delta \boldsymbol{x}}$ for the fitted observations and parameters using $\hat{\boldsymbol{l}} = \boldsymbol{l} + \hat{\boldsymbol{v}} = \hat{\boldsymbol{l}}^a + \widehat{\Delta \boldsymbol{l}}$ and $\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}^a + \widehat{\Delta \boldsymbol{x}}$. With these assumptions we can rephrase the optimization problem: minimize $\Omega(\widehat{\Delta \boldsymbol{l}}) = (\hat{\boldsymbol{l}}^a - \boldsymbol{l} + \widehat{\Delta \boldsymbol{l}})^{\mathsf{T}}\, \Sigma_{ll}^{-1}\, (\hat{\boldsymbol{l}}^a - \boldsymbol{l} + \widehat{\Delta \boldsymbol{l}})$ subject to $\boldsymbol{g}(\hat{\boldsymbol{l}}^a + \widehat{\Delta \boldsymbol{l}}, \hat{\boldsymbol{x}}^a + \widehat{\Delta \boldsymbol{x}}) = \mathbf{0}$. The approximate values are iteratively improved by finding best estimates for $\widehat{\Delta \boldsymbol{l}}$ and $\widehat{\Delta \boldsymbol{x}}$ using the linearized constraints

$$\boldsymbol{g}(\hat{\boldsymbol{l}}^a, \hat{\boldsymbol{x}}^a) + A \widehat{\Delta \boldsymbol{x}} + B^{\mathsf{T}} \widehat{\Delta \boldsymbol{l}} = \mathbf{0} \tag{26}$$

with the corresponding Jacobians $A$ and $B$ of $\boldsymbol{g}$ to be evaluated at the approximate values.

## Reducing the model

We now want to transform the model in order to allow for observations with singular covariances. For simplicity we assume the vectors $\boldsymbol{l}$ and $\boldsymbol{x}$ of all observations and unknown parameters can be partitioned into $I$ and $J$ individual and mutually uncorrelated observational vectors $\boldsymbol{l}_i$, $i = 1, \ldots, I$ and parameter vectors $\boldsymbol{x}_j$, $j = 1, \ldots, J$, referring to points, lines, planes, or transformations.

We first introduce the reduced observations $\boldsymbol{l}_{ri}$, the reduced corrections of the observations $\widehat{\Delta \boldsymbol{l}}_{ri}$, and the reduced corrections $\widehat{\Delta \boldsymbol{x}}_{rj}$:

$$\boldsymbol{l}_{ri} = J_r^{\mathsf{T}}(\hat{\boldsymbol{l}}_i^a)\, \boldsymbol{l}_i, \tag{27}$$

$$\widehat{\Delta \boldsymbol{l}}_{ri} = J_r^{\mathsf{T}}(\hat{\boldsymbol{l}}_i^a) \widehat{\Delta \boldsymbol{l}}_i, \tag{28}$$

$$\widehat{\Delta \boldsymbol{x}}_{rj} = J_r^{\mathsf{T}}(\hat{x}_j^a) \widehat{\Delta \boldsymbol{x}}_j, \tag{29}$$

where each Jacobian refers to the type of the entity it is applied to. The reduced approximate values are zero, as they are used to define the reduction, e.g. from (14) we conclude for 2D points $\hat{\mathbf{x}}_r^a = J_r^{\mathsf{T}}(\hat{\mathbf{x}}^a)\, \hat{\mathbf{x}}^a = \mathbf{0}$. We collect the Jacobians in two block diagonal matrices $J_l^{\mathsf{T}} = \{J_r^{\mathsf{T}}(\hat{\boldsymbol{l}}_i^a)\}$ and $J_x^{\mathsf{T}} = \{J_r^{\mathsf{T}}(\hat{\boldsymbol{x}}_j^a)\}$ in order to arrive at the reduced observations $\boldsymbol{l}_r = J_l^{\mathsf{T}} \boldsymbol{l}$, the corrections for the reduced observations $\widehat{\Delta \boldsymbol{l}}_r = J_r^{\mathsf{T}} \widehat{\Delta \boldsymbol{l}}$ and parameters $\widehat{\Delta \boldsymbol{x}}_r = J_r^{\mathsf{T}} \widehat{\Delta \boldsymbol{x}}$.

Second we need to reduce the covariance matrices $\Sigma_{l_i l_i}$. This requires some care: As a covariance matrix is the mean squared deviation from the mean, we need to refer to the best estimate of the mean when using it. In our context the best estimate for the mean at the current iteration is the approximate value $\hat{\boldsymbol{l}}_i^a$. Therefore we need to apply two steps: (1) transfer the given covariance matrix, referring to $\boldsymbol{l}_i$, such that it refers to $\hat{\boldsymbol{l}}_i^a$ and (2) reduce the covariance matrix to the minimal representation $\hat{\boldsymbol{l}}_{ri}$. As an example, let the observations be 2D lines with spherically normalized homogeneous vectors $\mathbf{l}_i$. Then the re-

duction is achieved by: $\Sigma^a_{l_r l_s} = J^a_i \Sigma_{l_i l_i} J^{aT}_i$ with $J^a_i = J_r^T(\mathbf{l}^a_i)\, R(\mathbf{l}_i, \hat{\mathbf{l}}^a_i)$ namely by first applying a minimal rotation from $\mathbf{l}_i$ to $\hat{\mathbf{l}}^a_i$ (McGlone et al. 2004), second reducing the covariance matrix following (15). The superscript $^a$ in $\Sigma^a_{l_r l_s}$ indicates the covariance to depend on the approximate values.

The reduced constraints now read as

$$\boldsymbol{g}(\hat{\boldsymbol{l}}^a, \hat{\boldsymbol{x}}^a) + A_r \widehat{\Delta \boldsymbol{x}}_r + B_r^T \widehat{\Delta \boldsymbol{l}}_r = \mathbf{0} \tag{30}$$

with

$$A_r = A J_x^T \quad B_r^T = B^T J_l^T. \tag{31}$$

Now we need to minimize the weighted sum of the squared reduced corrections $\hat{\boldsymbol{v}}_r = \hat{\boldsymbol{l}}^a_r - \boldsymbol{l}_r + \widehat{\Delta \boldsymbol{l}}_r = -\boldsymbol{l}_r + \widehat{\Delta \boldsymbol{l}}_r$. Thus we need to minimize $\Omega(\widehat{\Delta \boldsymbol{l}}_r) = (-\boldsymbol{l}_r + \widehat{\Delta \boldsymbol{l}}_r)^T (\boldsymbol{\Sigma}^a_{l_r l_r})^{-1} (-\boldsymbol{l}_r + \widehat{\Delta \boldsymbol{l}}_r)$ subject to the reduced constraints in (30).

The estimated corrections $\widehat{\Delta \boldsymbol{x}}_r$ and $\widehat{\Delta \boldsymbol{l}}_r$ to the reduced parameters and observations of the reduced linearised model are obtained from McGlone et al. (2004, Tab. 2.3). They are used to update the approximate values for the parameters and the fitted observations using the corresponding non-linear transformations, e.g. for an observed 2D point one uses (12), for an unknown 3D line (20). Using (15) and (22), for example, one can determine the covariance matrices of the non-reduced homogeneous coordinate vectors.

## 3.2 *Testing the Identity of Two Entities*

Testing geometric relations using homogeneous coordinates in most cases leads to constraints which are linear in the coordinates of each geometric entity. Examples are the incidence of a 2D point $\mathcal{X}$ and a 2D line $\mathcal{l}$, namely the constraint $\mathbf{x}^T\mathbf{l} = 0$, of a 3D line $\mathcal{L}$ and a plane $\mathcal{A}$, namely the constraint $\Gamma(\mathbf{L})\mathbf{A} = \mathbf{0}$ with the Plücker matrix $\Gamma(\mathbf{L}) = \mathbf{X}\mathbf{Y}^T - \mathbf{Y}\mathbf{X}^T$ for a line $\mathcal{L}$ through the points $\mathcal{X}$ and $\mathcal{Y}$, or the identity of two points $\mathcal{X}$ and $\mathcal{y}$, namely the constraint $\mathsf{S}(\mathbf{x})\mathbf{y} = \mathbf{0}$. Among all these tests, the tests on the identity of two homogeneous entities can be simplified using reduced coordinates.

Let the two entities be two lines $\mathcal{L}_i(\mathbf{L}_i, \Sigma_{\mathbf{L}_i\mathbf{L}_i})$, $i = 1, 2$, specified by their, not necessarily normalized, Plücker coordinates and their covariance matrices. Testing their identity would require to test $\boldsymbol{\mu}_{\mathbf{L}_1} = \lambda \boldsymbol{\mu}_{\mathbf{L}_2}$ for some unknown scale factor $\lambda$, which is cumbersome (McGlone et al. 2004). But this can be realized easily by testing the identity of their reduced coordinates as follows. (1) Spherically normalize the lines $\mathbf{L}^s_i = \mathbf{L}_i / |\mathbf{L}_i|$ and derive their $6 \times 6$-covariance matrix $\Sigma_{\mathbf{L}^s_i \mathbf{L}^s_i}$ by variance propagation similar to the normalization of 2D points in (8). The constraint to be tested now reads simply as $\mathbb{E}(\underline{\mathbf{L}}^s_2 - \underline{\mathbf{L}}^s_1) = \mathbf{0}$. (2) Change to reduced coordinates, as the covariance matrix of the difference is singular. This requires the choice of an approximate 3D line $\mathbf{L}^a$, close to the given ones, approximating the mean $\boldsymbol{\mu}_L$ and using $\underline{\boldsymbol{L}}_{ri} = \boldsymbol{J}_r^T(\mathbf{L}^a)\, \underline{\mathbf{L}}_i$, see (21). We may choose one of both lines. The constraint with reduced coordinates now reads as

$$\mathbb{E}(\underline{\boldsymbol{L}}_{r2} - \underline{\boldsymbol{L}}_{r1}) = \mathbf{0}. \tag{32}$$

In case the two lines are mutually statistically independent the optimal test statistic now is

$$\underline{T} = \underline{\boldsymbol{d}}^T \Sigma^{-1}_{dd} \underline{\boldsymbol{d}} \sim \chi^2_4, \tag{33}$$

with the difference and its regular covariance matrix

$$\underline{\boldsymbol{d}} = \underline{\boldsymbol{L}}_{r2} - \underline{\boldsymbol{L}}_{r1}, \tag{34}$$

$$\Sigma_{dd} = \Sigma_{L_{r1}L_{r1}} + \Sigma_{L_{r2}L_{r2}}. \tag{35}$$

The test statistic $\underline{T}$ is $\chi^2_4$-distributed in case the null-hypothesis, namely if the two lines are identical, holds. It obviously is the Mahalanobis distance of the two reduced line coordinates, thus can be used to measure the difference between two 3D lines.

The test on the identity of two entities using their minimal representation, thus their reduced coordinates can be applied to all geometric entities and also to all homogeneous transformations.

## 4 Example: Estimating 3D Line from Image Line Segments

The following example demonstrates the practical use of the proposed method: namely determining 3D lines from image line segments. Fig. 4 shows three images taken with a CANON 450D. The focal length was determined using vanishing points, the principal point was assumed to be the image centre, the images were not corrected for lens distortion. The images then were mutually oriented using a bundle adjustment program. Straight line segments were automatically detected and a small subset of 12, visible in all three and pointing in the three principal directions were manually brought into correspondence. From the straight lines $\ell_{ij}(\mathbf{l}'_{ij})$, $i = 1, \ldots, 12$; $j = 1, 2, 3$ and the projection matrices $\mathsf{P}_j$ we determined the projection planes $\mathbf{A}_{ij} = \mathsf{P}_j^{\mathsf{T}} \mathbf{l}'_{ij}$ of the line segments. For determining the ML estimates of the 12 lines $\mathbf{L}_i$, each from the three corresponding projection planes, we needed their covariance matrices. They were determined by variance propagation based on the covariance matrices of the image lines $\mathbf{l}_{ij}$ and the covariance matrices of the projection matrices.

As we did not have the cross-covariance matrices between any two of the projection matrices, we only used the uncertainty $\Sigma_{Z_jZ_j}$ of the three projection centres $Z_j$. The covariance matrices of the straight line segments were derived from the uncertainty given by the feature extraction. For this we fitted a straight line through the edge pixels, which was assumed to be determinable with a standard deviation of $\sigma_p = 0.3$ pixel]. The covariance matrix of the projection planes then is determined by variance propagation of $\mathbf{A}_{ij} = \mathsf{P}_j^{\mathsf{T}} \mathbf{l}'_{ij} = (I_4 \otimes \mathbf{l}'_{ij})\,\mathrm{vec}\,\mathsf{P}_j$ from $\Sigma_{\mathbf{A}_{ij}\mathbf{A}_{ij}} = \mathsf{P}_i^{\mathsf{T}} \Sigma_{\mathbf{l}'_i\mathbf{l}'_i} \mathsf{P}_i + (I_4 \otimes \mathbf{l}'_{ij})\,\Sigma_{\mathsf{p}_i\mathsf{p}_i}\,(I_4 \otimes \mathbf{l}'^{\mathsf{T}}_{ij})$.

We achieved the following results. First, the square roots $\hat{\sigma}_0$ of the estimated variance factors $\hat{\sigma}_0^2 = \Omega/(G-U)$ range between 0.03 and 3.2. As the degrees of freedom for each 3D line estimation is $G - U = 2I - 4 = 2 \cdot 3 - 4 = 2$, thus in this case is very low, such a spread is to be expected. The mean value for the 12 variance factors is 1.1, which confirms the model to fit to the data.


**Fig. 4:** Three images with 12 corresponding straight line segments used for the reconstruction of the 3D lines, forming three groups [1...4], [5...8], [9...12] for three main directions.

As a second result we analysed the angles between the directions of the 12 lines. As they are clustered into three groups corresponding to the main directions of the building, we should find values close to 0° within a group and values close to 90° between lines of different groups. The results are collected in Tab. 1.

The angles between lines in the same group scatter between 0° and 14.5°, the angles be-

**Tab. 1:** Result of determining 12 lines from the images in Fig. 4. Left column: minimal length $l$ of the three line segments involved. Upper right triangle: angles between the lines. Lower left triangle: values of the test statistic for the deviation from 0° or 90°.

| $l_{min}$ \# | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 173 pixel | – | 2.6° | 2.7° | 3.0° | 88.6° | 89.0° | 88.7° | 76.5° | 86.7° | 87.0° | 86.6° | 85.2° |
| 155 pixel | 0.7 | – | 0.7° | 1.6° | 89.9° | 89.7° | 87.3° | 75.1° | 89.0° | 89.2° | 88.9° | 87.4° |
| 72 pixel | 0.7 | 0. | – | 0.9° | 89.4° | 89.8° | 87.9° | 75.6° | 89.4° | 89.6° | 89.3° | 87.8° |
| 62 pixel | 1.2 | 0.3 | 0.1 | – | 88.5° | 88.9° | 88.7° | 76.4° | 89.8° | 90.0° | 89.7° | 88.2° |
| 232 pixel | 0.6 | 0.0 | 0.1 | 0.2 | – | 0.4° | 2.8° | 15.3° | 89.6° | 89.3° | 89.7° | 89.1° |
| 153 pixel | 0.3 | 0.1 | 0.0 | 0.1 | 0.1 | – | 2.4° | 14.9° | 89.5° | 89.2° | 89.6° | 89.1° |
| 91 pixel | 0.5 | 0.8 | 0.4 | 0.2 | 0.8 | 0.6 | – | 12.5° | 89.0° | 88.7° | 89.1° | 88.6° |
| 113 pixel | 1.0 | 1.1 | 1.1 | 0.9 | 1.1 | 1.1 | 0.8 | – | 87.0° | 86.7° | 87.2° | 87.0° |
| 190 pixel | 1.6 | 0.4 | 0.3 | 0.1 | 0.2 | 0.3 | 1.0 | 1.3 | – | 0.4° | 0.2° | 1.6° |
| 82 pixel | 1.4 | 0.3 | 0.2 | 0.0 | 0.4 | 0.4 | 1.2 | 1.4 | 0.5 | – | 0.5° | 1.8° |
| 103 pixel | 1.6 | 0.5 | 0.4 | 0.2 | 0.2 | 0.2 | 0.9 | 1.2 | 0.3 | 0.6 | – | 1.6° |
| 225 pixel | 2.4 | 1.1 | 1.1 | 1.0 | 0.5 | 0.5 | 1.3 | 1.5 | 3.2 | 3.6 | 4.0 | – |

tween lines of different orientation differ from 90° between 0° and 15°. The standard deviations of the angles scatter between 0.4° and 8.3°, this is why none of the deviations from 0 or 90° are significant.

The statistical analysis obviously makes the visual impression objective.

Further examples on the rigour and superiority of representation within vanishing point estimation and 3D line intersection are given in Förstner (2010 a,b).

## 5 Conclusions and Outlook

The paper proposes a minimal representation of uncertain homogeneous entities, vectors or matrices, useful for testing and estimation. It avoids the redundancy of the homogeneous representations. Therefore no additional constraints are required to enforce the normalization of the entities including the Plücker constraints for 3D lines. In addition we not only obtain a minimal representation for the uncertainty of the geometric elements, but also simple means to determine the Mahalanobis distance between two elements, which may be used for testing or for grouping. The covariance matrices of the minimal representation of observed entities are regular allowing a transparent definition of a maximum likelihood estimation. We demonstrated the rigour of the method with the estimation of 3D lines from projection planes in a multi-view setup.

The convergence properties when using the proposed reduced representation does not change as the solution steps are algebraically equivalent. The main advantage of the proposed concept is the ability to handle elements at or close to infinity and the full range of the transformations without loosing numerical stability and that the representation is minimal, which allows to use the representation for large estimation problems, such as the bundle adjustment. The concept can be extended to higher level algebras, such as the geometric or the conformal algebra (Gebken 2009) where the motivation to use minimal representations is even higher than in our context.

## Acknowledgement

I acknowledge the careful work of the reviewers, which helped to make the paper more readable.

## References

Åstrom, K. , 1998: Using Combinations of Points, Lines and Conics to Estimate Structure and Mo-

tion. – Proc. of Scandinavian Conference on Image Analysis.

Bartoli, A. & Sturm, P., 2004: Non Linear Estimation of the Fundamental Matrix With Minimal Parameters. – IEEE Transactions on Pattern Analysis and Machine Intelligence **26** (4): 426–432.

Bartoli, A. & Sturm, P., 2005: Structure-from-Motion Using Lines: Representation, Triangulation and Bundle Adjustment. – Computer Vision and Image Understanding **100:** 416–441.

Begelfor, E. & Werman, M., 2005: How to Put Probabilities on Homographies. – IEEE Trans. Pattern Anal. Mach. Intell. **27** (10): 1666–1670.

Bregler, C. & Malik, J., 1998: Tracking people with twists and exponential maps. – Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, Los Alamitos, CA, USA, 8–15.

Collins, R., 1993: Model Acquisition Using Stochastic Projective Geometry, PhD thesis. – Department of Computer Science, University of Massachusetts: Also published as UMass Computer Science Technical Report TR95-70.

Criminisi, A. , 2001: Accurate Visual Metrology from Single and Multiple Uncalibrated Images. – Springer-Verlag London Ltd.

Förstner, W., 2010a: Minimal Representations for Uncertainty and Estimation in Projective Spaces. – Proc. of Asian Conference on Computer Vision.

Förstner, W., 2010b: Optimal Vanishing Point Detection and Rotation Estimation of Single Images of a Legolandscene. – Int. Archives of Photogrammetry and Remote Sensing. ISPRS Symposium Comm. III, Paris.

Gebken, C., 2009: Conformal Geometric Algebra in Stochastic Optimization. – PhD thesis, Christian-Albrechts-University of Kiel, Institute of Computer Science.

Hartley, R.I. & Zisserman, A., 2000: Multiple View Geometry in Computer Vision. – Cambridge University Press.

Heuel, S., 2004: Uncertain Projective Geometry: Statistical Reasoning for Polyhedral Object Reconstruction, Volume 3008 of LNCS. – Springer: PhD Thesis.

Kanatani, K., 1996: Statistical Optimization for Geometric Computation: Theory and Practice. – Elsevier Science.

McGlone, J.C., Mikhail, E.M. & Bethel, J.S., 2004: Manual of Photogrammetry. – Am. Soc. of Photogrammetry and Remote Sensing.

Meidow, J., 2011: Efficient video mosaicking by multiple loop closing. – Proc. of Photogrammetric Image Analysis: 1–12.

Meidow, J., Beder, C. & Förstner, W., 2009: Reasoning with Uncertain Points, Straight Lines, and Straight Line Segments in 2D. – International Journal of Photogrammetry and Remote Sensing **64:** 125–139.

Rosenhahn, B., Granert, O. & Sommer, G., 2002: Monocular pose estimation of kinematic chains. – Dorst, L., Doran, C. & Lasenby, J. (eds.). – Applications of Geometric Algebra in Computer Science and Engineering. Proc. AGACSE 2001: 373–375, Cambridge, UK. Birkhäuser Boston.

Schneider, J., Schindler, F. & Förstner, W., 2011: Bündelausgleichung für Multikamerasysteme. – Proceedings of the 31th DGPF Conference.

Sturm, P. & Gargallo, P., 2007: Conic fitting using the geometric distance. – Proceedings of the Asian Conference on Computer Vision **2:** 784–795, Springer, Tokyo, Japan.

Address of the Author:

Prof. Dr.-Ing. Dr. h.c. mult. Wolfgang Förstner, Josef-Schell-Straße 34, 53121 Bonn, Germany, Tel.: +49-228-472223, e-mail: wf@ipb.uni-bonn.

Manuskript eingereicht: Dezember 2011
Angenommen: Februar 2012

PFG 2012 / 3, 0221 – 0236
Stuttgart, June 2012

Article

# Kreismarken in perspektiver Abbildung – im Bild und im Bündelblock

Bernhard P. Wrobel, Wachtberg

**Keywords:** Centres of perspective circular targets, correspondence models for circular targets and key points, bundle adjustment based on pixel bundles

**Summary:** *Circular targets in perspective projection – in an image and in bundle adjustment.* The perspective of a circular target generally is an ellipse, its centre **e′**, however, differs from the projected target centre **m′**. If **e′** is introduced into a bundle block adjustment instead of **m′** it may cause sensible deformations. A strict derivation of **m′** is presented by solving the polarity $\boldsymbol{N} = \mathbf{C}'\mathbf{m}'$, with $\boldsymbol{N}$ the normal of the target and $\mathbf{C}'$ the matrix of the target ellipse in the image. Its solution by single value decomposition (SVD) gives two values for $\boldsymbol{N}$,**m′**, which in the context of a bundle block can be resolved easily. In section 4 the treatment of overdetermined circular targets in a bundle adjustment is discussed. Three correspondence models are presented. The first two rely on the points of the elliptical target which were extracted from its digital image function. The centre of the circular target is determined in 3D, respectively, the target circle as a whole and their statistical uncertainties. The third model solves the correspondence by reconstructing the target in object space directly from its pixel grey values: the target plane and its brightness function, following the concept of facets stereo vision. It applies the original digital image function completely comprising the rays of all pixels around each target (called here a pixel bundle), and it derives the parameters by Gauß-Markov-optimization. In all, it is able to provide the optimal solution. This concept immediately may be transferred to general key points. Thus, the classical bundle block composed of all rays to object points may evolve to a block of pixel bundles, each referring to a key point or a circular target and their close surroundings.

**Zusammenfassung:** Die Perspektive einer Kreismarke $\mathcal{K}$ ist im Allgemeinen eine Ellipse, deren Mittelpunkt **e′** jedoch von der projizierten Kreismitte **m′** verschieden ist. Benutzt man **e′** in einer Bündelblockausgleichung anstelle von **m′**, so können spürbare Verformungen des Bündelnetzes auftreten. Es wird eine strenge Herleitung von **m′** vorgestellt durch Lösung der Polarität $\boldsymbol{N} = \mathbf{C}'\mathbf{m}'$, mit $\boldsymbol{N}$ der $\mathcal{K}$-Normalen und $\mathbf{C}'$ der Matrix der $\mathcal{K}$-Ellipse im Bild. Mit Hilfe einer SVD (single value decomposition) von $\mathbf{C}'$ erhält man zwei Wertepaare für $\boldsymbol{N}$,**m′**, das zutreffende davon ergibt sich leicht in einer Bündelausgleichung. Im Teil 4 wird die Behandlung von überbestimmten $\mathcal{K}$ in einer Bündelausgleichung mit drei Korrespondenzmodellen diskutiert. Die ersten beiden verwenden die Punkte der $\mathcal{K}$-Ellipse, die aus der digitalen Bildfunktion extrahiert worden waren. Es wird die $\mathcal{K}$-Mitte im 3D-Raum bestimmt bzw. die $\mathcal{K}$ als Ganzes einschließlich der statistischen Unsicherheit. Das dritte Modell löst die Korrespondenz durch die Rekonstruktion der $\mathcal{K}$ im 3D Raum direkt aus den originalen Pixelgrauwerten: die $\mathcal{K}$-Ebene und ihre Helligkeitsfunktion nach dem Konzept des Facetten-Stereo-Sehens. Dabei werden die Abbildungsstrahlen der Pixel der $\mathcal{K}$-Umgebung (hier genannt Pixelbündel) vollständig herangezogen und daraus die Modellparameter in einer Gauß-Markov-Optimierung abgeleitet. Das Modell ist in der Lage, die optimale Lösung zu erreichen. Das Konzept kann leicht auf die Verarbeitung von "natürlichen Punkten" (key points) übertragen werden. Daher könnte der klassische Bündelblock, bestehend aus der Gesamtheit von Strahlen zu Objektpunkten, weiter entwickelt werden zu einem Block aus Pixelbündeln, entweder zu einem natürlichen Punkt oder einer künstlichen Marke und deren enger Umgebung.


www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0113
1432-8364/12/0113 $ 4.00

## 1 Einführung

Kreismarken (künftig mit $\mathcal{K}$ bezeichnet) bilden sich in einer Kamera im Allgemeinen als Ellipsen ab (Abb. 1). Der Ellipsenmittelpunkt – leicht aus seinen Randpunkten zu berechnen – weicht jedoch vom projizierten Kreismittelpunkt ab. Gerade dieser wird aber benötigt, weil nur er mit Projektionszentrum und Kreismittelpunkt am Objekt kollinear ist. Dies ist die entscheidende Voraussetzung für die auf Bildpunkten basierende, klassische Bündellösung zur Orientierung von photogrammetrischen Bildern. Will man viele Bilder sehr genau gemeinsam orientieren und hat das erfasste Objekt eine Reihe von Kreismarken, so sind die projizierten Kreismitten anstelle der Ellipsenmitten unverzichtbar zur Vermeidung von systematischen Verformungen des Bündelblocks. Dies allein ist schon Grund genug, sich mit dieser Problematik zu befassen. Darüber hinaus stößt man auf die gleiche Aufgabe bei der Rekonstruktion von Kreisen als Elemente von hergestellten Objekten aus Industrie oder Kultur. Lösungsvorschläge zu diesem Thema gibt es seit Langem, siehe z. B. den strengen von Kager (1981) oder die Abschätzung nach Dold (1997). Dabei zeigt sich, dass mit algebraischer projektiver Geometrie Probleme wie das vorliegende durchsichtig werden und elegant lösbar sind.

Generell sind signalisierte Punkte am Objekt für Bündelblöcke des Nahbereichs wichtig, wenn höchste Ergebnisgenauigkeiten erreicht werden müssen. Die Gestaltung von Punktmarken und ihre Eigenschaften sind daher immer wieder und sehr gründlich untersucht worden (siehe Luhmann 2010 und die dort angegebene Literatur). Die aus Linien aufgebauten Marken, z. B. ein Quadrat, sind a priori frei von der Markenexzentrizität im Bild (Abb. 1): In der perspektiven Abbildung bleiben Geraden und Geradenschnitte erhalten, der Mittelpunkt eines Quadrates kann – mathematisch gleichwertig – vor oder nach der Abbildung aus seinen Randpunkten bestimmt werden. Dennoch scheinen Kreismarken auch heute noch bevorzugt zu werden, vermutlich wegen ihrer stabileren Rekonstruktion aus gestörten Punktmarken. Projizierte Kreismarken sind aber als quadratische Funktion nicht frei von Exzentrizitäten, auch wenn es oft nur um kleine Beträge geht. Selbst diese lassen sich minimieren, wenn man möglichst kleine Marken verwendet. Die Perspektive kommt im Differentiellen, wenn die $\mathcal{K}$ im Bild kleiner ist als die Kamerakonstante, der Affinabbildung recht nahe. Die Exzentrizität kann dann

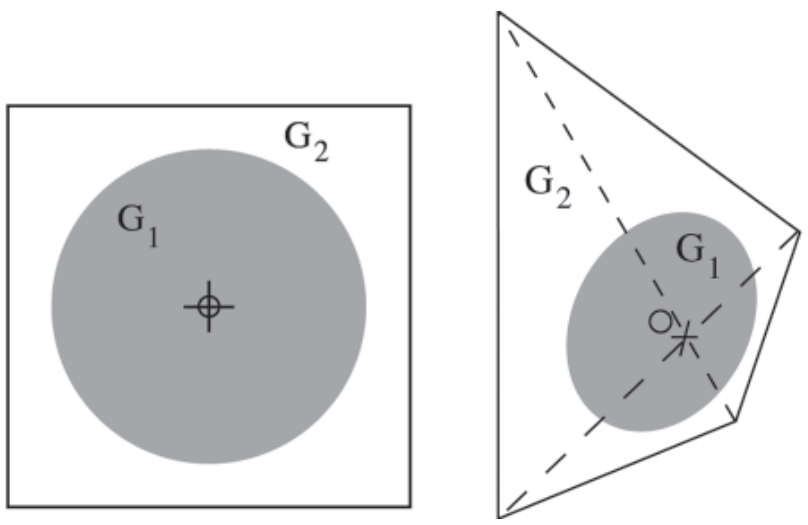



**Abb. 1:** Kreismarke im Objektraum und im Perspektivbild; links: Kreismarke im Objektraum: Quadrat und Kreis (Radius $r$) mit gemeinsamer Mitte $M$ + bzw. o und Objektgrauwerten $G_2$,$G_1$; rechts: Perspektivbild von $\mathcal{K}$; o = Ellipsenmitte $\boldsymbol{e'}$, + = projizierte $\mathcal{K}$-Mitte $\boldsymbol{m'}$ = Mitte von proj. Quadrat; – – – Hilfslinien. Die Bildfunktion der Ellipsenlinie besitzt meist eine Breite von wenigen Pixeln.

tolerierbar sein. Man hat aber das Risiko von Ungenauigkeiten (die Zahl an Pixeln pro Ellipse ist klein), zudem die Unbequemlichkeit, die $\mathcal{K}$ an Aufnahmeentfernungen und Kamera anpassen zu müssen. Mit einer strengen, für jede $\mathcal{K}$-Größe einsetzbaren Methode der $\mathcal{K}$-Zentrumsbestimmung im Bild – wie hier vorgestellt – gewinnt man Flexibilität und Genauigkeit und kann das an sich sehr hohe Potential an relativer Lokalisierungsgenauigkeit durch digitale Bildverarbeitung voll nutzen.

Dem skizzierten Problem wollen wir uns zuerst geometrisch anschaulich nähern und danach einen strengen algebraischen Lösungsweg herleiten. Man lege ein Auge des Betrachters in das Projektionszentrum $\boldsymbol{O}$ (Abb. 2) und schaue über die Ellipse im Bild hinaus in den Objektraum, wo auf einer Trägerebene eine Kreismarke mit frei gewähltem Radius $r$ (z. B. $r = 1$) liegen möge. Durch Drehen und Verschieben der Trägerebene kann man Kreismarke und Bildellipse zur Deckung bringen, also die Kreismarke rekonstruieren, jedoch – im Hinblick auf den beliebigen Kreisradius – nur mit unbekanntem Abstand zum Projektionszentrum. Mit dem Strahlenbündel eines Bildes allein kann man bekanntlich keine metrischen Aussagen erzielen. Das Ergebnis reicht aber aus, weil aus dem Strahlenbündel des Bildes nur der spezielle, zum Kreismittelpunkt führende Strahl mit seinen zwei Komponenten interessiert, die Distanz und damit der Kreisradius aber nicht.

Darüber hinaus ergibt der spielerisch experimentelle Umgang mit der Trägerebene von $\mathcal{K}$, dass es für sie eine zweite Stellung im Raume gibt, die ebenfalls mit der Bildellipse verträglich ist. Wenn die Ellipse zentrisch auf den Bildachsen $x'$ $y'$ liegt, ist die zweite Lösung sogar genau symmetrisch zur ersten (Abb. 2). Bei exzentrischer Ellipse zeigen die zwei Lösungen keine Symmetrie. Im Sonderfall einer $\mathcal{K}$ parallel zur Bildebene fallen beide Lösungen zusammen, das $\mathcal{K}$-Bild stellt einen Kreis dar und sein Mittelpunkt ist bereits die gesuchte Lösung. Die Vielfalt der Lösungen ist aber noch nicht erschöpft. Die vom Projektionszentrum zur Bildellipse verlaufenden Strahlen definieren v o r Betrachter und Bildebene einen Kegel, den es jedoch auch d a h i n t e r geben muss durch Verlängern der Strahlen nach hinten und dort die zentralsymmetrisch gelegene Kreismarke ergibt. Es ist klar, nur Lösungen mit positiven $Z$-

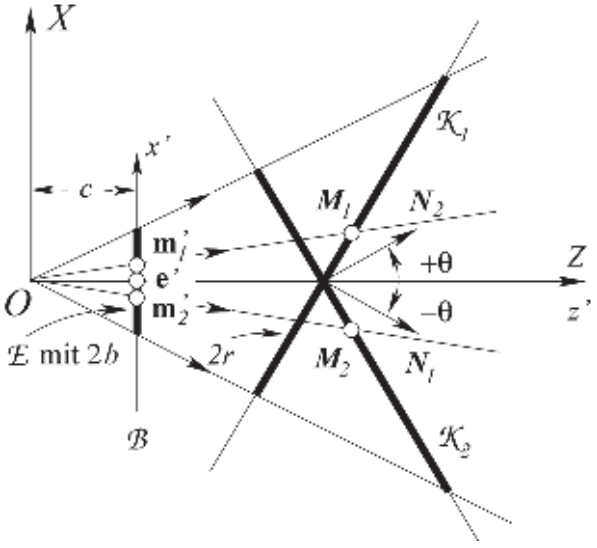



**Abb. 2:** Zwei Rekonstruktionen $\mathcal{K}_1$ und $\mathcal{K}_2$ einer Kreismarke im Strahlenkegel ihrer Ellipse $\mathcal{E}$ bei frei gewähltem $\mathcal{K}$-Radius $r$. $\mathcal{E}$ in der Bildebene $\mathcal{B}$ mit Kamerakonstante $c$ steht hier mit ihrem großen Durchmesser $2a$ senkrecht auf der Zeichenebene, nur der kleine Durchmesser $2b$ auf der $x'$-Achse ist sichtbar, ebenso wie die Koordinatenachsen $X, Z$ und $x', z'$; die Ellipsenmitte $\boldsymbol{e'}$ im Koordinatenursprung von $x'y'z'$; $\boldsymbol{M}_1, \boldsymbol{M}_2$: rekonstruierte $\mathcal{K}$-Mitten vor der Bildebene, ihre Bilder $\boldsymbol{m'}_1, \boldsymbol{m'}_2$ liegen immer auf dem kleinen Ellipsendurchmesser $2b$, hier auf $x'$ und symmetrisch zur $z'$-Achse; $\boldsymbol{N}_1, \boldsymbol{N}_2$ Normalen der $\mathcal{K}$ mit Neigungswinkeln $\theta$.

Koordinaten sind praktisch relevant. Es bleibt bei zwei Lösungen vor der Bildebene. Welche davon die richtige ist, lässt sich mit dem Einzelbild allein nicht entscheiden, wohl aber mit jedem weiteren Bild (mit unabhängigen Orientierungsparametern), in dem die $\mathcal{K}$als Ellipse vorkommt, insbesondere in einem Bündelblock. Halten wir fest, wir benötigen von einer $\mathcal{K}$nur ihre Normale und ihren Mittelpunkt.

Auf der Suche nach einem strengen Verfahren erweist sich die Literatur über die Rekonstruktion von Kreisen als sehr hilfreich (Kager 1981, d'Home et al. 1990, Forsyth et al. 1991, Kanatani 1993, Philip 1997). Mit diesen Grundlagen lassen sich die geometrischen Beziehungen restlos abklären. Hierüber wird mit Ergänzungen und numerischen Beispielen berichtet.

Insgesamt wird der Beitrag folgendermaßen gegliedert. Im Teil 2 werden die projektiven Grundlagen zusammen getragen und die Lösung der Aufgabe aufgezeigt. Danach werden die geometrischen Abhängigkeiten charakterisiert, numerische Beispiele demonstrieren die Notwendigkeit, die hergeleiteten Beziehungen in Bündelblöcken zu nutzen. Danach befassen wir uns mit Korrespondenzmodellen von Punktmerkmalen im Bündelblock, auch solchen, die ausschließlich mit Grauwerten formuliert sind. Wir vergleichen sie nach ihrer statistischen Qualität und ihrer Bedeutung für die künftige Bündelblocklösung.

## 2 Projektive Grundlagen: Der Kreis und sein projizierter Mittelpunkt

Wir haben das Ziel, die projizierte Kreismitte allein aus den Bilddaten der $\mathcal{K}$(Ellipsenmatrix $\mathbf{C}' = [c'_{ij}]$, $i, j = 1\ldots3$, innere Orientierung der Kamera) im Kamerakoordinatensystem zu bestimmen. Das heißt aber, es müssen Beziehungen zwischen den geometrischen Gebilden im Objektraum (Trägerebene des Kreises, Kreis und seine Mitte) und in der Bildebene (Ellipse, projizierte Kreismitte) gefunden werden.

Für die Herleitungen wird die algebraische projektive Geometrie herangezogen. Euklidische Vektoren des Bildraumes werden klein, geneigt und mit Apostroph geschrieben, z. B. $\boldsymbol{x}'$, und die des Objektraumes groß und geneigt, z. B. $\boldsymbol{X}$, dagegen homogene Vektoren und Matrizen aufrecht, z. B. $\mathbf{x}$, $\mathbf{C}$, stets alle fett. Das Zeichen = zwischen homogenen Vektoren bedeutet Proportionalität, sind sie euklidisch normiert (die letzte Komponente ist 1), so gilt die Gleichheit der Vektoren. Es werden Rechtskoordinatensysteme eingeführt (Abb. 2): $XYZ$ für den Objektraum, Ursprung im Projektionszentrum $\boldsymbol{O}$, parallel dazu in der Bildebene die auf den Hauptpunkt bezogenen Koordinaten $x'_{kal}, y'_{kal}$ einer kalibrierten Kamera mit Konstante $c_{kal}$. Wir gehen jedoch über zu euklidisch normierten Bildkoordinaten $\mathbf{x}' = [x'_{kal} / c_{kal}, y'_{kal} / c_{kal}, 1]^T = [x', y', 1]^T$, d.h., $c = 1$, und die $3 \times 3$-Kameramatrix wird einfach: $\mathbf{K} = \mathbf{I}$, vgl. (Förstner & Wrobel 2004, 222, Luhmann 2010, 253). Die positive Z-Achse ist zugleich optische Aufnahmeachse der Kamera.

Zwischen Punkten $\boldsymbol{X} = [X, Y, Z]^T$ des Objektraumes und der Bildebene $\boldsymbol{x}' = [x', y']^T$ bestehe die Perspektivbeziehung

$$x' = X / Z, \quad y' = Y / Z, \tag{1}$$

wobei gelten soll $Z > 1$. Es wird unterstellt, dass ein Datensatz von Ellipsenpunkten aus dem Bild extrahiert und erfolgreich in eine Ellipsengleichung samt Dispersionsmatrix $\mathbf{D}(c_{ij})$, siehe z. B. (Förstner & Wrobel 2004, 89), überführt wurde

$$c'_{11}x'^2 + 2c'_{12}x'y' + 2c'_{13}x' + c'_{22}y'^2 + 2c'_{23}y' + c'_{33} = 0 \tag{2}$$

oder in homogenen Größen

$$\mathbf{x}'^T\mathbf{C}'\mathbf{x}' = 0 \quad \mathbf{x}' = \left[x', y', 1\right]^T$$

$$\mathbf{C}'=\mathbf{C}'^T = \begin{bmatrix} c'_{11} & c'_{12} & c'_{13} \\ c'_{21} & c'_{22} & c'_{23} \\ c'_{31} & c'_{32} & c'_{33} \end{bmatrix}.$$

Die Bildkoordinaten $x', y'$ in (2) ersetzen wir durch Objektkoordinaten aus (1):

$c'_{11}X^2 + 2c'_{12}XY + c'_{22}Y^2 + 2c'_{13}XZ + 2c'_{23}YZ + c'_{33}Z^2 = 0$

Punkte $[X, Y, Z]$ auf dem Kreis in 3D erfüllen diese Gleichung, aber auch Punkte $-X, -Y, -Z$ auf einer zentralsymmetrisch gelegenen Kreismarke hinter der Bildebene. Wir haben

sie aber bereits ausgeschlossen. Im weiteren Vorgehen werden überschlanke oder gar entartetete Ellipsen mit Neigungswinkeln $\theta \rightarrow 90°$, Abb. 2, übergangen. Ihre Extraktion aus der Bildfunktion wäre grenzwertig. Sie wären in einer Bündelausgleichung nicht genau genug. Deshalb haben wir es nur mit regulären Ellipsenmatrizen $\mathbf{C}'$ zu tun.

Entscheidend für den Lösungsweg ist die Erkenntnis, dass zwischen den gesuchten Größen, nämlich der Normalen $\boldsymbol{N}$ der $\mathcal{K}$-Trägerebene und der projizierten $\mathcal{K}$-Mitte $\mathbf{m}'$, eine direkte Beziehung besteht in Form einer Polarität:

$$\boldsymbol{N} = \mathbf{C}'\mathbf{m}'. \tag{3}$$

Dies sei folgendermaßen gezeigt. Es ist bekannt (siehe z. B. Förstner & Wrobel 2004, 775), dass man mit geometrischen Fernelementen des Bildraumes die Orientierung von zugehörigen Elementen im Objektraum bestimmen kann, ein bemerkenswertes Potenzial der algebraischen projektiven Geometrie. Im Anhang 7.1 wird beispielsweise hergeleitet, dass allgemein die Normale $\boldsymbol{N}$ einer Ebene gleich dem Parametervektor $\mathbf{l}'_\infty$ ihrer abgebildeten Ferngeraden ist:

$$\boldsymbol{N} = \mathbf{l}'_\infty. \tag{4}$$

Nun wird man nicht garantieren können, für alle Werte von $\boldsymbol{N}$ die Ferngerade der $\mathcal{K}$-Ebene im Bild direkt anzutreffen. Deshalb muss man $\mathbf{l}'_\infty$ – was für alle sinnvollen Werte von $\boldsymbol{N}$ möglich ist – aus den projektiven Eigenschaften der hier als Marke gewählten Figur herleiten, dem Kreis und dessen Bild, die Ellipse, mit Hilfe von Polaritäten mit Ferngeraden. Auf der $\mathcal{K}$im Objektraum hat man die Polarität zwischen Kreis $\mathbf{C}$ , seinem Mittelpunkt $\mathbf{M}$ als Pol und der Ferngeraden $\mathbf{l}_\infty$ der Trägerebene als Polare (Anhang 7.2). Bei der projektiven Abbildung bleibt die Polarität invariant (Anhang 7.3). Im Bild existiert zwischen den abgebildeten Größen $\mathbf{m}'$, $\mathbf{C}'$ und $\mathbf{l}'_\infty$ wieder eine Polarität, also

$$\mathbf{l}'_\infty = \mathbf{C}'\mathbf{m}', \tag{5}$$

woraus mit (4) Gleichung (3) folgt. Die Polaren als Ferngeraden entziehen sich häufig der unmittelbaren Anschauung, obwohl sie algebraisch existieren. Bei einer allgemeinen Polarität $\mathbf{l}' = \mathbf{C}'\mathbf{x}'$ hat man oft, dass der Pol $\mathbf{x}'$ nicht allzu weit außerhalb der zugehörigen Ellipse $\mathbf{C}'$ liegt und dass die vom Pol ausgehenden Ellipsentangenten in Ellipsenpunkten von der Polare $\mathbf{l}'$ geschnitten werden. Auf diese Weise kann man die allgemeine Polarität auch herleiten. Eine Grenzsituation ergibt sich, wenn sich der Pol der Ellipse nähert und ihn berührt: die zwei Tangenten vereinigen sich mit Pol und Polare. Für Pole im Inneren einer Ellipse gehört eine außerhalb liegende Polare, falls Pol und Ellipsenmittelpunkt identisch sind, befindet sich die Polare im Unendlichen (Anhang 7.2). Das ist die hier gegebene Situation. In der Gruppe der projektiven Abbildungen (auch Kollineationen, Homographien oder projektive Transformationen genannt) gehört $\mathbf{l}' = \mathbf{C}'\mathbf{x}'$ zur Untergruppe projektiver Korrelationen und zwar speziell zu den Polaritäten wegen der Symmetrie von $\mathbf{C}'$. Eine reguläre Polarität transformiert dualisierend Punkte in Geraden oder umgekehrt. Jeder Punkt $\mathbf{x}'$ des 2D Raumes kann mit $\mathbf{C}'$ in eine Polare transformiert werden und umgekehrt. In höher dimensionalen Räumen als hier werden durch Korrelationen Punkte in Hyperebenen transformiert. In der Photogrammetrie ist bekanntlich eine singuläre Korrelation viel im Gebrauch, nämlich in Verbindung mit der Epipolargeometrie zweier Bilder: wenn ein Punkt eines Bildes in eine epi-polare Linie in einem Nachbarbild transformiert wird, um darauf den korrespondierenden Bildpunkt zu suchen.

Nach diesen projektiven Grundlagen befassen wir uns mit der Lösung der Gleichung $\boldsymbol{N} = \mathbf{C}'\mathbf{m}'$. Im Allgemeinen ist eine übliche Gleichungslösung nicht möglich, da beide Vektoren $\boldsymbol{N}$ und $\mathbf{m}'$ unbekannt sind. Im Sonderfall jedoch, dass die $\mathcal{K}$-Ebene von vornherein parallel zur Bildebene liegt, stellt Matrix $\mathbf{C}'$, leicht erkennbar, einen Kreis dar, und es gilt $\boldsymbol{N} = [0\ 0\ 1]^T$. Diesen Zustand kann man auch algebraisch herstellen mit Hilfe erlaubter Umformungen. Man erreicht dies in Schritten mittels Rotationen des Bildkoordinatensystems $x'\ y'\ z'$. Sie sind als konforme Transformationen zulässig, d.h. das Strahlenbündel des Bildes wird dabei nicht deformiert. Schritt 1: Rotation des Koordinatensystems auf die Eigenachsen $\boldsymbol{x''}$, $\boldsymbol{y''}$, $\boldsymbol{z''}$ der Matrix $\mathbf{C}'$. Sie selbst wird diagonal

mit Eigenwerten $\lambda_1 \geq \lambda_2 > \lambda_3$. Im Beispiel der Abb. 2 liegt die Ellipse auf den Koordinatenachsen, ihre große Achse auf $\boldsymbol{y}''$, die kleine auf $\boldsymbol{x}''$. Die $\mathcal{K}$ kann nun um den Winkel $+\theta$ oder $-\theta$ geneigt sein, die Winkel zeigen ihre zwei möglichen Lagen. Die Nicht-Null-Komponenten ihrer Normalen befinden sich in der $\boldsymbol{x}''\boldsymbol{z}''$-Ebene, ebenso wie die der projizierten $\mathcal{K}$-Mitte $\mathbf{m}''$. Schritt 2: Rotation des Koordinatensystems um $\boldsymbol{y}''$ mit $\theta$ unter der Bedingung, dass im neuen System die neue Matrix $\mathbf{C}'''$ identische Diagonalglieder $c'''_{11} = c'''_{22}$ erhält. Der Kegelschnitt $\mathbf{C}'''$ stellt sich nun als verschobener Kreis dar, sein Mittelpunkt $\mathbf{m}'''$ liegt auf der $\boldsymbol{x}'''$-Achse und die $\mathcal{K}$ befindet sich parallel zur Bildebene, d.h. die Normale ist $\boldsymbol{N}''' = [0\ 0\ 1]^T$. Die projizierte Kreismitte $\mathbf{m}'''$ folgt daher aus (3) zu $\mathbf{m}''' = \mathbf{C}'''^{-1}[0\ 0\ 1]^T$. Schritt 3: Die Rücktransformation von $\boldsymbol{N}'''$ und $\mathbf{m}'''$ in das Ausgangssystem beendet die Lösung.

Im Wesentlichen beruhen alle in der Literatur genannten Lösungsverfahren auf den drei Schritten. Im Folgenden werden die Berechnungsformeln nach Philip (1997) genannt, die besonders elegant und kompakt sind. Sie gelten für den Radius $r = 1$ der $\mathcal{K}$. Philip verwendet für $\mathbf{C}'$ anstelle der Eigenwertzerlegung die SVD (singular value decomposition) $\mathbf{C}' = \boldsymbol{U}\mathbf{S}\boldsymbol{V}^T$. $\mathbf{S}$ ist diagonal mit den nicht negativen, rellen Singulärwerten $s_1 \geq s_2 > s_3 > 0$. Diese Reihenfolge der $s_i$ und der Spalten von $\boldsymbol{V}$ wird allen Stellungen der $\mathcal{K}$-Ebene im 3D-Raum gerecht. Mit den $s_i$ ermittelt man den Neigungswinkel $\theta$ der $\mathcal{K}$-Ebene im System der Eigenachsen

$$\cos^2\theta = (s_2 + s_3) / (s_1 + s_3)\,. \qquad (6)$$

Die zwei Werte des **cos** ergeben die zwei Lösungen der Rekonstruktion. Für die $\mathcal{K}$-Normale $\boldsymbol{N}$ und die $\mathcal{K}$-Mitte $\boldsymbol{M}$ erhält man im ursprünglichen System $X, Y, Z$

$$\boldsymbol{N} = \pm\boldsymbol{V}[-\sin\theta\ \ 0\ \ \cos\theta]^T$$
$$\boldsymbol{M} = \pm\boldsymbol{V}\left[-\sin\theta\sqrt{s_3/s_1}\quad 0\quad \cos\theta\sqrt{s_1/s_3}\right]^T. \qquad (7)$$

Das Resultat für $\boldsymbol{M}$ kann mit $r$ multipliziert werden, falls der $\mathcal{K}$-Radius $r \neq 1$ bekannt ist. Dann ist $\boldsymbol{M}$ in 3D vollständig rekonstruiert worden. Mit (1), d.h. mit $c = 1$, folgt daraus die ins Bild projizierte $\mathcal{K}$-Mitte $\mathbf{m}'$. Für die ursprüngliche Kamera erhält man $[m'_1, m'_2]^T = [(X_M / Z_M)c_{kal}, [Y_M / Z_M)c_{kal}]^T$.

Die beiden Lösungen für die projizierte Mitte einer $\mathcal{K}$ kann man also ohne Näherungswerte allein aus ihrer Bildellipse bestimmen ohne irgendeine Approximation.

## 3 Der projizierte Kreismittelpunkt im Bild: Interpretationen und Beispiele

Das Perspektivbild einer $\mathcal{K}$ wird man in allen Größen antreffen, von nur wenige Pixel zählenden bis zum Bildformat füllenden oder das Format schneidenden Größen. In beiden Extremen kann die statistisch zuverlässige Bestimmung von $\mathbf{C}'$ unmöglich sein: es sind zu wenige Randpunkte für eine Ellipse vorhanden, weniger als fünf, oder die Ellipse erscheint im Bild mit einem zu kleinen Ausschnitt. Hier sei mit wenigen Simulationen ein Einblick in die Beträge der Exzentrizitäten gegeben, ohne allzu sehr ins Detail zu gehen. Dies sei dem interessierten Leser überlassen. Primär sind die $\mathcal{K}$ von Bündelblöcken im Visier, also kleine $\mathcal{K}$.

Es seien im Objektraum geeignete $\mathcal{K}$ definiert und in die Bildebene abgebildet worden. Daraus kann die Exzentrizität $\Delta(\mathbf{e}')$ der Ellipsenmitte $\mathbf{e}'$ gegenüber der projizierten $\mathcal{K}$-Mitte $\mathbf{m}'$ direkt berechnet werden oder mit Philip (1997) über

$$\Delta(\mathbf{e}') = \mathbf{m}' - \mathbf{e}' = \mathbf{C}'^{-1}\left[\boldsymbol{N} - [0\quad 0\quad 1]^T\right]. \qquad (8)$$

Für die Simulationen wählen wir $\mathcal{K}$-Ebenen mit Mittelpunkt $\boldsymbol{M} = [X_M, Y_M, Z_M]^T$ und Normale $\boldsymbol{N}$, definiert mit den Winkeln $\alpha$ und $\beta$, Abb. 3, so dass $\boldsymbol{N} = [\sin\beta, -\sin\alpha\sin\beta, \cos\alpha\cos\beta]^T$ und $\boldsymbol{N}^T\boldsymbol{N} = 1$.

Objektpunkte $\boldsymbol{X} = [X, Y, Z]^T$ auf der $\mathcal{K}$-Ebene erfüllen die Ebenengleichung $\boldsymbol{N}^T(\boldsymbol{X} - \boldsymbol{M}) = 0$. Der $\mathcal{K}$-Ebene wird ebenfalls ein 3D-Rechtskoordinatensystem $xyz$ zugeordnet mit der $z$-Achse auf der Normalen $\boldsymbol{N}$. Punkte haben die Koordinaten $\boldsymbol{x} = [x, y, z]^T$, mit $z \equiv 0$, ihr Ursprung sei in der $\mathcal{K}$-Mitte $\boldsymbol{M}$ . Eine der Achsen kann man frei wählen wegen der Rotationssymmetrie des Kreises: $\boldsymbol{y}$ liege hier horizontal und normal zu $\boldsymbol{X}$ und $\boldsymbol{N}$. Der $\mathcal{K}$-Kreis

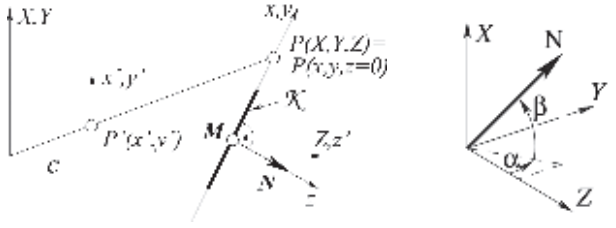



**Abb. 3:** Koordinatensysteme und Kreismarke: Objektkoordinatensystem $XYZ$, Bildkoordinatensystem $x'\ y'\ z'$ mit Kamerakonstante $c$, $\mathcal{K}$-Koordinatensystem $xyz$ mit Ursprung in $\mathcal{K}$-Mitte $\boldsymbol{M}$, $\mathcal{K}$-Durchmesser $2r$, Definition der Normalen $\boldsymbol{N}$ durch $\alpha$, $\beta$.

**Tab. 1:** Exzentrizitätswerte $|\Delta(\mathbf{e}')|$ in (µm) von elliptischen $\mathcal{K}$-Bildern in Abhängigkeit vom Ellipsendurchmesser $2a$ = 100 µm und der $\mathcal{K}$-Orientierung $\alpha°$, $\beta°$ in 3D. Kamera Nikon D700 mit $c$ = 12 mm und Bildwinkel ca. 55°.

| $\alpha° / \beta°$ | 20/0 | 40/0 | 60/0 | 70/0 | 20/20 | 40/40 | 60/60 | 70/70 |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{K}$-Mitte auf $Z = z'$ | 0.067 | 0.103 | 0.090 | 0.067 | 0.086 | 0.099 | 0.050 | 0.024 |
| $\mathcal{K}$-Mitte am Bildrand | 0.238 | 0.619 | 1.29 | 1.63 | 0.642 | 1.52 | 2.06 | 2.12 |

habe den Radius $r$ und mit homogenen $\mathcal{K}$-Koordinaten $\boldsymbol{x} = [x, y, 1]^T$ und homogener Matrix $\mathbf{C}$ die Gleichung

$$\mathbf{x}^{\mathbf{T}}\mathbf{C}\mathbf{x} = 0 \text{ mit } \mathbf{C} = \begin{bmatrix} 1/r^2 & 0 & 0 \\ 0 & 1/r^2 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \quad (9)$$

Es sei nun eine Beziehung zwischen Punkten $\mathbf{x}$ und ihren Bildpunkten $\mathbf{x}'$ in Form einer 2D-2D Homographie aufgestellt (mit fünf Parametern und $c$ = 1). Man findet sie durch Rotation mit $\boldsymbol{R}_M$ und Translation mit $\boldsymbol{M}_M$ von $\mathbf{x}$ nach $\boldsymbol{X}$ und Perspektivabbildung ins Bildsystem mit $[\mathbf{I}\ \ \mathbf{0}]$ zu

$$\mathbf{x}' = \begin{bmatrix} x't' \\ y't' \\ t' \end{bmatrix} = [\mathbf{I} \quad \mathbf{0}] \begin{bmatrix} \boldsymbol{R} & \boldsymbol{M} \\ \boldsymbol{0}^T & 1 \end{bmatrix}_M \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} c\cos\beta & 0 & cX_M \\ \sin\alpha\sin\beta & c\cos\alpha & cY_M \\ -\cos\alpha\sin\beta & \sin\alpha & Z_M \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (10)$$

oder

$$\mathbf{x}' = \mathbf{H}_M\mathbf{x}.$$

Mit dieser Homographie transformieren wir, vgl. Anhang 7.3, die Kreismatrix $\mathbf{C}$ ins Bild und erhalten die Ellipse: $\mathbf{C}' = \mathbf{H}_M^{-T}\mathbf{C}\mathbf{H}_M^{-1}$. Alles Weitere geschieht mit den Gleichungen (8).

Der Betrag der Exzentrizität $\Delta(\mathbf{e}')$ hängt dominant von der Ellipsen-Größe ab sowie von der relativen Lage und Orientierung der $\mathcal{K}$-Ebene zur Kamera. Natürlich wären überall im Bild die Exzentrizitäten Null, sofern $\alpha = \beta = 0$ wären. Die Neigungen der $\mathcal{K}$ spielen also immer eine Rolle. Wir bestimmen die Exzentrizitätsbeträge bei unterschiedlichen Winkeln $\alpha, \beta$ und berücksichtigen dabei, dass, falls $\mathcal{K}$ zentrisch auf der Aufnahmerichtung $Z = z'$ liegt, die Beträge der Exzentrizitäten Symmetrien besitzen (Tab. 1).

Symmetrien der Exzentrizitäten für Ellipsen in Bildmitte:
$|\Delta\,\mathbf{e}'(\alpha, \beta)| = |\Delta\,\mathbf{e}'(\beta, \alpha)| = |\Delta\,\mathbf{e}'(-\alpha, -\beta)|$;
projizierte $\mathcal{K}$-Mitte $\mathbf{m}'$ in (mm): in Bildmitte $[0\ 0\ 12]^T$, am Bildrand $[17\ 17\ 12]^T$.

Schon beim Ellipsen-Durchmesser von 100 µm hat man Exzentrizitäten, insbesondere am Bildrand, die bei genauen Bündellösungen (Punktmessgenauigkeit bei 0.2 µm (Vosselman & Förstner 1988, Luhmann 2010)) nicht vernachlässigbar sind. Im übrigen, Exzentrizitätswerte $|\Delta\,\mathbf{e}'|_i$ für andere Ellipsendurch-

messer $2a_i$ erhält man aus den Tabellenwerten recht gut über die Relation $|\Delta \mathbf{e}'|_i = (2a_i / 2\alpha)^2 |\Delta \mathbf{e}'|$ bei sonst übereinstimmenden Parametern. Exzentrizitätswerte für beispielsweise $2a_i = 50$ µm sind um den Faktor 1/4 kleiner als in der Tabelle, also am Bildrand nicht tolerierbar.

## 4 Korrespondierende Ellipsen einer Kreismarke in Bildern eines Bildverbands

Es fragt sich nun, wie man eine $\mathcal{K}$ in der Bündellösung eines Verbandes digitaler Bilder (iterative Gauß-Markov-Schätzung) behandeln sollte. Es werden drei Korrespondenzmodelle für überbestimmte $\mathcal{K}$ diskutiert und bei allen sei – gewissermaßen als Vorlauf – eine auf "Punkten im weiteren Sinne" ("natürliche Punkte") gegründete Bündellösung bereits errechnet worden. Unter derartigen "Punkten" versteht man lokale Bildfunktionen mit starken Gradienten, die eine genaue 2D-Lokalisierung ermöglichen, sowie mit individuellen Merkmalen, die das Auffinden von Korrespondenzen in anderen Bildern mit hoher Wahrscheinlichkeit leisten, siehe z. B. den SIFT-Operator (shift invariant feature transform) nach Lowe (2004). Oft begnügt man sich mit dem Ergebnis einer solchen Zuordnung, bleibt damit aber suboptimal.

Die Qualität eines photogrammetrischen Bündelblocks hängt bekanntlich nicht nur von seiner Makrostruktur ab (Zahl und Verteilung von Bildern und Objektpunkten, Bildmaßstäbe, etc.), sondern auch von der Schnittqualität korrespondierender Strahlenbündel und damit von den Korrespondenzmodellen mit der Mikrostruktur der Bilder, repräsentiert durch ihre Pixel. Anders als bei geodätischen Netzen (genaue Standpunkts-/Zielpunkts-Zentrierungen mit festen Marken), kann man die Korrespondenz in photogrammetrischen Netzen – wenn man die Mikrostruktur voll ausschöpfen will – erst im Prozess der Blockausgleichung realisieren; Korrespondenzmodelle stellen somit ein wesentliches Element des Konzepts digitaler Bündelblockausgleichungen dar.

Die weiteren Erörterungen sollen nun an folgender Stelle im Ablauf der Bündellösung beginnen: Die "korrespondierenden Punkte" des gesamten Bündelverbands seien mit ihrer lokalen Bildumgebung vorhanden. Dabei seien die Kreis-Marken als solche erkannt, ihre Ellipsenpunkte aus den digitalen Grauwerten (Bestrahlungsäquivalente) extrahiert und daraus Ellipsenmatrizen $\mathbf{C}'$ bestimmt worden. Man habe sodann die vorerst unkorrigierten Ellipsenmitten $\mathbf{e}'_k$, siehe (8), in $1\ldots K$ Bildern in der Bündellösung verwendet. Die Iterationen seien bereits fortgeschritten, so dass grobe Fehler behoben und recht gute Resultate für den Bildverband erreicht worden seien, also auch die 3D-Koordinaten einer $\mathcal{K}$-Mitte $\boldsymbol{M}$ und ihre Abbildungen $\mathbf{m}'_k$ in $K$ Bildern.

Die Genauigkeit dieser Daten ist auch deshalb relativ hoch, weil der Einfluss der noch nicht berücksichtigten Exzentrizitätswerte auf den Bildverband durch die Gaußsche Minimierung reduziert wird. Um die Exzentrizitäten mit Formel (8) bestimmen zu können, fehlt die Normale der $\mathcal{K}$-Ebene. Machen wir uns bewusst, die Objektkoordinaten eines Bildverbands sind bei normaler Iteration (die Zielfunktion der "Kleinsten Quadrate" ist dem globalen Minimum bereits recht nahe) eindeutig, also auch die Mitte $\boldsymbol{M}$ einer $\mathcal{K}$ und ihre projizierten Mitten $\mathbf{m}'_k$, so dass die $\mathcal{K}$-Ebenennormale $\boldsymbol{N}$ eindeutig aus Gleichung (3) errechnet werden kann, wegen zufälliger Fehler vorzugsweise als Mittelwert

$$\boldsymbol{N} = \frac{1}{K} \sum_k \mathbf{C}'_k \mathbf{m}'_k \qquad k = 1\ldots K. \tag{11}$$

In (11) müssen die homogenen Summanden normiert sein. Die Normale ist allerdings noch mit kleinen Exzentrizitätsfehlern aus den $\mathbf{m}'_k$ kontaminiert, die jedoch im Laufe der Folgeiterationen – $\boldsymbol{N}$ ist mit (11), (3) über $\mathbf{m}'_k$ und $\boldsymbol{M}$ Teil der Gaußschen Minimierung – gegen Null gehen sollten. Die mehrfachen Lösungen für die $\mathcal{K}$-Mitten, siehe (7), finden also im Bildverband eine unproblematische Abklärung; der gezeigte Weg ist einer von mehreren möglichen.

Aufbauend auf die einführende Betrachtung sei nun gezeigt, dass es heutzutage eine Reihe von photogrammetrischen Möglichkeiten gibt, die $\mathcal{K}$-Bearbeitung in die Bündellösung zu integrieren, solche, die mit den aus den Bildfunktionen extrahierten geometrischen Merkmalen (Punkte, Kreise, udgl.) ope-

rieren (klassische Bündellösung, siehe KAGER 1981, ANDRESEN 1991), und solche, die die digitalen Bildfunktionswerte (Grauwerte der Pixel) direkt in der Bündellösung verwenden und mittels digitaler Bildinversion die Geometrie-Extraktion mit der Grauwertberechnung der $\mathcal{K}$ gemeinsam durchführen (GRÜN 1985, WROBEL 1990, SCHNEIDER 1991). Dieser Ansatz ist in der Lage, die statistisch beste Lösung zu erreichen. Beide Möglichkeiten seien im Folgenden skizziert, Vieles existiert bereits. Es sei betont, bei allen Verfahren benötigen wir die Ergebnisse aus den Teilen 2 und 3.

## 4.1 *Korrespondenzmodelle der klassischen Bündellösung*

### 4.1.1 *Das Korrespondenzmodell mit $\mathcal{K}$-Mitten*

Wir betrachten im Folgenden wieder nur eine $\mathcal{K}$, ein Index für sie kann deshalb entfallen. Jede korrigierte Ellipsenmitte $(\mathbf{e}'+\Delta(\mathbf{e}'))_k$, siehe (8), steht als Repräsentant für die Ellipse $k$ im Bilde $k$ zum Objektpunkt $\mathbf{M}$. Für ihre statistische Unsicherheit wird angenommen, dass $\mathbf{D}(\mathbf{e}'+\Delta(\mathbf{e}'))_k \approx \mathbf{D}(\mathbf{e}'_k)$; sie kann durch Varianzfortpflanzung aus $\mathbf{D}(c'_{ij})_k$ hergeleitet werden, siehe (2). Damit könnte man $(\mathbf{e}'+\Delta(\mathbf{e}'))_k$ wie jeden üblichen Bildpunkt $P'$ des Bildverbandes behandeln, d.h. mit den homogenen Kollinearitätsgleichungen (siehe z. B. FÖRSTNER & WROBEL 2004, 222)

$$\left(\mathbf{e}'+\Delta\left(\mathbf{e}'\right)\right)_k + \hat{\mathbf{v}}_{e'_k} = \mathbf{K}\,\hat{\boldsymbol{R}}_k\left[\boldsymbol{I}\,|-\hat{\boldsymbol{X}}_{0,k}\right]\hat{\mathbf{M}}\;;$$
$$\mathbf{D}\left(\mathbf{e}'_k\right) \quad k = 1\ldots K \tag{12}$$

in den Bildverband eingliedern, also den $\mathcal{K}$-Mittelpunkt $\hat{\mathbf{M}}$ durch Strahlenschnitt aller seiner $K$ korrigierten Bildstrahlen $(\mathbf{e}'+\Delta(\mathbf{e}'))_k$ im Bündelprozess schätzen.

In (12) bedeuten:

$\hat{\mathbf{v}}_{e'_k}$, $\mathbf{D}\left(\mathbf{e}'_k\right)$ Zufallsvektor und Dispersionsmatrix von $\mathbf{e}'_k$

äußere und innere Orientierung von Bild $k$:

$\mathbf{K} = [\mathbf{I}]$ $3 \times 3$ Kameramatrix

$\hat{\boldsymbol{R}}_k$ geschätzte Rotationsmatrix

$\hat{\boldsymbol{X}}_{0,k}$ geschätztes Projektionszentrum.

Ansatz (12) umfasst die üblichen neun Parameter, die für die Eingliederung eines Bildstrahles ins Bündelnetz nötig sind. Die Dispersionsmatrizen können sich von Ellipse zu Ellipse – je nach ihrer Größe – unterscheiden; eine Varianzkomponentenschätzung empfiehlt sich daher. Für die numerische Durchführung der Gauß-Markov-Schätzung muss man die bekannte Linearform von (12) einsetzen. In heutigen Bündelblockprogrammen geht man dabei zu euklidischen Koordinaten über, künftig wäre aber vorteilhafterweise die direkte Verwendung homogener Koordinaten möglich (FÖRSTNER 2010), so dass man auch Fernelemente mit ihren Unsicherheitsmatrizen auf elegante Weise einbeziehen kann.

Die Korrektur $\Delta\left(\mathbf{e}'\right)_k$ kann erst im Laufe der Iterationen berechnet werden. Nach erreichter Konvergenz des Bildverbandes unterscheiden sich korrigierte Ellipsenmitte und projizierte $\mathcal{K}$-Mitte nur noch durch zufällige Fehler $\hat{\mathbf{v}}_{e'_k}$

$$\left(\mathbf{e}'+\Delta\left(\mathbf{e}'\right)\right)_k + \hat{\mathbf{v}}_{e'_k} = \hat{\mathbf{m}}'_k\;. \tag{13}$$

Bei diesem Korrespondenzmodell gehen sämtliche neun Parameter der Beobachtungsgleichungen (12) als Unbekannte in den Bildverband ein. Unbekannte, die nur für die $\mathcal{K}$ in (12) vorkommen, gibt es nicht. Vorhandene Bündelprogramme kann man also leicht mit der Exzentrizitätskorrektur ausstatten.

Es ist offensichtlich, dass mit der Ellipsenmitte $\mathbf{e}'_k$ allein nicht das Potenzial, das in den vielen Punkten $P'_l, l = 1\ldots L_k$, des Ellipsenrandes steckt, ausgeschöpft wird. $\mathbf{e}'_k$ ist insofern kein vollwertiger Repräsentant der Ellipse. Volle Repräsentanz wird erst mit dem folgenden Ansatz erreicht. Mit ihm hat man nicht mehr das Exzentrizitätsproblem der Ellipsenmitten, und die statistische Unsicherheit der gesuchten Größen gründet sich direkt auf die der extrahierten Ellipsenpunkte.

### 4.1.2 *Das Korrespondenzmodell mit rekonstruiertem Kreis*

In der Tat, die statistische Qualität des Ergebnisses lässt sich steigern, wenn man dieselben Datensätze von Ellipsen, die zu einer $\mathcal{K}$ gehören, siehe Teile 2 und 4.1.1, einem Korrespondenzmodell Kreis-Ellipse unterwirft: die Randpunkte $P'_{kl}$ aller Ellipsen $\mathcal{E}_k$ einer $\mathcal{K}$ können zur **gemeinsamen $\mathcal{K}$-Rekonstruktion** durch Gauß-Helmert-Schätzung herangezo-

gen werden. Der ganze Kreis der $\mathcal{K}$ mit seiner Normalen ist hier das korrespondierende geometrische Merkmal für die Ellipsen und der rekonstruierte Kreis im Objektraum das Ziel. Damit erreicht man eine optimale Auswertung der **vorhandenen Ellipsendaten**. Statt mehrerer, unabhängiger Ellipsenausgleichungen wie in 4.1.1 hat man hier nur eine einzige und damit bessere Schätzung, die den $\mathcal{K}$-Kreis im Raum festlegt, nicht nur seinen Mittelpunkt.

Geometrisch gesehen definiert jede Ellipse $\mathcal{E}_k$ einen eigenen Kegel. Seine Spitze ist jeweils das Projektionszentrum von Bild $k$, seine Mantellinien sind die Abbildungsstrahlen der Ellipsenpunkte. Der Kegel erhält aber nun einen – im Gegensatz zu 4.1.1 – mit allen anderen Kegeln streng gemeinsamen Schnittkreis in der $\mathcal{K}$-Ebene. Im Modell zuvor schnitten sich alle korrespondierenden Strahlen der korrigierten Ellipsenmitten $\mathbf{m}'_k$ gemeinsam in einem Objektpunkt $\mathbf{M}$, einem Gebilde mit Dimension 0. Hier nun schneiden und erzeugen damit alle Strahlen korrespondierender Ellipsen einen gemeinsamen Kreis im Objektraum, einem Gebilde der Dimension 1. Es ist eine Aufgabe mit deutlich weniger Parametern und damit höherer Redundanz als in 4.1.1. Die algebraisch-numerische Lösung von 4.1.2 gehört zu den statistischen Standardproblemen, wenige Hinweise ohne Details mögen daher genügen.

Analog zu Teil 3 wird eine 2D-2D Homographie zwischen Punkten $\mathbf{x}_l$ der $\mathcal{K}$-Ebene und ihren korrespondierenden Punkten $\mathbf{x}'_{k,l}$ im Bild $k$ angesetzt, die hier über ein Objektkoordinatensystem in beliebiger Lage erreicht wird. Man erhält mit homogenen Koordinaten

$$\begin{bmatrix} x't' \\ y't' \\ t' \end{bmatrix}_{kl} = \begin{bmatrix} \boldsymbol{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{R} & -\boldsymbol{R}\boldsymbol{X}_0 \\ \mathbf{0}^T & 1 \end{bmatrix}_k \begin{bmatrix} \boldsymbol{R} & \boldsymbol{M} \\ \mathbf{0}^T & 1 \end{bmatrix}_{\mathcal{K}} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}_l, \tag{14}$$

mit $k = 1 \ldots K$ = Zahl an Ellipsen für die $\mathcal{K}$, $l = 1 \ldots L_k$, $L_k$ = Zahl an Punkten der Ellipse $k$ und $z_l \equiv 0$ für alle $l$.

In (14) bewirkt die erste Matrix (von rechts) die Bewegung der Koordinaten $\mathbf{x}$ der $\mathcal{K}$-Ebene ins allgemeine 3D-System XYZ (5 Parameter: die Orientierung der $\mathcal{K}$ mit $\alpha, \beta$ in $\boldsymbol{R}_{\mathcal{K}}$, 3D-Mitte $\boldsymbol{M}_{\mathcal{K}}$), die zweite Matrix die weitere Bewegung zum Bildkoordinatensystem (6 Parameter: äußere Orientierung des Bildes $k$) und die dritte stellt die Perspektive her. In kompakter Schreibweise lautet die Beziehung

$$\mathbf{x}'_{kl} = \mathbf{H}_{\mathcal{K},k}\mathbf{x}_l, \tag{15}$$

worin $\mathbf{H}_{\mathcal{K},k}$ unter Berücksichtigung von $z_l \equiv 0$ das 3 × 3-Produkt der Matrizen in (14) ist. Da in der $\mathcal{K}$-Ebene nur Kreispunkte zugelassen sind, müssen die Punkte $\mathbf{x}_l$ die Kreisgleichung mit einem Radius-Parameter $r$ erfüllen, siehe (9). Die unbekannten Parameter in (14) können in einer Gauß-Helmert-Schätzung aus den Bedingungen (9) nach Substitution der Kreispunkte $\mathbf{x}_l$ durch gemessene Bildpunkte $\mathbf{x}'_{kl}$, siehe (16), (17), bestimmt werden

$$\hat{\mathbf{v}}'_{kl} + \mathbf{x}'_{kl} = \hat{\mathbf{H}}_{\mathcal{K},k} x_l \,; \quad \mathbf{D}\left(\mathbf{x}'_{kl}\right) \tag{16}$$

$$\left(\hat{\mathbf{v}}'_{kl} + \mathbf{x}'_{kl}\right)^T \hat{\mathbf{H}}_{\mathcal{K},k}^{-T} \hat{\mathbf{C}} \hat{\mathbf{H}}_{\mathcal{K},k}^{-1} \left(\hat{\mathbf{v}}'_{kl} + \mathbf{x}'_{kl}\right) = 0 \tag{17}$$

mit

| | |
|---|---|
| $\mathbf{x}'_{k,l}$ | Messpunkt $l$ der Ellipse $k$ zum Kreis $\mathbf{C}$ |
| $\mathbf{x}_l$ | Kreispunkt korrespondierend zum Bildpunkt $\mathbf{x}'_{k,l}$ |
| $\hat{\mathbf{v}}'_{kl}$, $\mathbf{D}(\mathbf{x}'_{kl})$ | Zufallsvektor und Dispersionsmatrix von $\mathbf{x}'_{k,l}$. |

Die Gesamtzahl an Gleichungen (17) von allen Ellipsen für eine $\mathcal{K}$ ist, $\Sigma L_k$, $k = 1 \ldots K$. Die Parameteranzahl für die $\mathcal{K}$ allein beträgt sechs (evtl. fünf, falls ihr Radius $r$ bekannt ist). Hinzu kommen 6K Parameter der äußeren Orientierung der Bilder, in denen die $\mathcal{K}$ vorkommt. Falls $r$ tatsächlich gegeben ist, benötigt man zur Bestimmung einer $\mathcal{K}$ nur eine Ellipse (mit mehr als sechs Punkten; dies ist die algebraische Entsprechung zum Experiment in der Einführung, Abb. 2), sonst zwei, zur Genauigkeitssteigerung eher mehr.

Die Näherungswerte der Parameter für die erforderliche Linearisierung von (17) lassen sich leicht beschaffen: die äußere Orientierung der Bilder und die $\mathcal{K}$-Mitte $\boldsymbol{M}$ liefert die Bündellösung, wenn man wie in 4.1.1 die ers-

ten Iterationen mit den Ellipsenmitten $\mathbf{e}'_k$ beginnt; für die $\mathcal{K}$-Normale benötigt man von allen beteiligten Ellipsen $\mathcal{E}_k$ eigentlich nur eine Matrix $\mathbf{C}'_k$, woraus mit (3) die Näherung für die Normale und daraus die Winkel $\alpha$, $\beta$ folgen. Mit denselben Daten kann man auch eine Näherung für den Radius $r$ der $\mathcal{K}$ erhalten. Die Linearisierung von (17) erscheint sehr kompliziert. Jedoch, Linearisierungen können auch relativ leicht numerisch erzeugt werden.

Das Ergebnis dieses Korrespondenzmodells ist die $\mathcal{K}$-Mitte zusammen mit Ebenennormale und Kreisradius; für alle kann man die statistische Unsicherheit schätzen. Wie im Modell zuvor ist bei unterschiedlich großen Ellipsen eine Varianzkomponentenbestimmung für die $\mathbf{D}(\mathbf{x}'_k)$ angebracht. Vorteilhafterweise erreicht man hier eine höhere Redundanz als in Kapitel 4.1.1: Aus derselben Menge an Ellipsenpunkten ist hier nur die gemeinsame $\mathcal{K}$-Ebene und ein Kreisparameter zu ermitteln anstelle der Parameter von allen beteiligten Ellipsen. Es ist daher zu erwarten, dass eine $\mathcal{K}$ mit dem Modell 4.1.2 im Bildverband eine übergeordnete Genauigkeit erreicht. Bei vielerlei Anwendungen würde dies den Absichten entgegenkommen, wenigstens einige Punkte eines Bildverbandes mit Kreismarken zu signalisieren, um die Stabilität zu stärken und/oder mit genauen geodätischen Koordinaten zu verbinden.

Ein Charakteristikum des Modells 4.1.2 soll uns noch beschäftigen. Seine Gleichungen (17) enthalten Unbekannte mit unterschiedlicher Bedeutung für den Bildverband: solche, die nur die Eigenschaften der benutzten $\mathcal{K}$ modellieren, also eine lokale Aufgabe zu erfüllen haben, und solche, die für den Aufbau des Bildverbandes und seine Stabilität unentbehrlich sind. Die Handhabung von Normalgleichungen der Gaußschen Schätzmethode erlaubt nun, hierauf mit Vorteil zu reagieren. Man kann die Normalgleichungen aus (17) vollständig in das Gesamtsystem des Bildverbandes integrieren oder die lokalen Parameter zuvor durch Gaußsche Eliminationsschritte entfernen, ohne die Strenge der Parameterschätzung und ihrer Genauigkeit zu beeinträchtigen – ein seit dem 19. Jahrhundert bekanntes Verfahren. Eine lokale Rechnung stellt am Ende den Gesamtzusammenhang wieder her. Es liegt nahe, von (17) die $\mathcal{K}$-Mitte $\boldsymbol{M}$ mit den Orientierungsparametern von Bild $k$ ins Gesamtsystem einzufügen, die anderen zu eliminieren. Der Punkt $\boldsymbol{M}$ ist daher als ein vollwertiger, starker Repräsentant der Ellipsen anzusehen. Dann ist man eigentlich wieder recht nahe beim ersten Korrespondenzmodell angelangt, siehe 4.1.1, aber mit viel besseren Eigenschaften.

## 4.2 *Statistisch optimale Bündellösung mit Pixel-Bündeln durch gemeinsame Rekonstruktion der $\mathcal{K}$-Geometrie und ihrer Helligkeitsfunktion*

Die bisher diskutierten Korrespondenzmodelle der klassischen Bündellösung 4.1.1 und 4.1.2 gehen von Ellipsenpunkten aus. Jedoch, sie stellen bekanntlich keine Originalmessdaten dar, sondern mussten erst durch lokale Schätzungen aus den digitalen Bildfunktionen abgeleitet werden, auch ihre a priori Standardabweichungen. Etwa seit den 1975er Jahren gibt es Vorschläge, die Methoden und Aufgaben der Photogrammetrie direkt mit den originalen Bildgrauwerten zu lösen (Überblicke in Luhmann 2010, Triggs et al. 2000, McGlone et al. 2004). Von Anfang an spielten die Kleinste-Quadrate-Zuordnungen eine besondere Rolle, zunächst von Bild zu Bild, z. B. Grün (1985). Seit 1985 gibt es den Vorschlag, sie konsequent als **digitale Inversion der Bildentstehung** vom Bild zurück zum Objekt anzusetzen, d.h. Rückprojektion aller Pixel-Grauwerte entlang ihrer Abbildungsstrahlen auf ein gemeinsames Rekonstruktions-Modell von Geometrie und Helligkeit des Objekts und Optimierung der Modellrechnung mit der Kleinste-Quadrate-Methode. Darin werden nicht mehr die zufälligen Fehler von Bildkoordinaten der extrahierten Bildpunkte minimiert, sondern die Zufallskomponenten $\mathbf{v}_{g'(x')}$ der originalen Bildgrauwerte $g'(\mathbf{x}')$. Näheres in Wrobel (1985/86), Wrobel (1987), Schlüter (1999a) und Schlüter (1999b) mit Verallgemeinerungen, topographischen Anwendungsbeispielen und vielen Referenzen. Mit der Leistungsfähigkeit der heutigen Computertechnik kann man diese direkten, aber aufwendigeren Wege gehen mit Vorteilen für die Automation und Qualität der Verfah-

ren. In der Tat, seit einiger Zeit befasst man sich erneut intensiv mit derartigen Ansätzen in großer Allgemeinheit und Vielfalt, siehe z. B. Strecha (2007), vor allem Cremers et al. (2010).

Die geometrische Anschauung zum Korrespondenzmodell hier unterscheidet sich fundamental von den anderen, da sie ausschließlich über die originalen Bildgrauwerte gebildet wird. Es sei unterstellt, dass eine $\mathcal{K}$ sich auf einer quadratischen Ebene befinde (Abb. 1), die in den Bildern $k$ je in ein anderes Viereck übergeht. Die Abbildungsstrahlen aller Pixel eines Vierecks (es sei Pixel-Bündel genannt) stellen eine "gefüllte Pyramide" dar, ihre Spitze liegt im Projektionszentrum des Bildes $k$. Das Ziel hier besteht darin, dass alle Pyramiden $k = 1 \ldots K$ einer $\mathcal{K}$ eine streng gemeinsame $\mathcal{K}$ im Objektraum rekonstruieren, mit identischen Oberflächenhelligkeiten und identischer Geometrie, unter der Gaußschen Minimumsbedingung bestmöglich aus allen Bildgrauwerten einer $\mathcal{K}$ geschätzt. Die Gleichheit von Bildgrauwerten (nach lokaler linearer Anpassung) mit geometrisch korrespondierenden Objektgrauwerten ist also das Prinzip. Korrespondenz, auf diesem Wege erzeugt, verwendet alle vorhandenen, ursprünglichen Messdaten. Im Vergleich zu den Korrespondenzmodellen zuvor hat das hier erzeugte Korrespondenzgebilde im Objektraum, die Helligkeitsfunktion und die Geometrie des Objekts, die Dimension 2.

Der Ansatz eines derartigen Korrespondenzmodells sei nun aufgezeigt (Kraut & Wrobel 1993). Nur schwarz-weiß Bilder werden behandelt. Es besteht aus drei klar unterscheidbaren Modellen: der optisch-photographischen Bildentstehung (bei kleinen Objekten in starker Vereinfachung meist hinreichend), der Objektoberfläche und ihrer Helligkeitsfunktion. Im Falle einer $\mathcal{K}$ hat man bereits einige Kenntnis über das zu rekonstruierende Objekt: die Helligkeitsfunktion hat eine sehr einfache Struktur (Parameter $G_1$, $G_2$, Kreisradius $r$, $s$ für die PSF (Punktstreufunktion); siehe Abb. 1 und Anhang 7.4), die Geometrie besteht aus der Trägerebene (fünf Parameter). Es gilt daher die Geometrie und die Näherungswertbeschaffung wie in 4.1.2. Die Rückprojektion der Bildgrauwerte geschieht sodann auf eine vorerst nur ungenau bekannte $\mathcal{K}$-Ebene mit ebensolchen Objektgrauwerten $G$. Der Zusammenhang der Bildgrauwerte mit den Änderungen aller Parameter des Korrespondenzmodells wird in einer Taylorreihe erreicht und die finale Lösung kann per Newton-Gauß-Iteration gefunden werden, vorausgesetzt die Näherungswerte befinden sich im Konvergenzkreis der Lösung.

Im Folgenden gehen wir von homogenen zu euklidischen Koordinaten über. Beispielsweise wird aus $\mathbf{x}' = [x't',y't',t']^T$ der euklidische Vektor $\boldsymbol{x}' = [x',y']^T$. Die Beziehung zwischen Punkten $\mathbf{x} = [x,y,1]^T$ der $\mathcal{K}$ und ihren Bildpunkten $\mathbf{x}'$, siehe (15), lautet nunmehr

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \boldsymbol{x}' = \begin{bmatrix} \mathbf{h}_{K,k,1}^T \mathbf{x} \\ \mathbf{h}_{K,k,2}^T \mathbf{x} \end{bmatrix} / \mathbf{h}_{K,k,3}^T \mathbf{x} \quad \text{mit}$$

$$\mathbf{H}_{\mathcal{K},k} = \begin{bmatrix} \mathbf{h}_1^T \\ \mathbf{h}_2^T \\ \mathbf{h}_3^T \end{bmatrix}_{\mathcal{K},k} . \qquad (18)$$

Wir betrachten die Pixel einer $\mathcal{K}$ im Bilde $k$. Der Bildstrahl $l$ des Pixels $\boldsymbol{x}'_{kl}$ mit Grauwert $g'_{kl}$ trifft die $\mathcal{K}$-Ebene im Punkt $x_{kl}$. Bei fehlerfreien Daten und einer $\mathcal{K}$ mit matter Oberfläche (Lambertstrahler) gelte zunächst folgende lineare Beziehung als Approximation des strahlungsphysikalischen Reflexions- und Sensorverhaltens

$$h'_{0,k} + h'_{1,k}\, g'\ (\boldsymbol{x}'_{kl}) = G\,(\boldsymbol{x}) \quad l = 1 \ldots L_k \qquad (19)$$

mit sechs Parametern

$h'_{0,k}, h'_{1,k}$ Parameter für den Transfer der $g'_{kl}$ in die $G_l$ auf $\mathcal{K}$, gültig für die Pixel $l = 1 \ldots L_k$ des $\mathcal{K}$–Bildes $k$

$G\,(\boldsymbol{x})$ $G\,(\boldsymbol{x})$ besteht aus drei Bereichen mit vier unbekannten Parametern, siehe Anhang 7.4.

Die Bildgrauwerte $g'(\boldsymbol{x}')$ sind jedoch Messdaten mit einer stochastischen Komponente $v_{g'}(\boldsymbol{x}')$, und die Parameter der Abbildungsmatrix $\mathbf{H}_{\mathcal{K},k}$ und von $G\,(\boldsymbol{x}_l)$ sind vorerst nur approximativ bekannt. Der Gesamtzusammenhang wird mit folgender Taylorreihe von (19), entwickelt nach den unbekannten Parametern $p_i$ in $\mathbf{H}_{\mathcal{K},k}$ und $G\,(\boldsymbol{x}_l)$, dargestellt

$$\hat{h}'_{0,k} + \hat{h}'_{1,k}\{g'(\boldsymbol{x}'_{kl}) + \hat{v}_{g'}(\boldsymbol{x}'_{kl})\} =$$
$$\hat{G}(\boldsymbol{x}_l) + \sum_{i=1}^{15} (d\hat{G}(\boldsymbol{x}_l)/d\boldsymbol{x}_{kl}) d\boldsymbol{x}_{kl}/d\hat{p}_{k,i}) d\hat{p}_{k,i};$$
$$\boldsymbol{D}(g'(\mathbf{x}'_{kl})), \qquad (20)$$

mit

| | |
|---|---|
| $l = 1 \ldots L_k$ | $L_k$ = Pixelzahl der $\mathcal{K}$ im Bilde $k$ |
| $i = 1 \ldots 15$ | $d\hat{p}_i$ = geschätzte Änderungen der $\hat{p}_i$: 11 in $\mathbf{H}_{\mathcal{K},k}$ (fünf für die $\mathcal{K}$ sowie sechs für die äußere Orientierung von Bild $k$) und vier in $\hat{G}(\boldsymbol{x}_l)$ |
| $\hat{v}_{g'}(\boldsymbol{x}'), \boldsymbol{D}(g'(\boldsymbol{x}'_{kl}))$ | Zufallskomponente und Dispersionsmatrix von $g'(\boldsymbol{x}'_{kl})$. |

Die unbekannten Parameter in (20) sind in Iterationen zu verbessern: $\hat{p}_i + d\hat{p}_i \rightarrow \hat{p}_i$ bis $d\hat{p}_i \rightarrow 0$. Dann ist der Summenterm von (20), rechts, verschwunden und man ersieht unmittelbar die hier gültige Korrespondenz. Wie zuvor in 4.1 wird auf die Wiedergabe der Linearform zum Modell (20) verzichtet.

Die Parameterzahl 17 in den $\mathcal{K}$-Gleichungen (20) eines Bildes hat im Hinblick auf die im Allgemeinen sehr hohe Pixelzahl eine Redundanz zur Folge, mit der das stochastische Modell der Bildgrauwerte sich auf einer sicheren Basis befindet. Es wird betont, dass bei diesem Modell sämtliche Pixel der abgebildeten $\mathcal{K}$-Bereiche die Modellparameter mitbestimmen und nicht etwa nur eine Auswahl wie in den Modellen zuvor. Hervorzuheben ist ferner, dass es hier eine einzige statistische Schätzung gibt und zwar mit den Originalmesswerten. Eine Folge von Schätzungen wie zuvor in 4.1 ist in der Regel mit Verlusten von Korrelationen verbunden. Das Korrespondenzmodell mit Pixelbündeln wird daher als optimal angesehen.

Da es die $\mathcal{K}$-Geometrie wie in 4.1.2 besitzt, also mit der $\mathcal{K}$-Mitte $\boldsymbol{M}$, kann man hier wie dort die gleiche Strategie zum effizienten Aufbau des Bildverbandes (Bündelnetzaufbau) benutzen, d.h. einen relativ hohen Aufwand für eine möglichst realistische Modellierung der $\mathcal{K}$-Korrespondenzen betreiben, die Auflösung des Gesamtsystems des Bildverbandes aber damit nicht belasten. Aus den Normalgleichungen einer $\mathcal{K}$ aus einem Bild kann man die Transferparameter $h_0$,$h_1$ sofort reduzieren, die Gleichungen der vier G-Parameter und der Normalen $\boldsymbol{N}$ erst, wenn alle Bilder, die dieselbe $\mathcal{K}$ enthalten, ausgewertet wurden. $\mathcal{K}$-Mitte $\boldsymbol{M}$ und Orientierungsparameter von Bild $k$ gehen ins Gesamtsystem ein. An allen diesen Eigenschaften müssen sich andere Modelle zum Aufbau von Bündelblöcken messen lassen.

## 5 Zusammenfassung und Ausblick

Der Artikel befasst sich mit abgebildeten Kreismarken (=$\mathcal{K}$) im Einzelbild und im Bildverband, ausgehend von kalibrierten Digitalbildern, und bestimmt die Exzentrizität der projizierten Kreismitte von der zugehörigen Ellipsenmitte. Der Schwerpunkt liegt in der Herleitung von mathematischen Grundbeziehungen; Verfeinerungen und Ausbau zu rationellen Verfahren für konkrete Anwendungen bleiben offen, sollten aber mit den vorgelegten Ergebnissen möglich sein.

In den Teilen 1 bis 3 wird die Exzentrizität der projizierten $\mathcal{K}$-Mitte $\mathbf{m}'$ von der $\mathcal{K}$-Ellipsenmitte $\mathbf{e}'$ hergeleitet und erläutert. Die Beziehungen beruhen auf der Polarität zwischen $\mathbf{m}'$ und der $\mathcal{K}$-Normalen $\boldsymbol{N}$ in Bezug auf die $\mathcal{K}$-Ellipse $\mathbf{C}'$. Mit Hilfe einer SVD der Ellipsenmatrix erhält man die zwei Lösungen für $\mathbf{m}'$, also allein aus den Ellipsendaten. Wegen der Strenge der Lösung kann man auch große $\mathcal{K}$ verwenden, deren Ellipsen aus mehr Bildpunkten genauer bestimmbar sind als kleine Ellipsen, und die insbesondere als Elemente von Industrieobjekten oder in Bildern von Tiefbefliegungen über Städten oder in Nahaufnahmen von Gebrauchsgegenständen aus heutiger oder historischer Zeit vorkommen können. Im Teil 4 wird die Behandlung von $\mathcal{K}$ im Verfahren der Bündelausgleichung dargelegt. Die Mehrfachlösungen für die projizierte $\mathcal{K}$-Mitte $\mathbf{m}'$ erfahren unter dem Minimierungseinfluss der Gaußschen Methode der Kleinsten Quadrate eine unproblematische Abklärung.

Das eigentliche Thema von Teil 4 ist der Frage gewidmet, welches Korrespondenzmodell einer $\mathcal{K}$ im Bildverband aktuell angemes-

sen ist. Es werden drei Modelle diskutiert, von einfach und reduziert genau bis zu aufwendig und optimal genau. Sie ermöglichen es, den sehr unterschiedlichen praktischen Bedürfnissen gerecht zu werden. Die beiden ersten stehen – nach der Ellipsenextraktion – ausschließlich auf dem Boden der projektiven Geometrie und verwenden die extrahierten Ellipsenpunkte einer $\mathcal{K}$ in den Bildern: aus den Ellipsenmitten wird nur die $\mathcal{K}$-Mitte im Objektraum bestimmt bzw. die Punkte aller Ellipsen legen gemeinsam – d.h. genauer, robuster – die ganze $\mathcal{K}$ im Objektraum fest.

Das dritte Modell verfolgt das Ziel, den Bildverband aus den Originalbilddaten direkt auf geometrisch-strahlungsphysikalischer Grundlage der Bilderzeugung wieder herzustellen. Es wird durch digitale Bildinversion direkt die $\mathcal{K}$ im Objektraum rekonstruiert: ihre Geometrie und ihr Helligkeitsverlauf. Das Pixelbündel, das zu einer Ellipse samt lokaler Umgebung gehört, wird komplett herangezogen. Damit ist eine maximale Informationsausnutzung und Genauigkeit erreichbar. Es darf dabei nicht übersehen werden, dass dies alles mit einer Kombination von merkmalsbasierter Bündellösung mit anschließender Pixelbündel-Lösung zu geschehen hat. Die erste bewältigt die gewiss nicht leichte Aufgabe der Näherungswertbeschaffung (Objektpunkte, Orientierungsdaten, Neigungswinkel von $\mathcal{K}$-Ebenen, …), die zweite schöpft das hohe relative Positionierungspotenzial digitaler Bilder aus.

Der Artikel hat sich auf Modelle zur Verarbeitung von Kreismarken in der Bündellösung eines Bildverbandes konzentriert. Kreismarken sind zwar eine spezielle Punktverkörperung, dennoch ist das Korrespondenzmodell mit Pixelbündeln mit allen Eigenschaften auf allgemeine Punktdefinitionen ("natürliche Punkte", "key points") übertragbar und damit auf die heute am häufigsten in Bildverbänden vorhandene Punktart. Die aktuellen Verfahren zur Auffindung natürlicher Punkte sind sehr weit entwickelt und erprobt worden und sind bereits auf recht allgemeine Objekte des Nahbereichs anwendbar. Dies gilt für die Deskriptoren von Merkmalen (Lowe 2004) und auch für ihren Zuordnungsprozess, der zunehmend robuster wird (Dickscheid 2011). Die dabei auftretenden "Punkte" benötigen im Korrespondenzmodell für die Oberflächen-Geometrie und -Helligkeit (anders als bei einer $\mathcal{K}$) flexible Freiformansätze, bestehend aus Facetten (Maschen) in angemessener Dichte. Die erfolgreiche Handhabung solcher Ansätze ist mit dem Facetten-Stereosehen bereits untersucht worden (Kempa 1995, Schlüter 1999a, b).

Damit könnte die Bündellösung von Bildverbänden, eine schon lange etablierte Methode, eine günstige Weiterentwicklung erfahren. Aus dem klassischen Bündelblock, bestehend aus der Gesamtheit von Strahlen zu Objektpunkten, entstünde ein Block von Pixelbündeln je zur lokalen Umgebung von Objektpunkten. Der Aufwand für die Pixelbündellösung ist sicherlich größer, aber die Methode ist in vielerlei Hinsicht zeitgemäß.

## 6 Literatur


Andresen, K., 1991: Ermittlung von Raumelementen aus Kanten im Bild. – Zeitschrift für Photogrammetrie und Fernerkundung **1991** (6): 212–220.

Cremers, D., Goldlücke, B. & Pock, T., 2010: Variational Methods in Computer Vision. – Tutorial, 11th European Conference on Computer Vision, Heraklion, Kreta.

Dickscheid, T., 2011: Robust Wide-Baseline Stereo Matching for Sparsely Textured Scenes. – Dissertation Uni Bonn.

d'Home, M., Laprestel, J.T., Rives, G. & Richetin, M., 1990: Spatial localization of modelled objects of revolution in monocular perspective vision. – ECCV 1, 1990. – Goos, G. & Hartmanis, J. (eds.): Lecture notes in computer science **427:** 475–485, Springer Verlag, Berlin, Heidelberg.

Dold, J., 1997: Ein hybrides photogrammetrisches Industriemeßsystem höchster Genauigkeit und seine Überprüfung. – Dissertation, Universität der Bundeswehr, München, Schriftenreihe Studiengang Vermessungswesen, Heft **54**.

Förstner, W., 2010: Minimal Representations for Uncertainty and Estimation in Projective Spaces. – Asian Conference on Computer Vision **II:** 619–633, Queenstown, New Zealand.

Förstner, W. & Wrobel, B., 2004: Basic image geometry. – McClone, J.Ch., Mikhail, E. & Bethel, J. (eds.): Manual of Photogrammetry. – 5th edition, 211 p., American Society for Photogrammetry and Remote Sensing, Bethesda (ML), USA.

Förstner, W. & Wrobel, B. (in Vorbereitung): Photogrammetric Computer Vision, vol. **1**: Geometry and Orientation.

Forsyth, D., Mundy, J.L., Zisserman, A., Coelhol, C., Heller, A. & Rothwell, C., 1991: Invariants descriptor for D object recognition and pose. – IEEE transactions on PAMI **13:** 971–991.

Grün, A., 1985: Adaptive least sqares correlation – a powerful image matching technique. – South African Journal of Photogrammetry, Remote Sensing and Cartography **14** (3): 175–187.

Kager, H., 1981: Bündeltriangulation mit indirekt beobachteten Kreiszentren. – Dissertation, Technische Universität Wien, Geowissenschaftliche Mitteilungen, Heft **19**.

Kanatani, K., 1993: Geometric Computation for Machine Vision. – Clarendon Press, Oxford.

Kempa, M., 1995: Hochaufgelöste Oberflächenbestimmung von Natursteinen und Orientierung von Bildern mit dem Facetten-Stereosehen. – Dissertation TH Darmstadt, Institut für Photogrammetrie und Kartographie, Darmstadt.

Kraut, A. & Wrobel, B., 1993: Target localization by least-squares image matching including deconvolution of image blur. – Conf. Optical D Measurement Techniques, Oct. 1993, Zürich, Switzerland.

Lowe, D.G., 2004: Distinctive image features from scale-invariant keypoints. – Intern. Journal of Computer Vision **60** (2): 91–110.

Luhmann, T., 2010: Nahbereichsphotogrammetrie – Grundlagen, Methoden und Anwendungen. – 3. Auflage, Wichmann Verlag, Berlin, Offenbach.

McGlone, J.C., Mikhail, E.M. & Bethel, J.S. (eds.), 2004: Manual of Photogrammetry. – 5th edition, American Society for Photogrammetry and Remote Sensing, Bethesda(ML), USA.

Philip, J.S., 1997: An algorithm for determining the position of a circle in d from its perspective d projection. – Technical report TRITA-MAT-1997-MA-1, Department of Mathematics, KTH (Royal Institute of Technology), Stockholm.

Schlüter, M., 1999a: Von der D- zur D-Flächenmodellierung für die photogrammetrische Rekonstruktion im Objektraum. – Dissertation TU Darmstadt, Deutsche Geodätische Kommission C (**506**), München.

Schlüter, M., 1999b: Automatische Oberflächenrekonstruktion durch digitale Bildzuordnung, eine Zusammenstellung von drei Beiträgen. – Mitteilungen des Bundesamtes für Kartographie und Geodäsie, Band **4**, Frankfurt/Main.

Schneider, C.T., 1991: Objektgestützte Mehrbildzuordnung. – Dissertation TU Braunschweig, Deutsche Geodätische Kommission, Reihe C (375).

Strecha, C., 2007: Multi-View Stereo as an Inverse Inference Problem. – Dissertation, Katholieke Universiteit Leuven.

Triggs, B., McLauchlan, P.F., Hartley, R.I. & Fitzgibbon, A.W., 2000: Bundle adjustment – a modern synthesis. – Triggs, B., Zisserman, A. & Szeliski, R. (eds.): Vision algorithms, theory and praxis. – Lecture Notes in Computer Science, **1883**, Springer Verlag, Berlin New York.

Vosselman, G. & Förstner, W., 1988: The Precision of a Digital Camera. – ISPRS 16th Congress: 148–157, Kyoto.

Wrobel, B., 1985/1986: Studie über die digitale Bildinversion zur Rekonstruktion von Objektraummodellen (Oberflächen von Objekten, Bündellösung mit dem Facetten-Stereosehen). – Institut für Photogrammetrie und Kartographie, TH Darmstadt, unveröffentlicht.

Wrobel, B., 1987: Digitale Bildzuordnung durch Facetten mit Hilfe von Objektraummodellen. Einige Überlegungen über die theoretischen Grundlagen der digitalen Photogrammetrie. – Zeitschrift für Bildmessung und Luftwesen **55** (3): 93–101, **55** (4): 129–140.

Wrobel, B., 1987: Digital image matching by facets using object space models. – Intern. Symposion on Optical and Optoelectr. Science and Engineering, 30 March – April, The Hague (NL) 1987. – Osterlinck, A. & Tescher, A.G. (eds.): Advances in Image Processing. Proceedings SPIE **804:** 325–334.

Wrobel, B., 1990: Zentrumsbestimmung von Punktsignalen mit digitaler Bildverarbeitung. – Festschrift Prof. Dr. mult. G. Konecny zum 60. Geburtstag, Universtät Hannover, Institut für Photogrammetrie und Ingenieurvermessung.

# 7 Anhang

## 7.1 *Normale N der $\mathcal{K}$-Ebene im Objektraum und ihre Ferngerade im Bildraum*

Es existiere ein *XYZ*-System im Projektionssystem *O* mit *XY* parallel zu $x'$, $y'$ in der Bildebene, Kamerakonstante sei $c = 1$, so dass $X = x' / Z$, $Y = y' / Z$. Die Trägerebene einer $\mathcal{K}$ im Objektraum sei definiert durch ihre Normale $\boldsymbol{N} = [N_1 \;\; N_2 \;\; N_3]^T$ und den beliebigen festen Punkt $P^*$ ($X^*$, $Y^*$, $Z^*$), also gilt für jeden Punkt $\boldsymbol{X}$ der Ebene

$$\boldsymbol{N}^T [\boldsymbol{X} - \boldsymbol{X}^*] = 0 \quad \text{oder}$$
$$N_1 (X - X^*) + N_2 (Y - Y^*) + N_3 (Z - Z^*) = 0. \quad (21)$$

Wir bilden Punkte $X$, $Y$, $Z$ der $\mathcal{K}$-Ebene in ihre Bildpunkte $x'$, $y'$ ab und erhalten

$$N_1x' + N_2y' + N_3 = (N_1X^* + N_2Y^* + N_3Z^*) / Z\,. \tag{22}$$

Lässt man $Z$ gegen Unendlich gehen, folgt daraus das Bild der Ferngeraden der $\mathcal{K}$-Ebene

$$N_1x' + N_2y' + N_3 = 0 \quad \text{oder} \quad \boldsymbol{N}^T\mathbf{x}' = 0 \quad \text{oder} \quad \mathbf{l}'^T_\infty\mathbf{x}' = 0\,. \tag{23}$$

$\mathbf{l}'_\infty$ ist nicht abhängig von $P^*$, sondern nur von $\boldsymbol{N}$. Man kann also die Normale $\boldsymbol{N}$ der $\mathcal{K}$-Ebene im Objektraum direkt aus dem Parametervektor $\mathbf{l}'_\infty$ ihrer Ferngeraden im Bild erhalten.

Im Übrigen, auch aus der Perspektive des Quadrats der $\mathcal{K}$ kann man die Ferngerade seiner Trägerebene direkt (aus den Fernpunkten gegenüber liegender Quadratseiten) herleiten.

## 7.2 *Die Polare zur Ellipsenmitte als Pol*

Der Mittelpunkt $\left[\boldsymbol{x}_0,1\right]^T = \left[-\boldsymbol{C}_{hh}^{-1}\boldsymbol{c}_{h0},1\right]$ der Ellipse $\boldsymbol{x}^T\boldsymbol{C}\boldsymbol{x} = 0$, $\mathbf{C} = \begin{bmatrix}\boldsymbol{C}_{hh} & \boldsymbol{c}_{h0}\\ \boldsymbol{c}_{0h} & c_{00}\end{bmatrix}$, wird in die Polargleichung $\mathbf{l} = \mathbf{C}\mathbf{x}$ eingesetzt und ergibt $[0\ 0\ C^*]^T \sim [0\ 0\ 1]^T = \mathbf{l}_\infty$, worin $C^* = c_{00} - \boldsymbol{c}_{h0}^T\boldsymbol{C}_{hh}^{-1}\boldsymbol{c}_{h0}$ eine skalare Konstante ist. Also: Die Polare des Mittelpunktes einer Ellipse ist die Ferngerade ihrer Ebene.

## 7.3 *Projektive Invarianz von Polaritäten*

Eine beliebige, reguläre Polarität $\mathbf{l} = \mathbf{C}\mathbf{x}$ mit Polare $\mathbf{l}$, Pol $\mathbf{x}$, Ellipsenmatrix $\mathbf{C} = \mathbf{C}^T$ wird einer regulären Homographie mit der 3 × 3 Matrix $\mathbf{H}$ unterworfen. Es ist zu prüfen, ob die Polare $\mathbf{l}$ in Bezug auf $\mathbf{C}$ abgebildet wird auf die Polare $\mathbf{l}'$ in Bezug auf $\mathbf{C}'$. Beliebige Punkte $\mathbf{x}$ bzw. Geraden $\mathbf{l}$ werden projektiv mit $\mathbf{H}$ transformiert in

$$\mathbf{x}' = \mathbf{H}\mathbf{x} \qquad \mathbf{l}' = \mathbf{H}^{-T}\mathbf{l}\,. \tag{24}$$

Ferner ergibt sich zunächst für die transformierte Ellipsenmatrix

$$\mathbf{C}' = \mathbf{H}^{-T}\mathbf{C}\mathbf{H}^{-1}. \tag{25}$$

Dies folgt mit Anhang 7.1 aus der Ellipsengleichung $\mathbf{x}'^T\mathbf{C}'\mathbf{x}' = 0$. Für die Polare im Bild erhalten wir $\mathbf{l}' = \mathbf{C}'\mathbf{x}' = \mathbf{H}^{-T}\mathbf{C}\mathbf{H}^{-1}\mathbf{H}\mathbf{x} = \mathbf{H}^{-T}\mathbf{C}\mathbf{x} = \mathbf{H}^{-T}\mathbf{l}$.

Dies ist die gesuchte Bestätigung. Die Pol – Polare – Relation in Bezug auf eine Ellipse ist also bei projektiven Transformationen invariant.

## 7.4 *$\mathcal{K}$-Bestimmung durch digitale Bildinversion mit gefaltetem Grauwertsprung*

Optische Bildfunktionen gehen aus zugehörigen Objektfunktionen durch die Perspektive hervor einschließlich einer Tiefpassfilterung (Faltungsintegral). Letztere ist charakterisiert durch das Punktbild oder die Punktstreufunktion (=PSF) der Kamera. Bereits mit der stark vereinfachten PSF als isotrope Rechteckfunktion ($2s$ = Basis, $1/(2s)$ = Höhe) wird aus dem Grauwertsprung zwischen $G_1$ und $G_2$ durch die Faltungsoperation eine lineare Funktion erzeugt. Die Breite des Übergangs und die Basis der PSF sind einander gleich. Die Parameter sind zu bestimmen aus den Pixel-Grauwerten der $\mathcal{K}$ in allen Bildern. Das Grauwertprofil $G(t)$ von der $\mathcal{K}$-Mitte bis an den $\mathcal{K}$-Rand lautet damit (Kraut & Wrobel 1993)

$G(t) = G_1$ für $0 \le t \le r - s$

$$G(t) = \frac{1}{2}(G_1 + G_2) - \frac{1}{2}(G_1 - G_2)\frac{t-r}{s}$$

für $r - s \le t \le r + s$

$G(t) = G_2$ für $r + s < t$

Die vier Parameter der Helligkeitsfunktion der $\mathcal{K}$ sind $G_1$, $G_2$, $\mathcal{K}$-Radius $r$ und $s$ für die PSF.

Es gilt ferner für (20): $dG/d\boldsymbol{x} = (dG/dt)(dt/d\boldsymbol{x})$ mit den Polarkoordinaten $t$, $\varepsilon$ von $\boldsymbol{x} = [t\cos\varepsilon, t\sin\varepsilon]^T$.

Anschrift des Autors:

Prof. Dr.-Ing. Bernhard P. Wrobel, TU Darmstadt, zurzeit Gast bei der Professur Photogrammetrie, Prof. Dr. mult. W. Förstner, Universität Bonn. Privat: Im Saufang 10, D-53343 Wachtberg, e-mail: wrobel.bernhard@t-online.de

Manuskript eingereicht: Januar 2012
Angenommen: März 2012

PFG 2012 / 3, 0237–0250
Stuttgart, June 2012

Article

# Calibration and Direct Georeferencing Analysis of a Multi-Sensor System for Cultural Heritage Recording

David Hernández-López, Miriam Cabrelles, Beatriz Felipe-García & José Luis Lerma, Albacete, Valencia, Spain

**Keywords:** Multi-sensor system, calibration, orientation, cultural heritage

**Summary:** A large amount of cultural heritage monuments and sites exist distributed worldwide that require easy to use, cheap and fast sensor orientation tools for recording, georeferencing, surveying and mapping. Ground control points are usually required on site for close range photogrammetry to achieve accurate surveys, limiting both the involvement of non-experts and the chance to know the right place of the monuments unless a ground reference system is considered. This paper presents the system calibration of an image-based multi-sensor system that integrates two consumer-grade cameras, one global navigation satellite system (GNSS) and one low-cost inertial system, i.e., a micro-electro-mechanical system (MEMS) based inertial measurement unit (IMU). The multi-sensor system is calibrated indoor regarding both the camera orientation parameters and the boresight (rotations and offsets) parameters of the two digital cameras. The boresight parameters will be used to correct the direct approach estimates. The performance of the system calibration is tested outdoor on an upside down pyramidal sculpture to deliver both accurate 3D points and high resolution 3D models. Both scenarios are considered free of magnetic anomalies. The results achieved with the GNSS/MEMS-IMU direct approach are compared with the indirect approach based on bundle block adjustment. Further extrapolations to object space through digital surface models are also determined. The testing of the system shows that GNSS/MEMS-IMU data are good enough to provide approximate exterior orientation parameters of the cameras but not to yield accurate 3D models (<1–2 cm) for cultural heritage applications.

**Zusammenfassung:** *Kalibrierung und Analyse der direkten Georeferenzierung eines Multisensorsystems zur Aufnahme von Kulturdenkmälern*. Weltweit gibt es eine große Zahl an Denkmälern, die zum Kulturerbe zählen und für deren Aufnahme, Vermessung und Kartierung einfach handhabbare, preiswerte und schnelle Sensor-Orientierungswerkzeuge benötigt werden. Für eine genaue Vermessung durch Nahbereichsphotogrammetrie werden vor Ort Passpunkte benötigt, was die Einbeziehung von Nichtexperten, aber auch die absolute Positionierung eines Denkmals ohne zusätzliche Information über das Referenzsystem behindert. Dieser Beitrag präsentiert die Kalibrierung eines Multisensorsystems, welches zwei übliche digitale Kameras, einen Empfänger für ein globales Satelliten-Navigationssystem (GNSS) und eine kostengünstige inertiale Messeinheit (IMU) auf Basis von mikroelektromechanischen Systemen (MEMS) integriert. Dieses Multisensorsystem wird im Labor sowohl hinsichtlich der Orientierungsparameter der Kameras als auch hinsichtlich deren Verschiebung und Verdrehung relativ zu den Positions- und Lagesensoren kalibriert. Die kalibrierten Werte dieser Verschiebungen und Verdrehungen werden bei der direkten Georeferenzierung zur Korrektur der durch diese Sensoren bestimmten Position und Lage des Systems verwendet. Die Güte dieser Systemkalibrierung zur Gewinnung von genauen 3D Punkten und hochaufgelösten 3D Modellen wird außerhalb des Labors anhand einer auf dem Kopf stehenden pyramidenförmigen Skulptur untersucht. Beide Szenarien werden als frei von magnetischen Anomalien angesehen. Die Ergebnisse der direkten Georeferenzierung basierend auf den GNSS/MEMS-IMU Sensoren werden mit einer indirekten Methode unter Verwendung einer Bündelblockausgleichung verglichen und Extrapolationen im Objektraum durch digitale Oberflächenmodelle bestimmt. Die Untersuchung des Systems zeigt, dass die Daten der GNSS/MEMS-IMU Sensoren gut genug sind, um genäherte Parameter für die äußere Orientierung der Kameras zu liefern, aber noch nicht dafür ausreichen, sehr genaue 3D Modelle (< 1–2 cm) für die Aufnahme von Kulturdenkmälern zu erreichen.


www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0114
1432-8364/12/0114 $ 3.50

## 1 Introduction

There are many cultural heritage architectural and archaeological objects, monuments and sites around the world that need appropriate recording. Accurate recording is mandatory for decay analysis, reconstruction, and monitoring over time. Close range photogrammetry has long been used to record cultural heritage. The classical approach was making use of metric cameras, acquiring stereo-pairs and manual plotting. With the advent of digital photogrammetry, two-dimensional rectifications and ortho image generation became popular in architectural recording. In the last years, there is a trend to automate image-based solutions either with metric cameras or with amateur cameras based on robust hierarchical detection and matching of image features (Haoa & Mayer 2003, Pollefeys et al. 2004, Remondino & Ressl 2006). The determination of the exterior orientation parameters is a conventional requirement whenever dealing with image-based photogrammetric datasets. An ideal scenario would be to achieve them quickly and accurately without any need of direct measurements based on total stations, GNSS and rulers, among others. However, direct georeferencing in close range photogrammetry is not widespread despite of its benefits to avoid measuring coordinated targets (Kirchhöfer et al. 2010, 2011).

Direct georeferencing using integrated relative kinematic GNSS/inertial systems has been investigated for more than one decade to determine the reliability and accuracy of directly measured orientation parameters in operational photogrammetric airborne environments (Cramer et al. 2000, Mostafa & Schwarz 2001) and terrestrial mobile mapping systems (El-Sheimy & Schwarz 1999, Da Silva et al. 2003). Skaloud (2006) reviews the essential features, methods and approaches in direct georeferencing of airborne sensors. There is no doubt that highly sophisticated integrated GNSS/inertial systems can be used to overcome traditional issues in aerial triangulation such as block design, determination of approximate exterior orientation parameters, reduction of interactive editing, more reliable feature matching and truly automatic processing. Nevertheless, GNSS/inertial data integration and uncorrected errors in the overall system calibration including not only the boresight misalignment and the GNSS offsets but the imaging sensors are the limiting factors to achieve high accuracy estimates in object space (Cramer 2001). The most crucial task is to select the right calibration parameters for the system calibration (Cramer & Stallmann 2002). An analysis of the stability of the misalignments is required to avoid significant changes in the rotations, GNSS offsets and camera calibration parameters (Jacobsen 2000). Furthermore, stereoplotting as well as some automatic image matching methods for digital elevation model (DEM) generation may be affected by y parallaxes (Yastikli & Jacobsen 2002, 2005).

There is an increased interest in the topic of low-cost inertial navigation especially for pedestrian positioning systems (Feliz et al. 2009, Chen et al. 2009) and human motion tracking (Frank 2010, Sabatini 2011). Standalone MEMS trajectory determination is an actual challenge despite of the recent improvements in the performance of small and lightweight systems (Woodman 2007). Compared to high-end tactical and navigation IMUs, with MEMS-IMU it is difficult to determine an accurate heading due to the drift of low-cost MEMS gyroscopes and unpredictable perturbance of the magnetic field (Chen et al. 2009). Recent scientific publications are tacking the latter effect mitigating the magnetic anomalies in environments such as urban canyons and indoor (Afzal et al. 2011a, b, Tome & Yalak, 2008).

Low-cost multi-sensor systems for positioning can integrate a diverse set of devices such as a laptop, a consumer GNSS, a digital compass, a video camera (Haala & Böhm 2003), a MEMS inertial sensor, a webcam and a display system (Portalés et al. 2010), a personal digital assistant (PDA) or a smartphone with integrated camera and MEMS based orientation and positioning sensors. The accuracy requirements for cultural heritage applications are nowadays neither fulfilled with consumer GNSS nor smartphone devices. The limiting factor is the theoretical positioning accuracy of the code-based GNSS, although its absolute accuracy can be improved after the system calibration due to the partial compensation of

the positional errors (Kirchhöfer et al. 2011). On the contrary, MEMS-IMU results integrated with differential GNSS (or better DGNSS) are promising for navigation-type applications such as mobile mapping systems (Schwarz & El-Sheimy 2004, García-Asenjo et al. 2008, Kirchhöfer et al. 2010).

This paper aims at an analysis of the performance of the overall system calibration of an image-based multi-sensor system for architectural/archaeological recording under scenarios without disturbances of the magnetic field. The GNSS will be used to estimate the position, while the low-cost MEMS-IMU will be used for orientation determination. The data acquisition will follow a stop-and-go strategy. This system might be used to survey other scenarios such as outcrops in geomorphology, large civil engineering structures and 3D city models namely in opened areas, among others. In section 2 the geometric relationship between the GNSS, IMU and the two cameras is briefly reviewed and particularized for the image-based multi-sensor system with two consumer-grade cameras. In section 3 the overall system calibration results are described. In section 4 the performance of the direct georeferencing approach with GNSS/MEMS-IMU and the indirect approach with bundle block adjustment when modelling a large archaeological sculpture is analysed and discussed. A conclusion of the research carried out for cultural heritage recording is presented in section 5.

## 2 Geometric Relationship of the Sensors

The direct georeferencing model sets up the geometric relationship of a multi-sensor system, i.e., digital camera (*c*-frame), IMU body frame (*b*-frame) and the GNSS antenna in the ground coordinate system (*m*-frame). The relationship among the different sensors can be written as follows (1):

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}_m = \begin{bmatrix} X_{GNSS} \\ Y_{GNSS} \\ Z_{GNSS} \end{bmatrix}_m - R_b^m \begin{bmatrix} ax \\ ay \\ az \end{bmatrix}_b^{GNSS} + R_b^m \begin{bmatrix} dx \\ dy \\ dz \end{bmatrix}_c^b + \alpha R_b^m dR_c^b \begin{bmatrix} x \\ y \\ -f \end{bmatrix}_c \quad (1)$$

where $X$, $Y$, $Z$ are ground coordinates, $X_{GNSS}$, $Y_{GNSS}$, $Z_{GNSS}$ are ground coordinates of the centre of phase of the GNSS antenna, $ax$, $ay$, $az$ the offsets of the centre of phase of the GNSS antenna relative to the IMU centre (*b*-frame origin), $R_b^m$ is the rotation matrix of the *b*-frame into the *m*-frame; $dx$, $dy$, $dz$ are the offsets of the projection centre relative to the IMU centre; $\alpha$ is a scale factor; $dR_c^b$ is the rotation matrix of boresight misalignment; $x$, $y$ are the image coordinates; and finally $f$ the principal distance.

Our multi-sensor system integrates two digital cameras, one virtual reference station (VRS)-GNSS rover receiver, one low-cost MEMS-IMU, and a laptop for operating, synchronizing and saving the data coming from the sensors. All the sensors are attached to a portable pitch and yaw rotating mount (Fig. 1a). This portable two axis multi-sensor rotating mount is an evolution of a previously designed 1 axis rotating mount (Lerma et al. 2010) that allows free image data acquisition without losing the satellite signals when shooting upwards or downwards the horizontal plane. Thus, this mount minimises the problem of antenna tilt and loss of signal or multipath effects when large pitch angles are required for recording.

The devices selected for the project presented herein are two single lens reflex (SLR) digital cameras, one Canon EOS 1Ds Mark III (21.9 MPixels) with a Canon EF 24 mm F2.8 and another low resolution camera Canon EOS D60 (6.3 MPixels) with a Sigma 15–30 mm F3.5-4.5 EX DG Aspherical (fixed at 15 mm and focused at infinity). A summary of the camera specifications is presented in Tab. 1. An Xsens MTx inertial system (Tab. 2) is used to provide the attitude information of the multi-sensor system. From the specifications, it is understood that the MEMS-IMU

(a) (b)

**Fig. 1:** (a) Close-up view of the image-based multi-sensor system, (b) outdoor testfield.

**Tab. 1:** Camera specifications.

| Camera | Canon EOS 1Ds Mark III | Canon EOS D60 |
|---|---|---|
| Resolution (pixel) | 5616 x 3744 | 3072 x 2048 |
| Sensor size (mm) | 36 x 24 | 22.7 x 15.1 |
| Pixel size (µm) | 6.4 | 7.4 |
| Sensor size | Full frame (36 mm x 24 mm) | APS-C (22.7 mm x 15.1 mm) |
| Focal length (mm) | 24 | 15 (24 full frame equivalent) |
| Image quality | RAW | RAW |
| Lens | Canon EF 24 mm F2.8 | Sigma 15–30 mm F3.5–4.5 EX DG |

**Tab. 2:** Xsens MTx specifications.

| | |
|---|---|
| Angular Resolution | 0.05 deg |
| Static Accuracy (Roll/Pitch) | < 0.5 deg |
| Static Accuracy (Heading) | < 1 deg |
| Dynamic Accuracy | 2 deg RMS |
| Maximum update rate onboard | 120 Hz |

static accuracy in roll and pitch is <0.5 deg and in yaw <1 deg (in homogeneous magnetic environment). A proprietary Xsens MTx sensor fusion algorithm to combine magnetometer and gyroscope data was used to orient the device relative to the global reference frame. The inertial sensor is placed inside the housing of the portable mount nearby the centre of rotation; the Trimble Zephyr antenna is fitted on top of the mount. A Trimble 5700 GPS receiver in VRS mode is used to determine the coordinates of the antenna phase centre. The expected positioning accuracy is approximately 2–3 cm (Retscher 2002). The accuracy depends on the number of collected data between the rover and the receiver to fix signal ambiguity with integers. When the ambiguity can not be fixed with integers, a float solution can be calculated that degrades the solution to a decimetre level (El-Rabbany 2002). Fig. 1 shows the image-based multi-sensor system at the Polytechnic University of Valencia testfield. The image-based multi-sensor system presented herein differs from those published by Lerma et al. (2010) and Kirchhöfer et al. (2011). The former only has got a free yaw-angle rotation while the latter integrates only one digital SLR camera.

The idea of integrating two cameras on the mount is five-fold: first, to improve reliability in the data acquisition (the simultaneous acquisition of two images avoids troubles in case of unexpected camera errors on site); second, it allows multiband image acquisition when requested, for instance, for building inspections with two cameras, one visible and

another thermal (such as in the solution presented by Alba et al. 2011 and Lerma et al. 2010); third, it facilitates image fusion due to the rigid body transformation between cameras; fourth, it eases the comparison and the extrapolation of output results; and fifth, the multi-sensor system might work as a stereoscopic system with an appropriate base and camera-object distance ratio. In this paper, the cameras are not supposed to work as stereopairs from each station but as a multi-acquisition system acquiring data from multiple stations.

## 3 Overall System Calibration

The overall system calibration is required to relate GNSS-derived positions, IMU-derived attitude parameters and imagery-derived attitude parameters. Fig. 2 shows the frame of the multi-sensor system without the GNSS-antenna and cameras; only the MEMS-IMU is inside the metal frame. Next, the indoor approach for calibrating the system is presented.

The system calibration was undertaken in a static close range fashion following three main steps:

1. camera calibration of the two cameras,
2. determination of the origin of the IMU placed inside the mount and some auxiliary marks, and
3. boresight calibration among the sensors.

For the first step, two self-calibration bundle adjustments were independently carried out, one for each camera, to determine the interior orientation parameters (for details, see Cabrelles 2010). To achieve high precision in the camera calibration parameters, an optimised object space full of well-distributed targets and a convergent image space configuration was conducted to strengthen the overall network geometry and the camera station configuration (Remondino & Fraser 2006). The indoor testfield full of coded targets used to calibrate the multi-sensor system is presented in Fig. 2.

The familiar eight-parameter model with principal distance, principal offset and corrections for radial and decentring distortion was selected to estimate the interior orientation parameters. The adjustment included 9 images for the Canon EOS 1Ds Mark III and 8 for the Canon EOS D60. The output interior orientation parameters for the two cameras are presented in Tab. 3. All the parameters were significant for the former camera; for the latter, radial distortion parameter K3 and decentring distortion parameter P1 proved insignificant and were removed by the self-calibration approach. The standard deviations were always better by one order of magnitude for the 1Ds Mark III than for the D60.

It is essential to model radial lens distortion to effectively achieve high accuracy, especially for consumer-grade digital cameras (Chan-

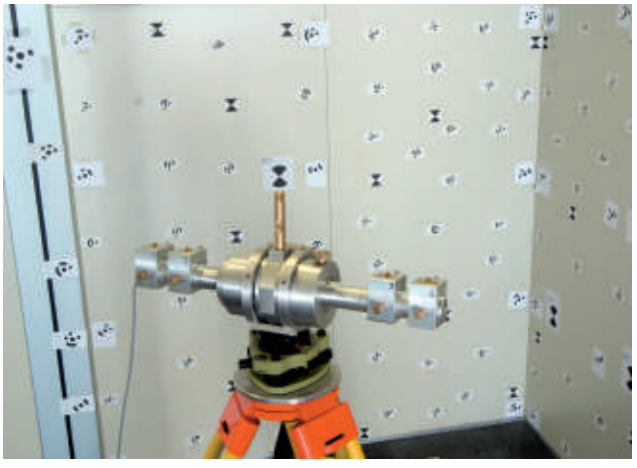


**Fig. 2:** Indoor testfield used to calibrate the multi-sensor system.

**Tab. 3:** Interior orientation parameters achieved for the two cameras and related standard deviations.

| Parameters | Canon EOS 1Ds Mark III | | Canon EOS D60 | |
|---|---|---|---|---|
| | Value | Standard deviation | Value | Standard deviation |
| f (pixel) | −3819.93 | 0.03 | −2099.53 | 0.19 |
| $x_0$ (pixel) | −26.47 | 0.03 | 12.26 | 0.12 |
| $y_0$ (pixel) | 25.53 | 0.04 | −17.79 | 0.31 |
| $K_1$ (pixel$^{-2}$) | 7.57E−09 | 5.29E−12 | 1.99E−08 | 8.30E−11 |
| $K_2$ (pixel$^{-4}$) | −4.60E−16 | 1.03E−18 | −1.98E−15 | 2.64E−17 |
| $K_3$ (pixel$^{-6}$) | −3.42E−24 | 6.10E−26 | | |
| $P_1$ (pixel$^{-1}$) | 6.33E−08 | 6.63E−10 | | |
| $P_2$ (pixel$^{-1}$) | 3.15E−08 | 8.87E−10 | 2.01E−07 | 2.15E−08 |

DLER et al. 2005, WACKROW et al. 2007, WACKROW & CHANDLER 2008).

The second step of the overall system calibration started with the determination of the origin of the IMU that is placed inside the mount. The multi-sensor system was dismantled and only the central metallic bar hosted the attached IMU as well as the four screws used to attach the two camera mounts. The four screws were used as auxiliary marks to relate the origin of the IMU and the position of the camera centres once the mount is assembled. The bar was photographed from multiple images and the marks measured on the images. A bundle block adjustment was undertaken to determine the spatial offsets between the origin of the IMU *b*-frame and the centres of the four screws (screws 1, 2 for the Canon EOS 1Ds Mark III and screws 3, 4 for the Canon EOS D60, Fig. 3).

The third and last step in the overall system calibration was the determination of the boresight parameters between the sensors: the IMU, the two cameras and the GNSS antenna. For this third step, the assembled multi-sensor system mount was carefully levelled and oriented parallel to the object space coordinate system to transform the raw IMU angles into the *m*-frame. The orientation parallel to the object space coordinates was required to transform the raw IMU κ reading (which refers to a global coordinate frame) to its equivalent in the local coordinate system that was used for calibration. Several images were taken with the Canon EOS 1Ds Mark III to determine: a) the centre of the GNSS screw; b) the centres of the four screws. Afterwards, the two SLR cameras were attached to the multi-sensor system mount and two images were taken, one with each camera, to determine the perspective centres and the rotations. The centre of the GNSS screw was modified to account for the shift between the base of the GNSS antenna and its centre of phase. This latter value was provided by the GNSS manufacturer and added to compute the GNSS offset between the GNSS antenna and the IMU *b*-frame, i.e., *ax*, *ay* and *az* in (1). The centres of the screws in b) were used as auxiliary marks to determine the camera-IMU offset parameters *dx*, *dy* and *dz* (1) owing to the position of the MEMS-IMU inside the chase. This was possible because the system was not moved when the cameras were attached to their mounts, so that the positions of the screws in the object coordinate system did not change in this process. A final bundle adjustment with all the images acquired in this step was carried out to determine the boresight parameters. The differences between the IMU rotation angles and the bundle adjustment orientation angles yielded the boresight misalignment $(dR_c^b)$. Tabs. 4 and 5 summarise the offsets and the boresight misalignments of the multi-sensor system, respectively, considering the IMU *b*-frame as the origin, Fig. 3. FOTOGIFLE photogrammetric software (developed in-house by the authors) was used to determine the estimates of the overall system calibration.

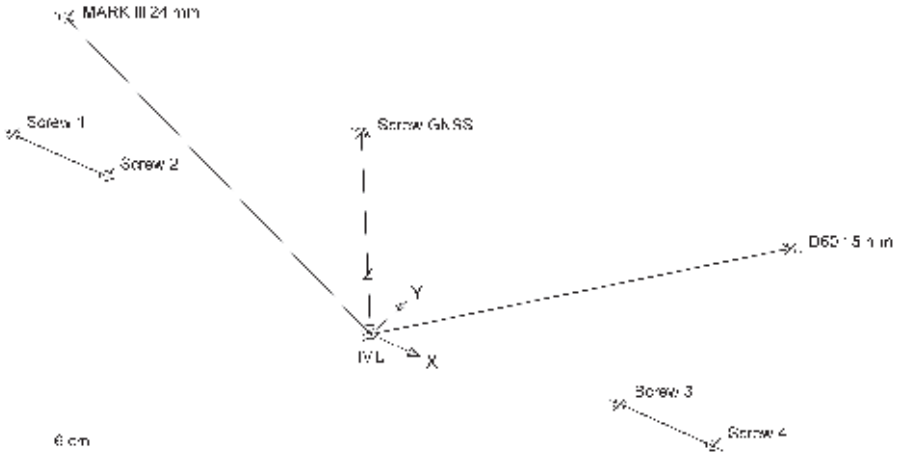



**Fig. 3:** Offsets among the camera perspective centres (Canon EOS 1Ds Mark III and Canon EOS D60), the centre of the screw of the GNSS antenna and the origin of the IMU; screws 1 up to 4 are used as auxiliary marks for the two cameras.

**Tab. 4:** Offsets and standard deviations after the overall system calibration.

| Point | $X$ (mm) | $Y$ (mm) | $Z$ (mm) | $\sigma_X$ (mm) | $\sigma_Y$ (mm) | $\sigma_Z$ (mm) |
|---|---|---|---|---|---|---|
| IMU | 0.0 | 0.0 | 0.0 | < 0.1 | 0.1 | < 0.1 |
| Centre 1Ds Mark III | −199.1 | 11.6 | 114.1 | 0.1 | 0.1 | < 0.1 |
| Centre D60 | 193.9 | 109.4 | 71.9 | 0.4 | 0.1 | 0.3 |
| Screw GNSS | −3.4 | 4.3 | 131.9 | < 0.1 | 0.1 | 0.1 |

**Tab. 5:** Boresight misalignments and standard deviations after the overall system calibration.

| Camera | $\Delta\omega$ (g) | $\Delta\varphi$ (g) | $\Delta\kappa$ (g) | $\sigma_\omega$ (g) | $\sigma_\varphi$ (g) | $\sigma_\kappa$ (g) |
|---|---|---|---|---|---|---|
| Canon EOS 1Ds Mark III | 1.3221 | 0.8161 | −0.3252 | 0.0012 | 0.0012 | 0.0016 |
| Canon EOS D60 | −0.0952 | 0.6256 | −1.0179 | 0.0074 | 0.0023 | 0.0093 |

As it can be checked in Tab. 4, the results of the overall system calibration on the indoor testfield are better than 0.2 mm (1 $\sigma$); only the perspective centre of the Canon EOS D60 yields standard deviation up to 0.4 mm in the X axis and 0.3 mm in the Z axis. Tab. 5 summarises the boresight misalignments between the bundle adjustment and the transformed IMU attitudes into ω, φ and κ. The corresponding standard deviations determined in the bundle adjustment solution are up to 6.2, 1.9 and 5.8 times larger for the Canon EOS D60.

## 4 Performance of Direct Georeferencing

The performance of direct georeferencing has been investigated on an upside-down pyramidal sculpture on the campus of the Polytechnic University of Valencia (Fig. 1b). This sculpture is considered as an excellent testfield due to its complex geometry full of texture for the image measurements. A total of suitable 36 coded-target ground control points were surveyed by a Topcon IS Total Station; the accuracy of signalised control points is better than 0.005 m. Four ground control

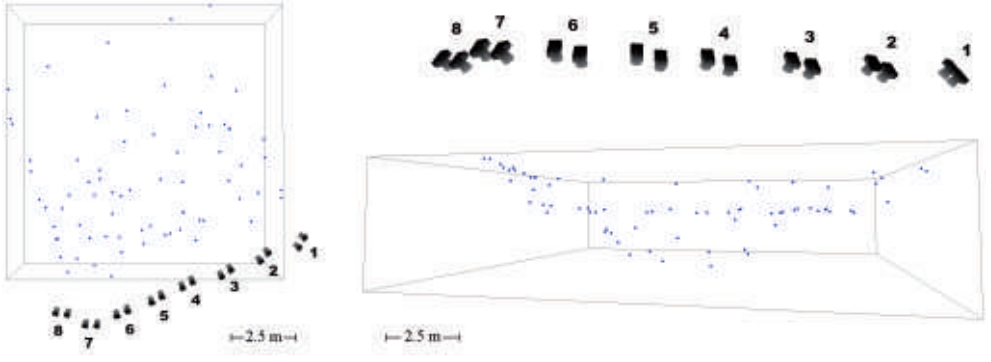
points nearby the sculpture were used before and after the measurements to set up the relationship between the local coordinate system and the european terrestrial reference system 1989 (ETRS89). A systematic shift of −0.025 m, −0.044 m and −0.120 m was found in the VRS-GNSS coordinates making use of the ERVA network (ERVA 2012).

As reported by Wackrow & Chandler (2008), a mildly convergent camera configuration can be used to minimise eventual systematic error surfaces caused by slightly inaccurate lens distortion parameters for DEM generation. Herein, a stop-and-go sequential acquisition mode at intervals of roughly 60 s at 8 stations following a strip with convergent axes at both extremes was carried out to verify the quality of the overall system calibration. Fig. 4 displays the 8 stations represented by the 16 camera positions.

An indirect approach was used to derive the exterior orientation parameters for each recorded image in a bundle adjustment integrating both cameras. A maximum root-mean-square error (RMSE) of 3 mm was achieved in object space on the ground control points. All the coded-target measurements in image-space were measured automatically with subpixel accuracy in PhotoModeler 6.0. The tie points between consecutive images corresponding to the two cameras were measured manually. All the points used in the bundle block adjustment (BBA) together with the exterior orientation of the cameras are displayed in Fig. 4.

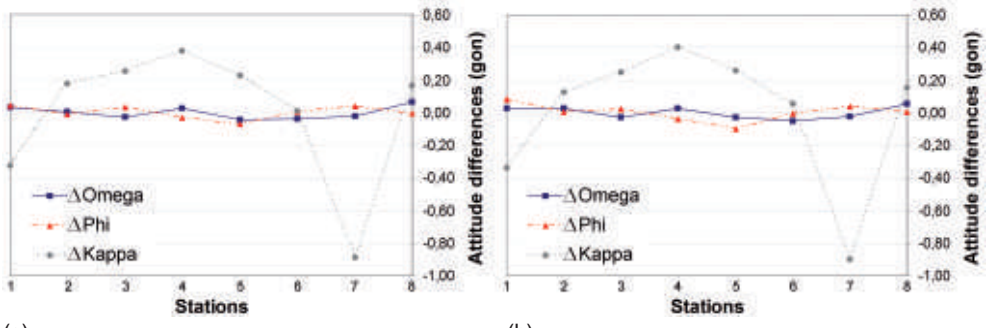
To assess the differences of the direct georeferencing approach and the photogrammetric image-based approach from multiple stations, the raw roll, pitch and yaw angles determined by the MEMS-IMU were converted into ω, φ and κ. The magnetic declination was considered to refer the κ angle to the geodetic north (García-Asenjo et al. 2008). The rotation matrix was corrected with the boresight misalignment parameters before transforming with two rigid body transformations (one for each camera) the GNSS coordinates measured at the antenna phase centre into the theoretical projection centres of the cameras. The mathematical routines were developed by the authors. In Figs. 5 and 6 the particular attitude and position differences after the overall system calibration carried out indoors are shown for the distinct camera stations taken on site around the sculpture. In other words, Figs. 5 and 6 present the residuals (differences) of the boresight misalignments and the GNSS offsets for the two cameras, respectively.

The differences of the boresight misalignments (Fig. 5) are quite consistent with zero mean for the three rotation angles in both cameras. The RMSEs for the Canon EOS 1Ds Mark III are 0.0136 gon in ω and φ, and more than one order of magnitude higher in κ, 0.1470 gon; for the Canon EOS D60 slightly worse, 0.0206 gon in ω, 0.0187 gon in φ and 0.1496 gon in κ. In both cameras, the instability of κ compared with ω and φ is apparent. It should be noted that the RMSEs for the three rotations are smaller than the RMSEs



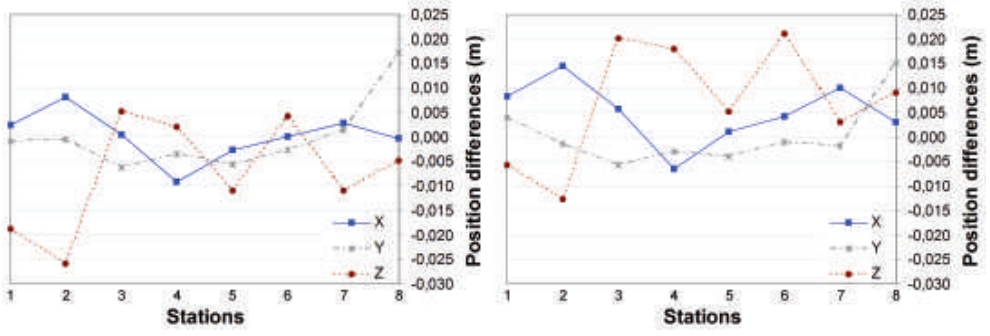

(a) (b)

**Fig. 4:** Two perspective views of the strip 8 stop-and-go stations, (a) top view, (b) front view. The blue dots are the tie points and the ground control points measured for the BBA.




(a) (b)

**Fig. 5:** Comparison of attitudes from reference bundle block adjustment and MEMS-IMU on site: (a) Canon EOS 1Ds Mark III, (b) Canon EOS D60.




(a) (b)

**Fig. 6:** Comparison of projection centres from reference bundle block adjustment and GNSS on site: (a) Canon EOS 1Ds Mark III, (b) Canon EOS D60.

of the attitudes provided by the manufacturer (Tab. 2).

Analysing the differences of the GNSS offsets after calibration presented in Fig. 6, the results for the Canon EOS 1Ds Mark III yield a zero mean in planimetry, and −7 mm in Z. However, for the Canon EOS D60, there is a mean value for the offsets of 5 mm in X, 0 mm in Y and 7 mm in Z, with RMSE values of 2 mm in X, 2 mm in Y and 4 mm in Z. The RMSEs for the Canon EOS 1Ds Mark III are 2 mm, 3 mm and 4 mm in X, Y and Z, respectively. It is worth mentioning that the departure from the mean of the height in both cameras points in opposite directions. To a lesser extent, the X-axis is also pointing in opposite directions due to the lack of accuracy of the angle κ of the MEMS-IMU. The achieved positioning accuracy is within the limits of the GNSS.

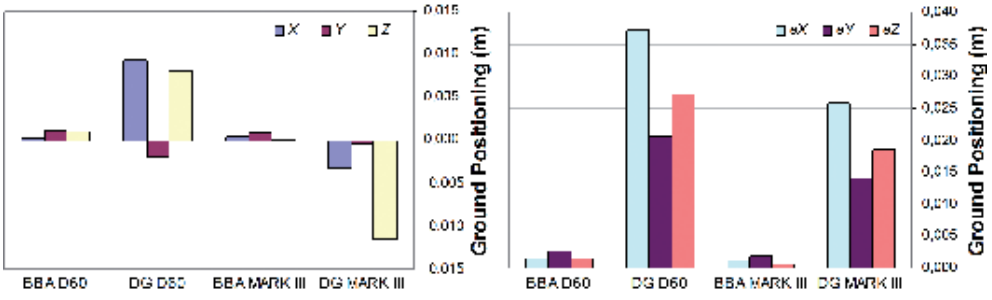
Based on the exterior orientation determined by the direct and indirect approaches, object coordinates of measured image points were computed by combined intersection independently for the two cameras. The combined spatial intersection corresponds to a bundle adjustment without control points, using the exterior orientation as fixed values and all available observations in the different images for the ground point measurement (Yastikli & Jacobsen 2005).

The results presented in Fig. 7a show the mean differences in X, Y and Z considering as ground truth the six check points measured by the image total station for both the direct and the indirect approaches with the two cameras independently; Fig. 7b concentrates on the corresponding RMSEs. On the one hand, the quality of the indirect approach with bundle adjustment is similar for both cameras despite

their specifications (Tab. 1). The mean differences in position are in the range of 1 mm and 2 mm, and RMSEs up to 3 mm in Y for the Canon EOS D60. On the other hand, the direct approach yields both larger differences and larger RMSEs. The mean differences are approximately 1 cm (Fig. 7a). However, the RMSEs up to 4 cm are larger for the Canon D60 (approx. 1.4 times for the three components). Without any doubt, the uncertainty of the VRS-GNSS in positioning (with RMSEs not better than 1–2 cm) is a limiting factor to improve the quality of the output coordinates. A review of the accuracy performance of VRS networks can be found in Retscher (2002).

As expressed in Jacobsen (2000), the absolute accuracy is only one result. The relative accuracy represented by the y parallax is important for the model setup, especially for the stereoscopic view. In addition, it might affect the performance of image matching. Tab. 6 presents the y parallax achieved after bundle block adjustment and direct georeferencing. The difference in quality is significant, below 1/3 of a pixel for bundle adjustment with maximum parallaxes below 1.5 pixels and higher for the direct referencing approach. The RMSE of y parallax is in the range of 4 pixels for the Canon EOS D60 and almost 8 pixels for the Canon EOS 1Ds Mark III are both unacceptable for stereo-plotting; the maximum y parallaxes go up to 19.7 and 34.2 pixels, respectively. The larger errors in the Canon EOS 1Ds Mark III digital camera are owing to its higher resolution. In fact, the y parallax values for the Canon EOS 1Ds Mark III are slightly below the height resolution ratio (that equals 1.82).

Regarding derived XYZ coordinates in object space, despite of the mean differences (Fig. 7a) being within the accepted tolerance in large architectural documentation projects with typical mapping scales of 1:100 or 1:200, the large RMSEs can yield computations with large deviation errors (up to 10 cm). This fact is clearly visible when extrapolating the exterior orientation parameters to dense image matching, e.g. for 3D modelling. A statistical analysis shows the metric differences on the digital surface models for the two different solutions, direct referencing and indirect referencing (Fig. 8). 43.2% of the digital surface model is in the range of 0 and −2 cm, with maximum differences up to 2 cm (24.9%) and −10 cm (29.3%). These results coming from direct referencing allow us to confirm that the




(a) (b)

**Fig. 7:** (a) Mean differences and (b) RMSEs at the check points considering blocks with 8 images after bundle block adjustment (BBA) and direct georeferencing (DG).

**Tab. 6:** RMS y parallax errors of the blocks.

| Approach | Number of images | Number of check points | RMSE of y parallax (pixel) | Max. y parallax (pixel) |
|---|---|---|---|---|
| BBA Canon EOS D60 | 8 | 55 | 0.3 | 1.5 |
| DG Canon EOS D60 | 8 | 55 | 4.3 | 19.7 |
| BBA Canon EOS 1Ds Mark III | 8 | 61 | 0.3 | 1.3 |
| DG Canon EOS 1Ds Mark III | 8 | 61 | 7.6 | 34.2 |

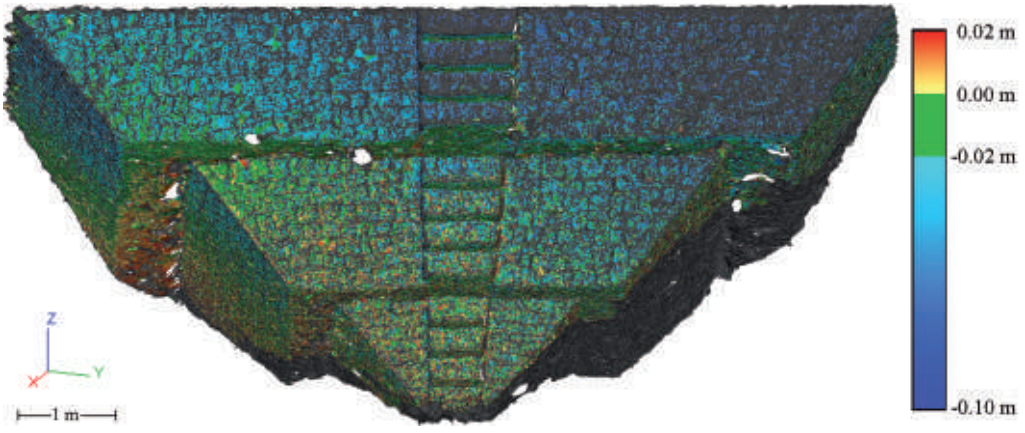



**Fig. 8:** Comparison of the digital surface models coming from direct and indirect georeferencing.

estimates are good as initial estimation for the exterior orientation problem and visualisation but not good enough to achieve accurate 3D models in close range applications following the presented methodology.

Three main factors affect the estimates of both relative and absolute accuracy with direct georeferencing: first, the influence of a VRS-GNSS systematic shift namely in the Z axis; second, the IMU rotation errors namely in κ; and third, the inherent errors due to the overall system calibration carried out indoor. Regarding the relative accuracy, the non-zero mean difference when positioning with the direct approach (particularly in Z upwards for one camera and downwards for the other, Fig. 6) affects directly the quality of the output results, not only in object space (Fig. 7) but also in the image space (Tab. 5). In addition to the previous statement, the rotation estimates determined by the MEMS-IMU with their corresponding errors also affect the outputs in object space, namely the XY due to the κ error (Fig. 7) and slightly in Z due to ω and φ rotation errors. From a photogrammetric point of view, one way to improve the overall accuracy will be to carry out a combined block adjustment (also known as integrated sensor orientation). Alternatively, a combination of global direct georeferencing and local indirect orientation approaches can also be considered as an optimised way to perform georeferencing on monuments and sites. The alternative approaches dealing with global and local orientation approaches to improve the reliability of the georeferenced data will be presented in future papers. Nevertheless, the results presented in this paper yield better accuracy estimates than other recording solutions (Da Silva et al. 2003, García-Asenjo et al. 2008, Kirchhöfer et al. 2010, 2011), even without combined block adjustment. More tests are necessary to confirm the reliability of the presented recording system for cultural heritage documentation.

## 5 Conclusions

This paper addresses the indoor overall calibration of an image-based multi-sensor system that integrates two SLR digital cameras, a low cost MEMS-IMU and a VRS-GNSS. The overall calibration includes the determination of both the interior orientation parameters and the boresight parameters. Once calibrated, the system is ready to carry out photogrammetric surveys outdoor despite of its relative accuracy is not appropriate for stereoscopic viewing with VRS-GNSS/MEMS-IMU direct georeferencing. The results presented in this paper demonstrate that mean differences of 1 cm and RMSEs in the range of 2–3.5 cm can be achieved with an image-based multi-sensor system performing direct georeferencing both in orientation and in spatial intersection. The multi-sensor system can be used to integrate different imaging devices, for instance,

cameras with different focal lenses or cameras with different spectra such as visible and near infrared. Furthermore, the multi-sensor system can be used either as a conventional stereoscopic system from each station or as a stereoscopic solution from multiple stations based on direct georeferencing. Using cameras of similar specifications except resolution, a slight improvement in quality by a factor of 1.4 is possible when increasing the resolution of a SLR digital camera from 6.3 up to 21.9 MPixel. The influence of the overall system calibration is of paramount importance to achieve maximum accuracy in daily projects. The limiting factors in the error budget are the κ angle and the VRS-GNSS solution. The bundle adjustment approach delivers highly reliable results with strong network geometry and enough number of images. Nevertheless, the inner camera geometry has to be properly modeled. The presented results proved that the quality of direct exterior orientation measurements using VRS-GNSS/MEMS-IMU for medium accuracy (1–4 cm) heritage documentation projects is feasible; projects demanding more accuracy require different processing to improve the quality of the exterior orientation parameters. One way to improve the quality is to carry out a combined block adjustment (integrated sensor orientation) with at least one control point or preferably without ground control points.

Further research is required to confirm the feasibility of the presented image-based multi-sensor system in a variety of complex scenarios namely with GNSS shortages and with magnetic anomalies. For those circumstances, different indirect photogrammetric approaches will be used to bridge the gap to direct georeferencing.

## Acknowledgements

The authors would like to thank for the support provided by the Spanish Ministry of Science and Innovation to the project HAR2010-18620. Special thanks need to be expressed to Dr. Luis García-Asenjo and Mr. Pascual Garrigues from the Polytechnic University of Valencia for their contributions.

## References

Alba, M.I., Barazzetti, L., Scaioni, M., Rosina, E. & Previtali, M., 2011: Mapping Infrared Data on Terrestrial Laser Scanning 3D Models of Buildings. – Remote Sensing **3:** 1847–1870.

Afzal, M.H., Renaudin, V. & Lachapelle, G., 2011a: Multi-Magnetometer Based Perturbation Mitigation for Indoor Orientation Estimation. – Navigation **58** (4): 279–292.

Afzal, M.H., Renaudin, V. & Lachapelle, G., 2011b: Use of Earth's Magnetic Field for Mitigating Gyroscope Errors Regardless of Magnetic Perturbation. – Sensors **11** (12): 11390–11414.

Cabrelles, M., 2010: Calibración geométrica de cámaras en dispositivos métricos y parametrización de sus descentrados ópticos. – Diploma Thesis, Universitat Politècnica de València.

Chandler, J.H., Fryer, J.G. & Jack, A., 2005: Metric capabilities of low-cost digital cameras for close range surface measurement. – The Photogrammetric Record **20** (109): 12–26.

Chen, W., Fu, Z., Chen, R., Chen, Y., Andrei, O., Kroger, T. & Wang, J., 2009: An integrated GPS and multi-sensor pedestrian positioning system for 3D urban navigation. – 2009 Joint Urban Remote Sensing Event: 1–6, Shanghai.

Cramer, M., Stallmann, D. & Haala, N., 2000: Direct georeferencing using GPS/inertial exterior orientations for photogrammetric applications. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **33** (B3): 198–205.

Cramer, M., 2001: Performance of GPS/inertial solutions in photogrammetry. – Fritsch, D. & Spiller, R. (eds.): Photogrammetric Week **2001:** 49–62, Wichmann, Heidelberg, Germany.

Cramer, M. & Stallmann, D., 2002: System calibration for direct georeferencing. – The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **34** (3A): 79–84.

Da Silva, J.F.C., Camargo, P.D.O. & Gallis, R.B.A., 2003: Development of a low-cost mobile mapping system: a South American experience. – The Photogrammetric Record **18** (101): 5–26.

El-Rabbany, A., 2002: Introduction to GPS: the Global Positioning System. – Artech House, Boston, MA.

El-Sheimy, N. & Schwarz, K.P., 1999: Navigating urban areas by VISAT – A mobile mapping system integrating GPS/INS/digital cameras for GIS applications. – Navigation **45** (4): 275–285.

ERVA 2012: <http://icverva.icv.gva.es:8080> (2012-03-09).

Feliz, R., Zalama, E. & Gómez García-Bermejo, J., 2009: Pedestrian tracking using inertial sensors. – Journal of Physical Agents **3** (1): 35–42.

Frank, K., Vera Nadales, J. & Angermann, M., 2010: Reliable Real-Time Recognition of MotionRelated Human Activities Using MEMS Inertial Sensors. – ION GNSS 2010, Portland, Oregon, USA.

García-Asenjo, L., Lerma, J.L., Garrigues, P., Baselga, S., Cabrelles, M., Hernández, D., Buchón, F. & Navarro, S., 2008: Integración de GNSS y un sistema de navegación inercial de bajo coste para la georeferenciación directa de imágenes fotogramétricas. – International Congress on Geomatic & Surveying Engineering: 9 p., Valencia [on CD-ROM].

Haala, N. & Böhm, J., 2003: A multi-sensor system for positioning in urban environments. – ISPRS Journal of Photogrammetry & Remote Sensing **58:** 31–42.

Haoa, X. & Mayer, H., 2003: Orientation and autocalibration of image triplets and sequences. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **34** (3/W8): 73–78.

Jacobsen, K., 2000: Potential and limitation of direct sensor orientation. – The International Archives of the Photogrammetry and Remote Sensing **33** (B3/1): 429–435.

Kirchhöfer, M.K., Chandler, J.H. & Wackrow, R., 2010: Testing and application of a low-cost photogrammetric recording system suitable for cultural heritage recording. – Proceedings of RSPSoc and Irish Earth Observation Symposium: 330–337 [on CD-ROM].

Kirchhöfer, M.K., Chandler, J.H. & Wackrow, R., 2011: Cultural heritage recording utilising low-cost close-range photogrammetry. – Proceedings of CIPA 23rd International Symposium: 8 p., Prague, Czech Republic [on CD-ROM].

Lerma, J.L., Navarro, S., Cabrelles, M. & Seguí, A.E., 2010: Camera calibration with baseline distance constraints. – The Photogrammetric Record **25** (130): 140–158.

Mostafa, M.M.R. & Schwarz, K.P., 2001: Digital image georeferencing from a multiple camera system by GPS/INS. – ISPRS Journal of Photogrammetry & Remote Sensing **56:** 1–12.

Pollefeys, M., Van Gool, L., Vergauwen, M., Verbiest, F., Cornelis, K. & Tops, J., 2004: Visual Modeling with a Hand-Held Camera. – International Journal of Computer Vision **59** (3): 207–232.

Portalés, C., Lerma, J.L. & Navarro, S., 2010: Augmented reality and photogrammetry: A synergy to visualize physical and virtual city environments. – ISPRS Journal of Photogrammetry and Remote Sensing **65:** 134–142.

Remondino, F. & Fraser, C., 2006: Digital camera calibration methods: considerations and comparisons. – The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (5): 266–272.

Remondino, F. & Ressl, C., 2006: Overview and experiences in automated markerless image orientation. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **36** (3): 248–254.

Retscher, G., 2002: Accuracy Performance of Virtual Reference Station (VRS) Networks. – Journal of Global Positioning Systems **1** (1): 40–47.

Sabatini, A.M., 2011: Kalman-Filter-Based Orientation Determination Using Inertial/Magnetic Sensors: Observability Analysis and Performance Evaluation. – Sensors **11:** 9182–9206.

Schwarz, K.P. & El-Sheimy, N., 2004: Mobile mapping systems – state of the art and future trends. – The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **35** (B): 10 p.

Skaloud, J., 2006: Reliability in direct georeferencing: An overview of the current approaches and possibilities. – EuroSDR workshop EuroCOW on Calibration and Orientation, Castelldefels, Spain. isprs.org/commission1/euroCOW06/euroCOW06_files/papers/Eurocow06Skaloud.pdf (2011-11-08).

Tome, P. & Yalak, O., 2008: Improvement of Orientation Estimation in Pedestrian Navigation by Compensation of Magnetic Disturbances. – Navigation **55** (3): 179–190.

Wackrow, R., Chandler, J.H. & Bryan, P., 2007: Geometric consistency and stability of consumer-grade digital cameras for accurate spatial measurement. – The Photogrammetric Record **22** (118): 121–134.

Wackrow, R. & Chandler, J.H., 2008: A convergent image configuration for DEM extraction that minimises the systematic effects caused by an inaccurate lens model. – The Photogrammetric Record **23** (121): 6–18.

Woodman, O.J., 2007: An introduction to inertial navigation. – Computer Laboratory. University of Cambridge. Technical Report number 696. UCAM-CL-TR-696.

Yastikli, N. & Jacobsen, K., 2002: Investigation of direct sensor orientation for DEM generation. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **34** (1): 298–304.

Yastikli, N. & Jacobsen, K., 2005: Direct sensor orientation for large scale mapping–potential, problems, solutions. – The Photogrammetric Record **21** (111): 274–284.

Addresses of the Authors:

Asso. Prof. Dr.-Ing. David Hernández, Universidad de Castilla-La Mancha, Departamento de Ingeniería Geológica y Minera, E-02071 Albacete, Tel.: +34-96-7599200, Fax: +34-96-7599233, e-mail: david.hernandez@uclm.es

Ing. Miriam Cabrelles, Universitat de València, Departamento de Prehistoria y Arqueología, E-46022 Valencia, Tel.: +34-96-3877550, Fax: +34-96-3877559, e-mail: miriam.cabrelles@uv.es

Dr.-Ing. Beatriz Felipe, Universidad de Castilla-La Mancha, Instituto de Desarrollo Regional, E-02071 Albacete, Tel.: +34-96-7599200, Fax: +34-96-7599233, e-mail: beatriz.felipe@uclm.es

Asso. Prof. Dr.-Ing. José Luis Lerma, Universitat Politècnica de València, Department of Cartographic Engineering, Geodesy and Photogrammetry, E-46022 Valencia, Tel.: +34-96-3877550, Fax: +34-96-3877559, e-mail: jllerma@cgf.upv.es

Manuskript eingereicht: November 2011
Angenommen: März 2012

PFG 2012 / 3, 0251 – 0266
Stuttgart, June 2012

Article

# Assessment of Radiometric Correction Methods for ADS40 Imagery

Lauri Markelin, Eija Honkavaara, Masala, Finland, Daniel Schläpfer, Wil, Switzerland, Stéphane Bovet, Wabern, Switzerland & Ilkka Korpela, Helsinki, Finland

**Keywords:** Reflectance, aerial images, radiometric correction, vicarious calibration

**Summary:** This article presents the results of an assessment of radiometric correction methods of images taken by the large-format aerial, photogrammetric, multispectral pushbroom camera Leica Geosystems ADS40. The investigation was carried out in the context of the multi-site EuroSDR project "Radiometric aspects of digital photogrammetric images". Images were collected at the forestry research test site Hyytiälä, Finland, in August 2008. Two processing workflows were evaluated: one based on the photogrammetric software Leica XPro, which in radiometric processes relies on physical modelling and information collected from the imagery only, and one based on ATCOR-4, which is software dedicated to physical atmospheric correction of airborne multi-, hyperspectral and thermal scanner data, and can be operated either with or without in-situ reflectance and atmospheric observations. Outputs of these processes are reflectance images. Three participants processed the data with several processing options which resulted in a total of 12 different radiometrically corrected reflectance images. The data analysis was based on field and laboratory reflectance measurements of reference reflectance targets and field measurements of permanent targets (asphalt, grass, gravel). Leica XPro provided up to 5% reflectance accuracy without any ground reference and ATCOR-4 provided reflectance accuracy better than 5% with vicarious in-flight radiometric calibration of the sensor. The results show that the radiometric correction of multispectral aerial images is possible in an efficient way in the photogrammetric production environment.

**Zusammenfassung:** Dieser Beitrag präsentiert die Ergebnisse einer Untersuchung von Methoden zur radiometrischen Korrektur von Aufnahmen der großformatigen photogrammetrischen Luftbild-Zeilenkamera Leica Geosystems ADS40. Die Untersuchung wurde im Rahmen des EuroSDR Projekts "Radiometric aspects of digital photogrammetric images" durchgeführt. Im August 2008 wurden hierzu Bilder über dem forstwirtschaftlichen Testgebiet Hyytiälä in Finnland erfasst. Zwei Verfahren der Prozessierung wurden evaluiert. Das erste Verfahren baut auf der photogrammetrischen Software Leica XPro auf, welche sich in radiometrischen Prozessen auf physikalische Modelle bzw. Informationen verlässt, die ausschließlich aus den Bildern abgeleitet werden können. Das zweite Verfahren baut auf ATCOR-4 auf, einem Software-Paket, welches zur physikalischen atmosphärischen Korrektur flugzeuggestützter multi- und hyperspektraler Bilder sowie Bildern von Thermalscannern dient und sowohl mit als auch ohne in-situ Beobachtungen des Rückstreuverhaltens bzw. der atmosphärischen Bedingungen operieren kann. Das Ergebnis dieser Prozesse sind Bilder, welche den Reflexionsgrad darstellen. Drei Teilnehmer prozessieren die Daten mit verschiedenen Optionen, was insgesamt zu 12 verschiedenen radiometrisch korrigierten Bildern führt. Die Analyse der Daten basiert auf Messungen des Reflexionsgrads von Referenzsignalen im Feld und im Labor sowie auf Feldmessungen von permanenten Objekten (Asphalt, Gras, Schotter). Leica XPro liefert eine Genauigkeit des Reflexionsgrads von bis zu 5% ohne Verwendung von auf dem Boden erfassten Referenzdaten. Für ATCOR-4 war die Genauigkeit des Reflexionsgrads mit radiometrischer Selbstkalibrierung des Sensors besser als 5%. Diese Ergebnisse zeigen, dass eine effiziente radiometrische Korrektur von multispektralen Luftbildern in einer photogrammetrischen Produktionsumgebung möglich ist.


www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0115
1432-8364/12/0115 $ 3.75

## 1 Introduction

Multispectral digital aerial images are collected and used in huge amounts daily around the world. They are geometrically referenced, but the colour manipulations are relative, user dependent, and often local. Even if the images are used in automatic classification and interpretation tasks, the methods and results are valid only for the images used. If the images can be converted to spectral reflectance, which is a non-ambiguous surface property, the automatic use of images would become easier. In this study, the objective of the radiometric correction is the reflectance image generation.

Converting the digital number (DN) of a pixel to a surface reflectance using atmospheric models is well known and a standard procedure with satellite images (Richter 1990, Chavez 1996). Such methods have also been established for airborne imaging spectroscopy (Richter 1996). However, radiometric correction methods are rare with aerial photogrammetric images. Wider field of view of sensors, smaller ground sample distance (GSD), large number of images and significant BRDF-effects (bidirectional reflectance distribution function) makes the radiometric correction of aerial images more demanding compared to satellite and small FOV imaging spectroscopy images. In recent years, radiometric calibration and correction has been investigated for several photogrammetric sensors (Beisl et al. 2008, Markelin et al. 2008, Ryan & Pagnutti 2009, Martinez et al. 2010). A method called radiometric aerial triangulation has been proposed for relative and absolute radiometric correction of the frame image mosaics (Chandelier & Martinoty 2009, Collings et al. 2011).

In May 2008, EuroSDR (European Spatial Data Research) launched a project called "Radiometric aspects of digital photogrammetric images" to investigate the issues of accurate radiometric processing in the photogrammetric image production line. The first phase of the project consisted of a literature review and a questionnaire (Honkavaara et al. 2009). The major conclusions of the questionnaire were that improvements were desired for the entire process: sensors, calibration, data collection, data post-processing and data utilization. The basic radiometric end products requested by image users were true colour images and reflectance images. The expected benefits of a more accurate radiometric processing included a more automatic and efficient image post-processing, better visual image quality, more accurate and automatic interpretation, and quantitative use of image data. Based on the results of the questionnaire, an empirical phase was launched to study the following topics: 1) radiometric calibration and characterization, 2) spatial resolution assessment, 3) radiometric correction and image block equalization, 4) colour enhancement of the calibrated data, 5) application oriented studies. This article studies the topics 1 and 3.

The objective of this paper is to present the main results of the performance evaluation of the radiometric correction methods for ADS40 imagery collected on 23rd of August 2008 in Hyytiälä, Finland. This evaluation was carried out in the context of the EuroSDR project. The data was processed using two commercially available processing lines: the XPro software of Leica Geosystems and the ATCOR-4 software of ReSe Applications Schläpfer. Three participants carried out the processing as follows: Swisstopo (ST) (XPro), ReSe Applications Schläpfer (ATCOR-4) and Finish Geodetic Institute (FGI) (XPro and ATCOR-4). The same dataset is used also for tree species classification in Korpela et al. (2011) and Heikkinen et al. (2011). The evaluated image versions present different scenarios for performing flight campaigns aimed for producing reflectance images: with or without accurate in-situ measurements of atmospheric parameters and with or without a possibility for vicarious in-flight radiometric calibration of the sensor.

Comprehensive results of all participants and detailed description of the evaluated methods will be published in the final report of the EuroSDR-project. This study is a continuation of the first results presented in Markelin et al. (2010) and Honkavaara et al. (2011).

The article is arranged as follows: in section 2, theories of the radiometric correction methods are briefly described. Sections 3 and 4 present the materials and methods used. Section 5 presents the main results, results are discussed

in section 6 and, finally, conclusions are given in section 7.

## 2 Radiometric Correction Theory

### 2.1 *Radiative Transfer Theory*

In a simplified case, the key formula to radiometric correction is the model for surface reflectance:

$$\rho = \frac{\pi\left(\left(c_0 + c_1 DN\right) - L_0\right)}{T_{down} T_{up} S \cos\theta_i} \quad (1)$$

where ρ is the surface reflectance, $c_0$ and $c_1$ are sensor calibration parameters, DN is target recorded digital number, $L_0$ is the path radiance, $T_{down}$ is the total downward transmittance from the top of the atmosphere (TOA) to the ground, $T_{up}$ is total upward transmittance from ground to sensor, S is the mean extraterrestrial solar irradiance and $\theta_i$ is the solar zenith angle (Beisl et al. 2008, Richter & Schläpfer 2011).

Beisl et al. (2008) give the following model for the sensitivity of the surface reflectance model (1): if the multiple reflection is not taken into consideration, the error in surface reflectance Δρ caused by the path radiance uncertainty $\Delta L_0$ is:

$$\Delta\rho = \frac{\partial\rho}{\partial L_0}\Delta L_0 \approx -\frac{\pi}{T_{down} T_{up} S \cos\theta_i}\Delta L_0 \quad (2)$$

Based on (1) and (2), a number of important conclusions can now be drawn. First, an accurate radiometric calibration of the sensor is required. Secondly, an accurate estimate of the main atmospheric parameters (aerosol type, visibility or optical thickness, water vapour) is necessary, because these influence the values of path radiance, transmittance and global flux. Also, if the main atmospheric parameters and the reflectance of two reference targets are known, the quantities $L_0$, $T_{up}$, $T_{down}$, S and ρ are known. So, a vicarious in-flight radiometric calibration of the sensor can be performed (Richter & Schläpfer 2011). In order to keep the output reflectance error small, the path radiance error $\Delta L_0$ has to be kept as small as possible. Especially the dark surfaces of low reflectance are sensitive in this respect. (2) also shows that the absolute reflectance error becomes larger for smaller transmission, i.e., for a hazy atmosphere.

### 2.2 *Leica XPro*

Leica XPro is a photogrammetric software used for the entire post-processing workflow of the ADS-imagery from data download to the generation of stereo models and orthoimages. The default product of the XPro is calibrated DN, which relates the pixel data to at-sensor radiances (ASR). In radiometric terms, the main feature of the XPro is the option to produce radiometrically corrected reflectance images. This option for atmospheric correction for reflectance image production is based on the physical modelling of radiative transfer equations and parametrization of the atmospheric parameters. To speed up the calculations, some simplifications are made in the physical modelling and parameterization of the atmosphere. In surface reflectances, the reflected radiance is divided by the incoming solar irradiance which results in a surface property. Additionally, BRDF correction based on a modified Walthall model is implemented in XPro. All corrections in XPro rely entirely on a priori sensor calibration information and atmospheric information derived from dark pixels (and bright pixel statistics for BRDF-correction) in the image data. The atmospheric correction algorithm of XPro is based on the radiative transfer equation by Kaufman & Sendra (1988). (2) shows, that the accurate radiometric correction in XPro requires a careful selection of the dark pixels. The details of the XPro radiometric correction methods are given in Beisl et al. (2008) and the method and its limitations are further studied in Heikkinen et al. (2011).

### 2.3 *ATCOR-4*

The second radiometric correction method evaluated within this paper is based on the technology of the ATCOR-4 atmospheric compensation procedure (Richter & Schläpfer 2011). This program is one of the established standards for atmospheric compensa-

tion of optical and thermal airborne remote sensing imagery. It follows a physical approach by inverting the MODTRAN® radiative transfer code and includes the correction for terrain influences, adjacency effects, spatial water vapour distribution, aerosol content variations and variation of direct and diffuse illumination. The software uses a precompiled look-up-table (LUT) for this inversion. This LUT has been resampled to the ADS spectral response using the standard description of the spectral bands. The aerosol distribution has been either derived by the dark dense vegetation approach (DDV) or by a generic constant (which mostly affects the blue spectral band). The output in the optical domain is the surface reflectance cube and in the thermal domain it is the surface (brightness) temperature and emissivity spectrum. ATCOR-4 is for wide FOV airborne scanner imagery and for all terrain types, and it includes the capability for radiometric correction in rugged terrain with cast shadow and illumination calculations. The in-flight radiometric calibration tool within ATCOR-4 allows user to feed a number of ground reference spectra using a good estimate of the atmospheric parameters from the given boundary conditions in order to find new gains and offsets for all spectral bands of an instrument.

In the optical domain (wavelength $<2.5$ µm), assuming a flat terrain, and avoiding the specular and backscattering regions, an accuracy of the retrieved surface reflectance of $\pm 0.02$ for reflectances below 0.1 and $\pm 0.04$ for reflectances above 0.4 is reported to be achievable for ATCOR-4 (Richter & Schläpfer 2002).

## 3 Materials

### 3.1 *Imagery*

A flight campaign was carried out at the Hyytiälä forestry research station in Finland on 23rd of August 2008 using a Leica ADS40 SH52 digital photogrammetric camera to validate the sensor performance and to evaluate data performance in forestry applications. A total of 15 flight lines were collected from four flying heights (1, 2, 3 and 4 km, resulting in

**Tab. 1:** ADS40 flight description. Line = name of the image line, F.h. = flying height, GSD = ground sample distance, F. Dir. = flying direction (north = 0), CF = configuration (N = nadir, B = 16° backward looking line), SenZen = average sensor zenith angle for tarps, IT = integration time, Sun Alt. and Az. = sun altitude from horizon and sun azimuth angle (0 = north), Vis. = horizontal visibility (meteorological range), T = air temperature, AOT = aerosol optical thickness. All angles are in degrees.

| Line | F. h. (km) | GSD (cm) | F. Dir. | CF | SenZen (N / B) | IT (ms) | Start time | End time | Sun Alt. | Sun Az. | Vis. (km) | T (°C) | $CO_2$ (ppm) | $O_3$ (g/cm²) | $H_2O$ (g/cm²) | AOT 500 nm |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1A | 1 | 10 | 349 | N | 14.7 | 1.94 | 9:56 | 10:00 | 27.1 | 119.2 | 47.8 | 13.6 | 379 | 6.54E-04 | 1.45 | 0.17 |
| 1B | 1 | 10 | 349 | N | 15.0 | 1.94 | 10:25 | 10:28 | 30.0 | 126.6 | 49.6 | 14.6 | 373 | 6.54E-04 | 1.41 | 0.16 |
| 1C | 1 | 10 | 169 | B | 21.4 | 1.94 | 10:33 | 10:36 | 30.8 | 128.7 | 44.4 | 14.7 | 373 | 6.54E-04 | 1.44 | 0.17 |
| 2A | 2 | 20 | 349 | N | 3.5 | 2.77 | 10:45 | 10:48 | 31.8 | 131.9 | 42.8 | 14.9 | 372 | 6.54E-04 | 1.40 | 0.15 |
| 2B | 2 | 20 | 349 | B | 16.2 | 2.77 | 11:00 | 11:03 | 33.1 | 136.0 | 48.9 | 14.8 | 374 | 6.54E-04 | 1.45 | 0.15 |
| 3A | 3 | 30 | 349 | N B | 3.3/16.0 | 4.16 | 11:18 | 11:20 | 34.5 | 141.0 | 50.0 | 15.0 | 374 | 6.54E-04 | 1.45 | 0.17 |
| 3B | 3 | 30 | 169 | N B | 28.8/32.2 | 4.16 | 11:25 | 11:28 | 35.0 | 143.0 | 47.8 | 15.5 | 374 | 6.54E-04 | 1.43 | 0.17 |
| 3C | 3 | 30 | 260 | N B | 1.1/15.7 | 4.16 | 11:33 | 11:35 | 35.6 | 145.3 | 46.8 | 15.4 | 373 | 6.54E-04 | 1.40 | 0.17 |
| 4A | 4 | 40 | 169 | N B | 10.7/18.8 | 5.54 | 11:43 | 11:47 | 36.2 | 148.3 | 50.0 | 15.5 | 372 | 6.54E-04 | 1.39 | 0.16 |
| 4B | 4 | 40 | 260 | N B | 0.2/15.6 | 5.54 | 11:52 | 11:54 | 36.8 | 150.9 | 50.0 | 15.5 | 373 | 6.54E-04 | 1.38 | 0.16 |

GSDs of 10, 20, 30 and 40 cm, respectively). Reference targets were visible on 10 of these flight lines (Tab. 1). The MS channels (red (R), green (G), blue (B), NIR (N), both nadir and 16° backward directions) were recorded in raw (uncompressed) mode.

The weather conditions were mostly clear, but some small clouds occurred during the capturing of the 3 km and 4 km flight lines. The detailed information of the images used and atmospheric conditions during the campaign are shown in Tab. 1. Visibility, Temperature and $CO_2$ values were provided by the local SMEAR-II station and $O_3$, $H_2O$ and AOT values from the AERONET station. The measurement accuracies of these molecules and parameters were not considered in this study.

## 3.2 *Ground Reference Measurements*

During the campaign, nadir spectra of the reference reflectance targets (portable tarpaulins, called tarps) and several other targets (asphalt road, gravel road, beach volley field sand, football field grass) were measured using an ASD Field Spec Pro FR spectroradiometer (Tab. 2, Fig. 1). Each target was measured 10–20 times on different places, and these measurements were averaged to get the final nadir field reference spectra ($\rho_{field\_nadir}$). The tarps are flat, well defined targets whose colours are made as lambertian as possible. The other targets used present typical flat homogenous targets that can be found on an average campaign area.

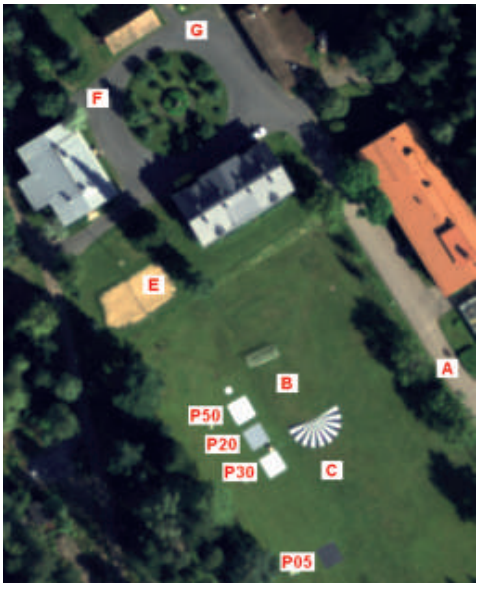



**Fig. 1:** Reference targets in field. For abbreviations, see Tab. 2.

**Tab. 2:** Ground reference targets. SH = short name for target, n = number of spectra measured, Time = measurement time (UTC+3), Sun Alt. and Az. = sun altitude from horizon and sun azimuth angle (0 = north), Refl. = average target reflectance on green channel (550 nm), CV% = ground measurement coefficient of variation (100*stdev/mean) for green channel.

| SH | Target | n | Time | Sun Alt. | Sun Az. | Refl. | CV% |
|---|---|---|---|---|---|---|---|
| A | asphalt | 13 | 9:59 | 27.4 | 120.0 | 0.140 | 2.8 |
| B | grass1 | 15 | 10:08 | 28.4 | 122.2 | 0.078 | 4.9 |
| C | grass2 | 10 | 10:16 | 29.1 | 124.3 | 0.068 | 9.6 |
| E | sand | 20 | 10:44 | 31.7 | 131.6 | 0.187 | 21.0 |
| F | gray gravel1 | 14 | 10:54 | 32.6 | 134.4 | 0.090 | 8.7 |
| G | gray gravel2 | 14 | 11:00 | 33.1 | 136.0 | 0.090 | 8.3 |
| P05 | tarpaulin 05 | 15 | 10:21 | 29.6 | 125.6 | 0.057 | 4.9 |
| P20 | tarpaulin 20 | 15 | 10:25 | 30.0 | 126.6 | 0.181 | 2.8 |
| P30 | tarpaulin 30 | 12 | 10:29 | 30.4 | 127.7 | 0.261 | 5.6 |
| P50 | tarpaulin 50 | 20 | 10:33 | 30.8 | 128.7 | 0.442 | 2.9 |

# 4 Methods

## 4.1 *General*

A total of 12 different reflectance image products were created using the XPro and ATCOR-4 software systems. Options for sensor radiometric calibration were laboratory calibration by the sensor manufacturer and vicarious in-flight calibration performed during the data processing. Atmospheric correction parameters were either based on in-situ measurements or estimated from the images. Tab. 3 shows all the different processing versions evaluated in this article. All image lines (Tab. 1) were processed with XPro, resulting in versions XA1 and XF1 by FGI. FGI also processed four nadir looking image lines (1B, 2A, 3A and 4A) with ATCOR-4, resulting in versions AL1, AL2, AV1, AV2 and AV3. Swisstopo processed the nadir looking lines 2A, 4A and 4B with XPro (versions XA2 and XF2), and ReSe processed the nadir looking line 2A with ATCOR-4, referred to as versions AL3, AV4 and AV5.

All image versions obtained by XPro and the ATCOR-4 versions AL2 and AL3 present a typical flight campaign with a laboratory-calibrated sensor, when no ground reference measurements or in-situ atmospheric observations exist. Version AL1 is the same but with detailed in-situ atmospheric observations. The ATCOR-4 processed versions AV2, AV4 and AV5 are examples of the situation when vicarious calibration of the sensor can be performed with ground reference targets, but there is no accurate information of atmospheric parameters. AV1 presents the case when extensive in-situ reference measurements provide a possibility to vicarious calibration of the sensor and atmospheric correction with in-situ atmospheric data. Finally, the ATCOR-4 processed version AV3 presents the situation where detailed vicarious calibration of the sensor is performed using a number of four reference spectra in one image line and then the same sensor calibration is used in the processing of the rest of the images.

## 4.2 *XPro Processing*

FGI used the Leica XPro version 4.1 for the entire post-processing workflow of the ADS-imagery from data download to the generation of stereo models and orthoimages (versions L2). The later version 4.2 of XPro was used for BRDF-corrections. Geometric processing was done using 12 ground control points (GCPs) and 47 check points. The RMS residuals (in m) for check points after triangulation were the following: x: 0.067, y: 0.050, z: 0.090 (1 km flying height); x: 0.050, y: 0.044, z: 0.113 (2 km); x: 0.058, y: 0.061 , z: 0.105 (3 km); x: 0.062, y: 0.078, z: 0.171 (4 km). These results indicate excellent geometric accuracy.

In radiometric processing, all images were processed without any atmospheric corrections to at-sensor radiance data. The accuracy of the XPro ASR-product was studied in Markelin et al. (2010). Next, both FGI and Swisstopo processed radiometrically corrected surface reflectance images (Tab. 3 XA1-2) and radiometrically corrected versions with BRDF-correction (Tab. 3 XF1-2), i.e. surface reflectance data corrected to nadir looking

**Tab. 3:** Processing parameters of all evaluated image versions (X for XPro, A for ATCOR-4); Cal.: origin of the sensor radiometric calibration (lab = laboratory, vic = vicarious in-flight radiometric calibration with tarps P05 and P50), Atm.: origin of the atmospheric parameters used (imag. = derived from the imagery, in-situ = in-situ measurements). Other: BRDF = with empirical BRDF-correction, cal.1B = sensor calibration based on image line 1B and all four tarps, shd. = with shadow removal.

| | XA1 | XA2 | XF1 | XF2 | AL1 | AL2 | AL3 | AV1 | AV2 | AV3 | AV4 | AV5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Participant | FGI | ST | FGI | ST | FGI | FGI | ReSe | FGI | FGI | FGI | ReSe | ReSe |
| Cal. | lab | lab | lab | lab | lab | lab | lab | vic. | vic. | vic. | vic. | vic. |
| Atm. | imag. | imag. | imag. | imag. | in-situ | imag. | imag. | in-situ | imag. | imag. | imag. | imag. |
| Other | | | BRDF | BRDF | | | | | | cal.1B | | shd. |

geometry. FGI used the XPro version 4.2 for BRDF-correction because it included an updated water masking algorithm compared to the previous version. The default settings of XPro were used for the BRDF-correction.

Swisstopo used the Leica XPro version 4.2 for the data processing. Since their aim was to test the software and not the know-how of the user, Swisstopo decided to keep the default settings given by Leica Geosystems for the statistics generation and all applied corrections.

### 4.3 *ATCOR-4 Processing*

The starting point for the ATCOR-4 processing was the nadir looking L2 orthorectified ASR-versions of the ADS-imagery created with XPro. For the processing, standard values for the water vapour amount has been taken into account as this parameter is of minor influence in the ADS spectral wavelength range.

ReSe used the following processing steps with ATCOR-4. First, the solar and flying geometries were defined for each image. Then ATCOR-4 was started with the gain values provided in the sensor specific Leica calibration files, and the derived surface reflectance outputs were visually checked (Tab. 3 AL3). Next the in-field targets were used to perform an in-flight vicarious calibration, which lead to updated gain/offset values. The atmospheric correction was repeated using the in-flight calibration coefficients (Tab. 3 AV4). Finally the cast-shadow correction and building/tree correction was tested and applied on the basis of in-flight calibrated data (Tab. 3 AV5).

The FGI processing was done using the ATCOR-4 software version 5.1. Small subsets of the size about 2 km x 2 km were cropped from the original image lines. Because of limitations with the available system memory, images from flying heights 1, 2 and 4 km were resampled to 40 cm GSD and the image from the 3 km flying height to 60 cm GSD. Five different radiometrically corrected versions were then calculated for each flying height (Tabs. 3 and 4). In versions AL1 and AL2, the sensor radiometric calibration parameters were taken from Leica Geosystems (i.e., how the DNs of ASR-images were converted to at-sensor radiances). In versions AV1 and AV2 the sensor calibration parameters were determined individually for each image by using the ATCOR-4 in-flight calibration module and spectra of reference reflectance targets P05 and P50. In AV3, the sensor calibration parameters were determined from the 1 km flying height image 1B and using in-flight calibration with all the four reference reflectance targets. Atmospheric parameters (aerosol type, water vapour and visibility) for the versions AL1 and AV1 were set based on the in-situ atmospheric measurements (Tab. 3). For the versions AL2, AV1 and AV3 the atmospheric parameters were derived from the image data using the ATCOR-4 modules; also the variable visibility option was activated. Tab. 4 shows the final atmospheric parameters used in the ATCOR-4 processing (compare to in-situ measurements in Tab. 1).

The influence of topography was not considered in the ATCOR-4 processings as the test area is mostly flat.

**Tab. 4:** Atmospheric parameters used in ATCOR-4 processing for flight lines 1B, 2A, 3A and 4A; aer. = aerosol type (rural, urban, maritime or desert), vis. = horizontal visibility in km, AOT = aerosol optical thickness at 550 nm.

| | 1B | | | 2A | | | 3A | | | 4A | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | aer. | vis. | AOT | aer. | vis. | AOT | aer. | vis. | AOT | aer. | vis. | AOT |
| AL1, AV1 | rural | 49.6 | 0.165 | rural | 42.8 | 0.187 | rural | 50.0 | 0.164 | rural | 50.0 | 0.164 |
| AL2 | rural | 83.1 | 0.108 | mari. | 88.9 | 0.098 | mari. | 95.4 | 0.095 | mari. | 95.4 | 0.094 |
| AV2 | rural | 62.3 | 0.137 | rural | 69.4 | 0.125 | rural | 73.4 | 0.119 | rural | 73.4 | 0.123 |
| AV3 | rural | 62.3 | 0.137 | rural | 73.4 | 0.123 | rural | 78.0 | 0.113 | rural | 83.1 | 0.103 |
| AL3, AV4, AV5 | | | | urban | 80.0 | 0.110 | | | | | | |

### 4.4 *Analysis Methods*

Reflectance of all ground reference targets (Tab. 2) were measured from the images; the size of the measurement window was 3 m × 3 m in object coordinates. These measurements provided the target data reflectance ($\rho_{data}$). The exact positions of image measurements of targets other than tarps were ambiguous (measurement on different images, and difference between field reference measurement and image measurement), especially on grass, which may add some uncertainty to the results.

As a reference reflectance, laboratory measurements with the FIGIFIGO goniospectrometer (SUOMALAINEN et al. 2009) in the exact imaging geometry, scaled with the nadir field measurements were used. This scaling was done to match the laboratory measurements to the actual imaging conditions. The reference reflectance for tarps is:

$$\rho_{ref} = \rho_{lab\_exact} \frac{\rho_{field\_nadir}}{\rho_{lab\_nadir}} \quad (3)$$

where $\rho_{field\_nadir}$ is the target nadir reflectance measurement during the imaging campaign, $\rho_{lab\_nadir}$ is the target nadir reflectance measured at the laboratory, and $\rho_{lab\_exact}$ is the target reflectance at the exact imaging geometry (illumination and viewing angles) of the respective image.

The difference of the target image reflectance and the reference reflectance was calculated to obtain the reflectance error in reflectance units:

$$E_{refl} = \rho_{data} - \rho_{ref} \quad (4)$$

Next this difference was divided by the reference and multiplied by 100 to get the reflectance error in percents:

$$E_{refl\%} = 100 \frac{\rho_{data} - \rho_{ref}}{\rho_{ref}} \quad (5)$$

The $E_{refl}$ and $E_{refl\%}$ were calculated for all tarps, images, and colour channels. From these errors, root mean square error values ($RMSE_{refl}$ and $RMSE_{refl\%}$) were calculated using all of the four tarps (except when the targets were used in the sensor calibration) for each image line and channel. The $RMSE_{refl\%}$ is:

$$RMSE_{refl\%} = \sqrt{\frac{\Sigma E_{refl\%}^2}{n}} \quad (6)$$

where $n$ is the number of targets used.

For a relative comparison of the different flying heights, 2 km, 3 km and 4 km data were compared to 1 km data, which was used as a reference. For tarps, the data was first scaled to nadir view using laboratory calculated anisotropy factors:

$$\rho_{data\_nadir} = \rho_{data} \frac{\rho_{lab\_nadir}}{\rho_{lab\_exact}} \quad (7)$$

Next, relative differences were calculated between 1 km reference data and 2 km, 3 km and 4 km data using (5), and finally $RMSE_{refl\%}$ was calculated using (6).

For targets other than tarps (asphalt, sand, gravel, grass) the same kind of calculations for $E_{refl}$ and $E_{refl\%}$, $RMSE_{refl}$, $RMSE_{refl\%}$ and relative flying height comparison were carried out, but nadir field measurements were used as a reference. This means that the anisotropy effects of targets on images with large viewing angles are included in the results, which can add some error. The grass was processed separately from the other non-organic targets.

The reflectance error is given in most cases as percents of magnitude of the reflectance; we specify the cases separately where the error is given in reflectance units.

The dependency of $E_{refl}$ and $E_{refl\%}$ on the magnitude of reflectance was evaluated by plotting these errors against the reference reflectance. Linear fit parameters (gain, offset and $R^2$-value) were calculated for the data and the dependency was analyzed based on the $R^2$-values.

## 5 Results

The main results of the study are presented in the following subsections. The most thorough analysis is shown for the tarps, because they had the most accurate reference and tarps are unambiguous as targets. Only selected results

are shown for other targets. Comprehensive results of all participants and all processing methods will be published in the final report of the EuroSDR project.

## 5.1 *General Results*

First we made some general evaluations of the data to remove identical results.

a. The reflectance accuracy was evaluated both in reflectance error given in reflectance units (4) and reflectance error in percents (5). Both methods lead to the same conclusions, so only results with reflectance error in percent (and $RMSE_{refl\%)}$ are shown.
b. The results with the dark tarp (P05) are presented separately from the bright tarps (P20, P30 and P50) because the reflectance of dark target is sensitive to errors in path radiance modelling. Also the detection accuracy of dark pixels is essential in the evaluated XPro methods and a small inaccuracy will cause a large relative error in dark areas.
c. The XPro processing results of Swisstopo and FGI were practically the same, so only the FGI results are shown in most cases.
d. The XPro correction versions XA and XF provided a mostly similar accuracy. Because the Hyytiälä area is mostly forest and the XPro BRDF-correction is based on DN averages over large areas, the BRDF-correction is not expected to work accurately enough with tarps. So results with the XF-versions are considered only with the grass target.
e. The XPro results for nadir and 16° backward looking lines were mostly similar (especially when both views were collected simultaneously, i.e., for the 3 and 4 km flying heights), so in most cases only nadir results are shown.
f. For the cases where similar results were obtained on different passes, averages per flying height were calculated for XPro with tarps.
g. The ATCOR-4 processing results with atmospheric parameters set by the user (AL1 and AV1) and parameters derived automatically from the images (AL2 and AV2) provided similar results, so only results with AL2 and AV2 are shown.

The reflectance error (4) was dependent on the magnitude of the reflectance, which could be compensated to a large extent by evaluating reflectance errors in percent (5).

Some small clouds moved over the test area during capturing the 3 km and 4 km flying height lines. By visual inspection, these clouds or their shadows did not overshadow the reference targets, but it is possible that the clouds had some effect on the results.

The shadow removal method used in the ATCOR-4 processing version AV5 erroneously interpreted many dark objects (e.g. tarp P05) as shadows (Fig. 3 top).

The results with tarps for ATCOR-4 processings version AV3 are not independent, because the same tarps were also used with the sensor calibration. These results are still shown for

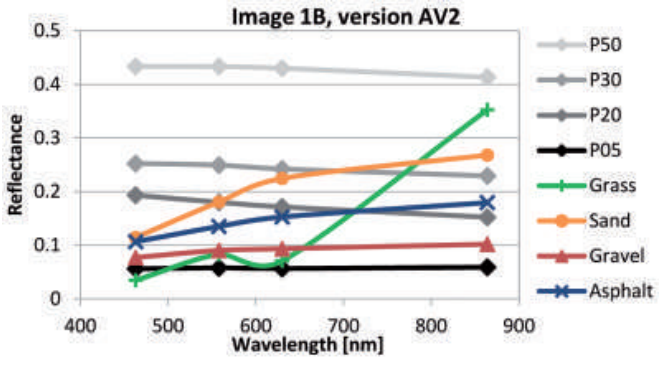



**Fig. 2:** Reference target reflectance spectra for image 1B (1 km flying height) with AV2 processing. Marks show ADS40 channels.

comparison. The accuracy of the method AV3 (as well as all the methods) is evaluated also with other (independent) targets.

Fig. 2 shows an example of the reflectance spectra of the reference targets derived from the imagery.

## 5.2 *Reflectance Accuracy with Tarps*

A reflectance accuracy better than 5% was achievable with all the evaluated methods. Results varied depending on the target, colour channel, origin of the sensor radiometric calibration, flying height and weather conditions. Fig. 3 shows the reflectance error for the dark tarp and $RMSE_{refl\%}$ for the bright tarps for the XPro and ATCOR-4 nadir looking lines. In Fig. 4, $RMSE_{refl\%}$ of all lines processed with XPro are shown.

The reflectance error was on an average higher for the dark tarp (often over 10%, Fig. 3 top) than for the bright tarps (mostly below 10%, Fig. 3 bottom). On the other hand, the reflectance error in reflectance units was smaller (below 0.01) for the dark tarp than for the bright tarps (below 0.04). This behaviour was expected, because even small reflectance errors for a dark object become large when scaled to %.

On the bright tarps, the green, red and sometimes NIR-channel provided the best results and the blue channel the worst (Fig. 3 bottom). The green and red channels behaved similarly compared to each other. The NIR channel was the worst on the 2 km flying height on ATCOR-4 methods AL2, AL3, AV4 and AV5. On the dark tarp, the NIR channel gave the worst results for the 3 and 4 km flying heights (Fig. 3 top).

The vicarious calibration of the sensor improved the results with ATCOR-4 on all methods, except for the NIR channel with the dark tarp (Fig. 3).

The effect of flying height was minor with XPro except for the blue channel, on which the reflectance accuracy was clearly worse for the 3 and 4 km flying heights compared to the 1 and 2 km flying heights (Fig. 4). With ATCOR-4, the reflectance errors from the 3 and 4 km flying heights were slightly higher (over 5%) than for the 1 and 2 km flying heights (below 5%) (Fig. 3).

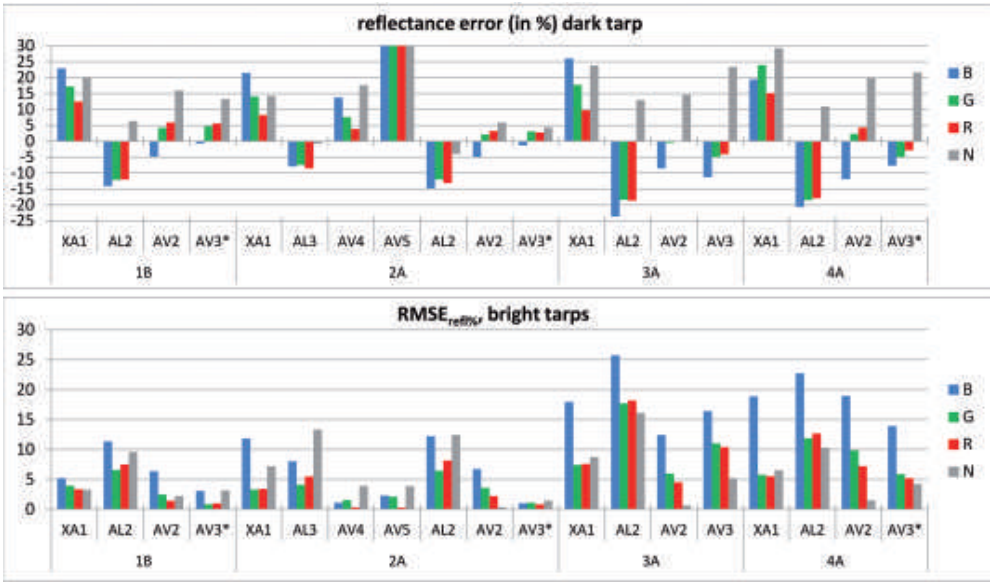



**Fig. 3:** Reflectance error (in %) for the dark tarp (top) and $RMSE_{refl\%}$ for the bright tarps (bottom). The values for XPro are averages of all image lines for a certain flying height. * denotes that the results for AV3 are not independent. For AV5 dark tarp image 2A, the reflectance error values are B: 62%, G: 81%, R: 79% and NIR: 121%.

With the XPro there appeared a clear dependency of the reflectance error to the magnitude of reflectance – both in reflectance units and in percent (modelled as linear in this investigation). As an example, in Fig. 5 reflectance errors in reflectance units and in percent are shown as a function of the reflectance for the XPro XA1 versions of lines 4A and 4B. Fig. 5 includes tarpaulins with exact reference and non-organic targets with nadir field reference. The dependency of the reflectance errors (in reflectance units) to magnitude of reflectance was clear for all XPro and ATCOR-4 methods and all colour channels with flying heights 3 km and 4 km. For XPro, this dependency was detected also for the red and green channels with the 2 km flying height and the blue-channel with the 1 km flying height imagery. After scaling to reflectance errors (in %), the $R^2$ values for linear fit lowered

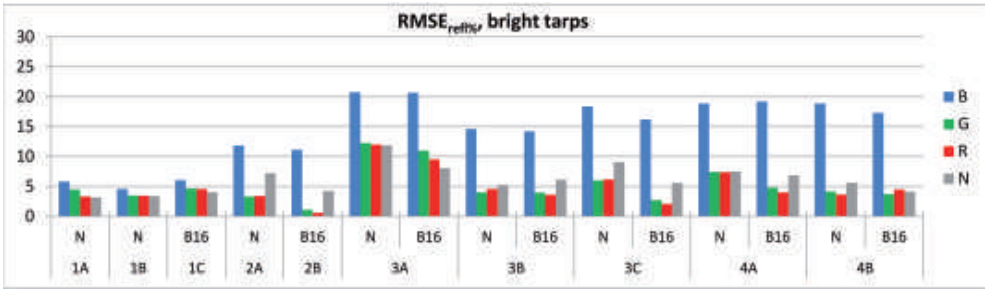



**Fig. 4:** $RMSE_{refl\%}$ for XPro, bright tarps, all lines, XA1 versions. N = nadir looking, B16 = backward looking image line.

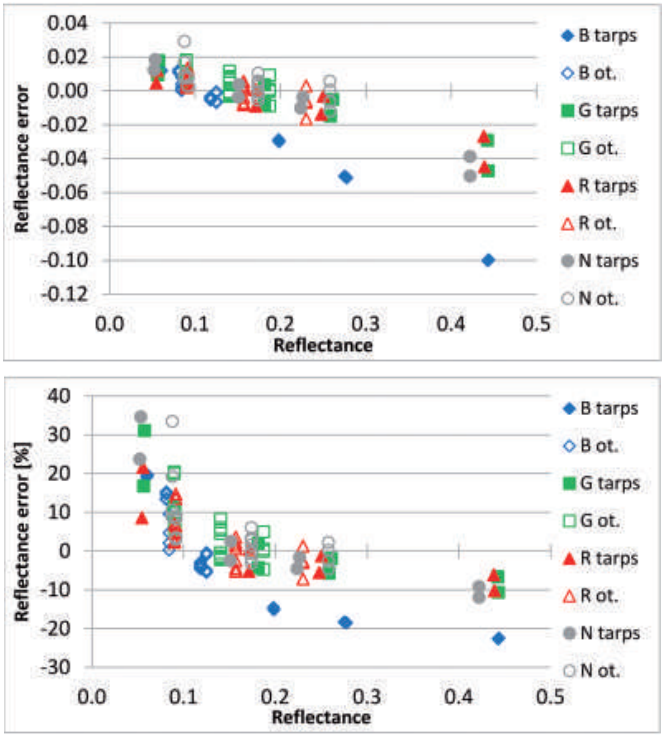



**Fig. 5:** Reflectance error for the XPro XA1 versions of nadir looking lines 4A and 4B. Reflectance error in reflectance units (top) and reflectance error in percent (bottom). Filled marks are for tarps, open marks for non-organic targets (asphalt, gravel, sand).

clearly with all ATCOR-4 methods, except for the blue and green channels for the 2 km flying height and the NIR-channel for the 3 and 4 km flying heights with the methods including sensor laboratory calibration (AL1–AL3). For XPro, the dependency of reflectance error to magnitude of reflectance remained after the scaling of reflectance errors to percents.

### 5.3 *Reflectance Accuracy with other Targets*

The $RMSE_{refl\%}$ results for non-organic (asphalt, gravel, sand) and grass targets are shown in Fig. 6. Both XPro and ATCOR-4 provided reflectance accuracy better than 5% at best. For non-organic targets, XPro and ATCOR-4 with vicarious calibration of the sensor (AV1–AV5) produced stable $RMSE_{refl\%}$ between 5 and 10% on all flying heights with some exceptions. ATCOR-4 with sensor laboratory calibration (AL1–AL3) provided a slightly higher reflectance error (5 to 20%) compared to versions with vicarious calibration of the sensor (AV1–AV5) (Fig. 6 top).

The grass targets provided variable results on all methods, both in errors in reflectance units and in percent. The NIR-channel gave the largest reflectance errors in reflectance units, between 0.04 to 0.13. On the other channels, the errors were mostly below 0.02. In $RMSE_{refl\%}$ the XPro blue channel gave the worst results, 40% and more; the green and red channels gave a $RMSE_{refl\%}$ below 25% and the NIR-channel below 20%, sometimes even below 10%. On the ATCOR-4 methods, the blue and green channels provided the lowest errors, sometimes even below 5%. $RMSE_{refl\%}$ of the red and NIR-channels were mostly between 10 and 20%. For blue and green channel, methods AL2, AV2 and AV3 provided sometimes even better $RMSE_{refl\%}$ results with grass targets than with tarps or other targets. The empirical XPro BRDF-correction on the FULL-versions (XF1) did not provide a significant improvement compared to the radiometric correction (XA1) only (Fig. 6 bottom).

The reflectance accuracy of different objects can be compared based on the Figs. 3 (bottom), 6 (top) and 6 (bottom). The results with tarps and non-organic targets were quite similar: on low flying altitudes (1 and 2 km) tarps were mostly better; for higher altitudes the results were similar on green and red channels, tarps provided better results on

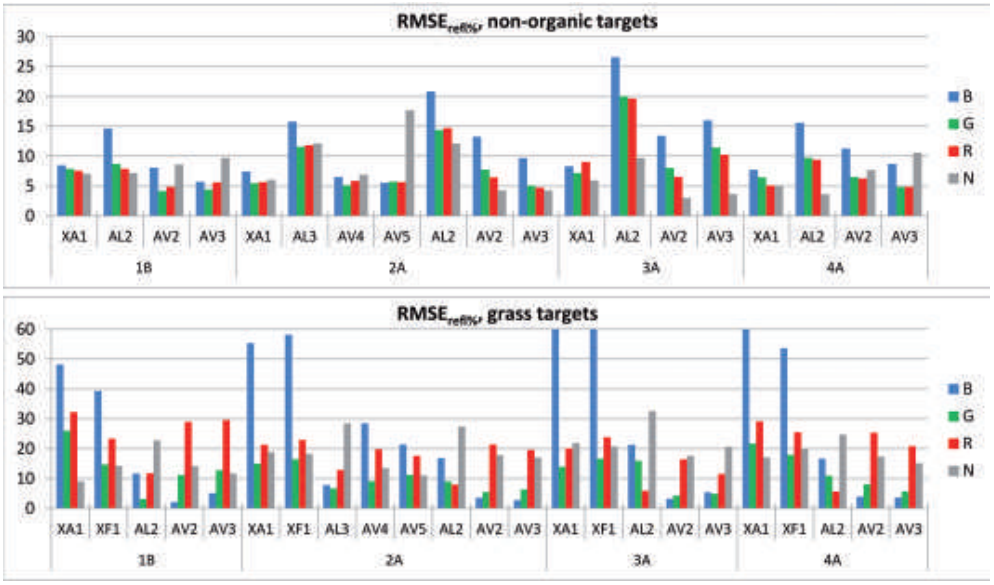



**Fig. 6:** $RMSE_{refl\%}$ results for non-organic (top; asphalt, gravel and sand) and grass targets (bottom). The $RMSE_{refl\%}$ values for grass targets, blue channel, image 3A method XA1 is 66%, XF1 71% and image 4A method XA1 61%.

NIR-channel and other targets on blue channel. Variability of results was clearly higher with the grass target, which is caused mostly by the vagueness of the object.

### 5.4 *Comparison of Methods*

The $RMSE_{refl\%}$ results for the bright tarps of all processing methods evaluated for the 2 km flight line 2A are shown in Fig. 7. All but the AL2 could produce a $RMSE_{refl\%}$ below 5% at least on some channels. For XPro, the green and red channels provided the best results and the blue the worst. For the ATCOR-4 versions, the methods with vicarious calibration of the sensor (AV1–AV5) outperformed methods with sensor laboratory calibration (AL1–AL3).

The internal repeatability of results with different processing methods was evaluated by comparing the results from the 2, 3 and 4 km flying heights to 1 km flying height (Fig. 8). Both XPro and ATCOR-4 provided a difference of 5% or less for the 2 km flying height. The difference for the 3 and 4 km flying height varied between 5 to 15% for the blue, green and red channels, and between 2 to 10% for the NIR-channel.

## 6 Discussion

The main focus of this article was to evaluate accuracies of reflectance images produced from the images of multispectral aerial line scanner data. In total, 12 different approaches and corrected image versions were compared. The evaluated methods included two different software systems (photogrammetric software Leica XPro and ATCOR-4 software dedicated to atmospheric correction) and the radiometrically corrected images represented both a typical photogrammetric campaign without any ground reference information and a campaign with a comprehensive reference including targets, field measurements and atmospheric observations.

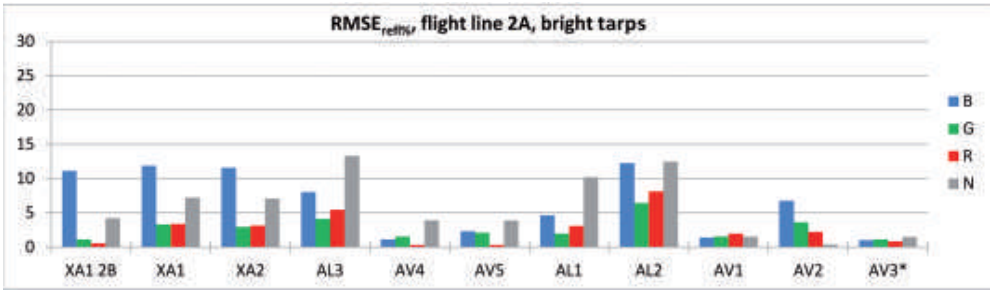



**Fig. 7:** $RMSE_{refl\%}$ results for the bright tarps (all methods, image line 2A for nadir, 2B for XA1 backward looking line) (dark tarp behaved similarly to Fig. 3). * denotes that the results for AV3 are not independent.

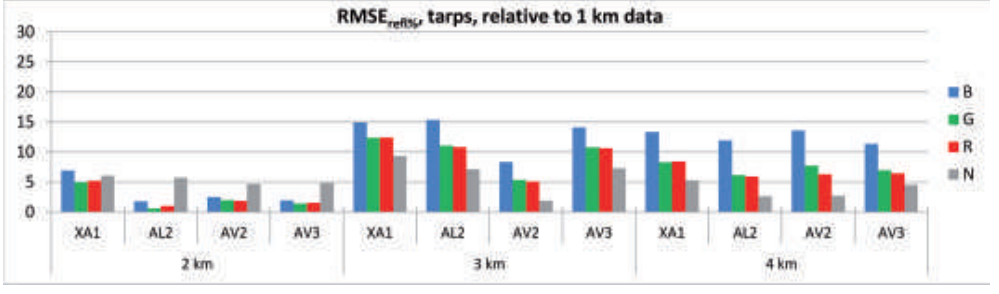



**Fig. 8:** Internal repeatability of methods, $RMSE_{refl\%}$ comparison relative to 1 km data.

Even though the campaign was flown in mainly excellent weather conditions, small clouds that occurred on the test area during the imaging of the 3 km and 4 km image lines may have some effect on the results.

A prerequisite of the processing accuracy was the radiometrically and spectrally calibrated stable sensor. In the case of ADS40 the quality is ensured by careful lens and filter design, using temperature stabilization and fixed aperture, and by accurate calibration processing. However, the reflectance products showed variation between different flying heights. In this investigation, these were most likely due to the inaccurate atmospheric and BRDF modeling, but also the sensor instability is one possible reason. Uniformity of reflectance product over the entire block area could be improved by using radiometric block adjustment methods; one approach was developed by Chandelier & Martinoty (2009) and another by Collings et al. (2011).

Both evaluated software systems provided good results without accurate in-situ measured atmospheric parameters, i.e., both methods were able to estimate the needed parameters automatically from the image data itself. It may also help that Leica ADS40 is a multispectral sensor with wide spectral bands (compared to hyperspectral sensors), so the system is not so sensitive to small changes in aerosol and molecule (e.g. $H_2O$, $CO_2$, $O_3$) concentrations. XPro produced stable results on all channels and flying heights with the laboratory calibration of the sensor. With ATCOR-4, the vicarious radiometric in-flight calibration of the sensor improved the reflectance accuracies of all processed images compared to the image versions with the sensor laboratory calibration. Further studies are needed to find out if this was due to sensor instability or some other reasons.

An important issue in interpretation applications is the natural variation of objects. The accuracies of the methods were tested using well-defined, flat, nearly-lambertian reference reflectance targets and typical permanent flat targets (asphalt, gravel road, sand, grass) that can be found from the coverage of the images. In our results, the reflectance errors of grass were 1 to 60 times, and on average 4.7 times higher than the reflectance errors of uniform asphalt and sand targets or tarps. The same ADS40 reflectance datasets were recently used by Korpela et al. (2011) in forestry applications and in tree species classification. Evaluations showed that the variation coefficients of tree species in forests were 13–31% of reflectance values. High variability will decrease the separability of tree species in classification (Korpela et al. 2011, Heikkinen et al. 2011).

An important advantage of the XPro reflectance image production was its efficient workflow with minimal user interaction. Also the XPro is able to process large amounts of images automatically, an essential requirement for a practical workflow, which was proven in operational conditions by Swisstopo. On the other hand, the effect of the few adjustable statistical parameters in the BRDF-correction would require further studies. For a broadband multispectral sensor like the ADS a rather simple approach of XPro for atmospheric correction gave results comparable to ATCOR-4. However, if ground reference data are available the results can be improved by the use of ATCOR-4. Adding an in-flight calibration procedure to XPro could further improve the results.

Important features of the ATCOR-4 were the possibility to vicarious in-flight radiometric calibration of the sensor and advanced methods for setting and detecting atmospheric parameters from the images, which are useful in applications with high reliability requirements. Because of the sophisticated modules and options, ATCOR-4 requires an experienced operator to produce reliable and accurate results. The results show that shadow correction method used in ATCOR-4-based method AV5 is not applicable to high spatial resolution multispectral instruments as it had been developed for 5–10 m resolution hyperspectral imagery, and requires further development for multispectral photogrammetric applications.

An improvement for both evaluated software products would be an option for radiometric block adjustment.

## 7 Conclusions

In this article, two methods and in total 12 different processing versions for correcting the airborne multispectral ADS40 at-sensor radiance images to surface reflectance images were evaluated. The results showed that reflectance accuracy level of 5% is possible with sensor laboratory calibration and even without any reference measurements. With vicarious radiometric in-flight calibration of the sensor, reflectance accuracies even better than 5% were achieved.

## Acknowledgements

Our sincere acknowledgements go to Leica Geosystems, Estonian Landboard and National Land Survey of Finland for their support. We are grateful for all the colleagues who helped in the field work, especially Teemu Hakala and Leena Matikainen for the field work and Juha Suomalainen for the reflectance library.

## References

Beisl, U., Telaar, J. & Schönermark, M.v., 2008: Atmospheric correction, reflectance calibration and BRDF correction for ADS40 image data. – The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences **37** (B7).

Chandelier, L. & Martinoty, G., 2009: Radiometric aerial triangulation for the equalization of digital aerial images and orthoimages. – Photogrammetric Engineering & Remote Sensing **75** (2): 193–200.

Chavez, R., 1996: Image-based atmospheric corrections – revisited and improved. – Photogrammetric Engineering & Remote Sensing **62** (9): 1025–1036.

Collings, S., Cacetta, P., Campbell, N. & Wu, X., 2011: Empirical models for radiometric calibration of digital aerial frame mosaics. – IEEE Transactions on Geoscience and Remote Sensing **49** (7): 2573–2588.

Heikkinen, V., Korpela, I., Tokola, T., Honkavaara, T. & Parkkinen, J., 2011: An SVM classification of tree species radiometric signatures based on the Leica ADS40 sensor. – IEEE Transactions on Geoscience and Remote Sensing **49** (11): 4539–4551.

Honkavaara, E., Arbiol, R., Markelin, L., Martinez, L., Cramer, M., Bovet, S., Chandelier, L., Ilves, R., Klonus, S., Marshall, P., Schläpfer, D., Tabor, M., Thom, C. & Veje, N., 2009: Digital airborne photogrammetry – A new tool for quantitative remote sensing? – A state-of-the-art review on radiometric aspects of digital photogrammetric images. – Remote Sensing **1** (3): 577–605.

Honkavaara, E., Arbiol, R., Markelin, L., Martinez, L., Bovet, S., Bredif, M., Chandelier, L., Heikkinen, V., Korpela, I., Lelegard, L., Pérez, F., Schläpfer, D. & Tokola, T., 2011: The EuroSDR project "Radiometric aspects of digital photogrammetric images" – Results of the empirical phase. – Proceedings of the ISPRS Hannover Workshop 2011, 8 p.

Kaufman, Y.J. & Sendra, C., 1988: Algorithm for automatic atmospheric corrections to visible and near-IR satellite imagery. – International Journal of Remote Sensing **9** (8): 1357–1381.

Korpela, I., Heikkinen, V., Honkavaara, E., Rohrbach, F. & Tokola, T., 2011: Variation and directional anisotropy of reflectance at the crown scale – implications for tree species classification in digital aerial images. – Remote Sensing of Environment **115** (8): 2062–2074.

Markelin, L., Honkavaara, E., Peltoniemi, J., Ahokas, E., Kuittinen, R., Hyyppä, J., Suomalainen, J. & Kukko, A., 2008: Radiometric calibration and characterization of large-format digital photogrammetric sensors in a test field. – Photogrammetric Engineering and Remote Sensing **74** (12): 1487–1500.

Markelin, L., Honkavaara, E., Beisl, U. & Korpela, I., 2010: Validation of the radiometric processing chain of the Leica ADS40 airborne photogrammetric sensor. – International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences **38** (7A): 145–150.

Martínez, L., Arbiol, R. & Pérez, F., 2010: CASI Characterization and atmospheric correction for EuroSDR Banyoles08 dataset. – Proceedings of the EuroCOW 2010, 4 p.

Richter, R., 1990: A fast atmospheric correction algorithm applied to Landsat TM images. – International Journal of Remote Sensing **11** (1): 159–166.

Richter, R., 1996: Atmospheric correction of DAIS hyperspectral image data. – Computers & Geosciences **22** (7): 785–793.

Richter, R. & Schläpfer, D., 2002: Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: atmospheric / topographic

correction. – International Journal of Remote Sensing **23** (13): 2361–2649.

Richter, R. & Schläpfer, D., 2011: Atmospheric / Topographic Correction for Airborne Imagery (ATCOR-4 User Guide, Version 6.0.2, August 2011). – 194, DLR report DLR-IB 565-02/11, Wessling, Germany, rese.ch/download/.

Ryan, R.E. & Pagnutti, M., 2009: Enhanced absolute and relative radiometric calibration for digital aerial cameras. – Fritsch, D. (ed.): Photogrammetric Week **2009:** 81–90, Wichmann, Heidelberg, Germany.

Suomalainen, J., Hakala, T., Peltoniemi, J. & Puttonen, E., 2009: Polarised multiangular reflectance measurements using the Finish geodetic institute field goniospectrometer. – Sensors **9** (5): 3891–3907.

Addresses of the Authors:

Lauri Markelin, Eija Honkavaara, Finnish Geodetic Institute, Geodeetinrinne 2, P.O. Box 15, FI-02431 Masala, Tel.: +358-9-29555239, +358-40-1920835, e-mail: first name.last name@fgi.fi

Daniel Schläpfer, ReSe Applications Schläpfer, Langeggweg 3, CH-9500 Wil, Tel.: +41-71-911-4614, e-mail: info@rese.ch

Stéphane Bovet, Swisstopo – Land survey of Switzerland, Seftigenstrasse 264, CH–3084 Wabern, Tel.: +41-31-963-2186, e-mail: stephane.bovet@swisstopo.ch

Ilkka Korpela, University of Helsinki, Department of Forest Sciences, P.O. Box 27, FI-00014 University of Helsinki, Tel.: +358-400-218305, e-mail: ilkka.korpela@helsinki.fi

Manuskript eingereicht: Oktober 2011
Angenommen: Februar 2012

PFG 2012 / 3, 0267–0280
Stuttgart, June 2012

Article

# A Decision Fusion Framework for Hyperspectral Subpixel Target Detection

Hamed Gholizadeh, Mohammad Javad Valadan Zoej & Barat Mojaradi, Tehran, Iran

**Keywords**: Target detection, decision fusion, hyperspectral data, remote sensing

**Summary:** Target detection is one of the most challenging issues of remotely sensed data. Due to high spectral resolution of the hyperspectral images and their limited ground sampling distance, targets of interest occur at subpixel level. In such cases, spatial characteristics of targets are hard to acquire and the only way to overcome such problem is to take advantage of the spectral information. Based on the spectral characteristics of background and the targets to be detected, several methods have been proposed. Some of these methods assume a physics-based approach, while the other may be purely statistical. So, all of these methods are based on some assumptions each of which can be challenged in one way or another. One possible way to take advantage of these differences to improve the final results is the fusion of the detectors' outputs. In this paper, eight subpixel target detectors are employed as the ensemble detectors. It is also worth mentioning that the detectors should be different from each other; otherwise the overall decision will not be better than the individual detectors. So, we suggest using the genetic algorithms to select the most suitable detectors for a given decision fusion rule. Experimental results on a real world hyperspectral data as well as a synthetic dataset show the efficiency of the proposed method to improve the detection performance.

**Zusammenfassung:** *Die Objekterkennung in Fernerkundungsszenen ist bislang nur teilweise gelöst.* Im Fall von Hyperspektraldaten steht der hohen spektralen eine begrenzte räumliche Auflösung gegenüber. Daher sind viele Objekte kleiner als ein Pixel, so dass eine Aussage über die geometrischen Eigenschaften nur eingeschränkt möglich ist. Daher kommt der spektralen Information eine erhöhte Bedeutung zu. In der Vergangenheit wurden viele Analysemethoden vorgeschlagen, die die Objekterkennung nach den spektralen Charakteristiken der gesuchten Objekte und ihrer Umgebung, dem Bildhintergrund, erlauben, Einige der Methoden verfolgen modellbasierte Ansätze während andere rein statistisch arbeiten. Alle Methoden erfordern spezifische Annahmen, die eine zusätzliche Unsicherheit für das Ergebnis bedeuten. Ein Ansatz zur Verbesserung des Gesamtergebnisses ist die Verschneidung (Fusion) der mit den unterschiedlichen Methoden (Detektoren) gefundenen Einzelergebnisse.

In diesem Artikel werden acht typische Detektoren beispielhaft untersucht und gezeigt, wie mit Hilfe der Methode Genetischer Algorithmus die für eine gegebene Fragestellung geeignetste Kombination gefunden werden kann. Die Methode wird sowohl an echten als auch an synthetischen Hyperspektraldaten erprobt. Die Untersuchung zeigt, dass die vorgeschlagene Methode die Erkennbarkeit von Objekten verbessert.

## 1 Introduction

Spectral subpixel detection in hyperspectral image (HSI) data aims to identify a target smaller than the size of a pixel using only spectral information (Broadwater & Chellappa 2007). This *a priori* information may be obtained from *in situ* measurements, a spectral library or can be image derived.

Subpixel target detectors are completely dependent on how the scene endmembers are modeled. If the total surface area is conceived to be divided proportionally according to the fractional abundances of the constituent subspaces, then there exists a linear relation between the fractional abundances of the substances comprising the area being imaged, and the spectra in the reflected radiation


www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0116
1432-8364/12/0116 $ 3.50

(Manolakis et al. 2003). Hence, this is called the linear mixing model (LMM), and is expressed as

$$r = \sum_{j=1}^{p} \alpha_j m_j + n = M\alpha + n \quad (1)$$

where $r$ is the $l \times 1$ received pixel spectrum vector, $m_j$ which represents the $j$-th endmember, is the $j$-th $l \times 1$ column of $M$, $\alpha$ is the $P \times 1$ fraction abundance column vector, $n$ is the $l \times 1$ additive observation noise vector, $M$ is the $l \times P$ matrix whose columns are $m_j$, $l$ is the number of spectral bands and $P$ is the number of endmembers.

We may confront a more complicated scenario in which the substances comprising the medium are organized proportionally on the surface. This intimate mixture results when each component is randomly distributed in a homogeneous way. As a result, the incident radiation may no longer uphold the linear properties of the constituent substance spectra. This scenario is referred to as non linear mixing (Manolakis et al. 2003).

Conventional subpixel target detectors can be divided into four categories (Broadwater & Chellappa 2007): One of the earliest methods uses array processing techniques to nullify the background signatures as one would nullify an interfering signature when performing beamforming (Broadwater & Chellappa 2007). The orthogonal subspace projection (OSP) (Harsanyi & Chang 1994), constrained energy minimization (CEM) (Chang & Heinz 2000) and target constrained interference minimized filter (TCIMF) (Ren & Chang 2000) algorithms are examples of such methods. Another approach uses the linear mixing model to directly estimate the abundance values and uses the estimated target abundances for detection purposes (Broadwater & Chellappa 2007). Fully constrained least squares (FCLS) (Heinz et al. 1999), non-negatively constrained least squares (NCLS) (Bro & Jong 1997) and sum-to-one constrained least squares (SCLS) fall within this category. Since FCLS and NCLS attempt to address the phenomenological constraints in the linear mixing model, they can be considered physics-based (Broadwater & Chellappa 2007). There are other kinds of subpixel target detectors which are based on a statistical hypothesis test. Adaptive matched subspace detector (AMSD) (Manolakis et al. 2001) is a statistical target detector based on structured background. Unlike AMSD, adaptive cosine/coherent estimate (ACE) (Kraut et al. 2005), which is another statistical target detector, assumes no structured background.

It is also well known that in many situations combining the output of several classifiers leads to an improved classification result. This happens because the subset of the input space that one classifier will attribute a correct label will differ from the other. This implies that by using information from more than one classifier it is probable that a better overall accuracy can be obtained for a given problem (Alexandre et al. 2001). A single classifier is generally unable to handle the wide variability and scalability of the data in any problem domain (Mangari et al. 2010). There are mainly three types of fusion strategies (Dasarathy 1994), namely, information/data fusion (low-level fusion), feature fusion (intermediate-level fusion), and decision fusion (high level fusion). In this paper the fusion process is carried out on decision level.

The aim of this paper is the decision fusion of 8 subpixel target detectors. Since there might be cases in which some detectors show similar performance, we employed the genetic algorithms (GA) to choose the most optimum detectors as the fusion input.

The remainder of this paper is organized as follows. Section 2 describes eight subpixel detection methods. Section 3 briefly reviews decision fusion approaches and the proposed method. Performance comparison of the proposed method is provided in section 4 and the conclusions are included in section 5.

## 2 Conventional Subpixel Target Detectors

### 2.1 *Orthogonal Subspace Projection (OSP)*

In OSP, the endmember matrix $M$ in (1) is divided into two parts: desired signature vector ($d = m_p$) and undesired signature matrix ($U = [m_1, m_2, \ldots, m_{p-1}]$. Then, we can rewrite (1) as

$$r = d\alpha_p + U\gamma + n \quad (2)$$

where $\alpha_p$ is the abundance fraction of the desired signature and $\gamma$ is a $(p-1) \times 1$ abundance fraction vector of the undesired signatures. To suppress the undesired target, the orthogonal projection operator ($P_U^{\perp}$), which maps data onto a subspace orthogonal to the undesired signatures space, is used.

$$P_U^{\perp} = I - U(U^T U)^{-1} U^T \quad (3)$$

where $I$ denotes the $L \times L$ identity matrix.

Applying $P_U^{\perp}$, under the white-noise assumption (Du et al. 2003), results in the OSP detector

$$\delta_{OSP} = d \;\; P_U^{\perp} r \quad (4)$$

## 2.2 *Constrained Energy Minimization (CEM)*

In order to implement OSP, knowledge of all target signatures of interest is required. Such knowledge is generally difficult to obtain in practice (Chang & Heinz 2000). So, CEM is developed for the case that the only available knowledge is the signature to be classified. The CEM algorithm tries to maximize the response of the target spectral signature while minimizing the response of the unknown background signatures. The algorithm uses an estimate of the sample correlation matrix as a basis for determining the unknown background signatures and is computationally efficient. It also does not assume the linear mixture model or any noise characteristics (Du et al. 2003).

The CEM detector can be given by

$$\delta_{CEM}(r) = (d^T R_{L\times L}^{-1} d)^{-1} (R_{L\times L}^{-1} d)^T r \quad (5)$$

where $R$ is the sample correlation matrix.

The CEM generally outperforms the OSP in terms of eliminating unidentified signal source and suppressing noise, but it has a poor generalization property since it is very sensitive to the knowledge used for the desired target as well as the noise (Chang & Heinz 2000).

## 2.3 *Target Constrained Interference Minimized Filter (TCIMF)*

Compared to the OSP that only deals with desired and undesired signal sources and the CEM that only considers the desired signal source without taking into account the undesired ones, the TCIMF combines both the OSP and the CEM into one filter operation. Interestingly, the CEM and the TCIMF can be also interpreted as various versions of the OSP operating different degrees of target knowledge (Chang 2005).

Consider $D = [d_1, d_2, \ldots, d_p]$ as the desired target signatures and $U = [u_1, u_2, \ldots, u_q]$ be the undesired target signature matrix, the TCIMF detector can be stated as

$$\delta_{TCIMF}(r) = (w^{TCIMF})^T r \quad (6)$$

where

$$w_{TCIMF} = R_{L\times L}^{-1}[DU]([DU]^T R_{L\times L}^{-1}[DU])^{-1} \begin{bmatrix} 1_{p\times 1} \\ 0_{q\times 1} \end{bmatrix} \quad (7)$$

## 2.4 *Fully Constrained Least Squares (FCLS)*

To make the LMM have physical meaning, two constraints are defined: the abundance nonnegative constraint (ANC) and the abundance sum-to-one constraint (ASC) (Zhang et al. 2010).

$$\sum_{j=1}^{p} \alpha_j = 1 \quad (8)$$

$$\alpha_j \geq 0 \qquad for \qquad 1 \leq j \leq p \quad (9)$$

With these two criteria, exact abundance fractions would be extracted.

What makes the FCLS algorithm different from the previously mentioned algorithms, is that both the nonnegativity and sum-to-one constraints are handled simultaneously (Heinz & Chang 2001). Therefore, the FCLS solution provides abundance estimates that meet the linear mixing model constraints, but does not allow a closed-form mathematical solution due to the non-negativity constraints. Instead, a numerical solution is proposed (Broadwater

& Chellappa 2007). In FCLS, the endmember matrix ($M$) and pixel signatures ($r$) are extended such that:

$$N = \begin{bmatrix} \sigma M \\ 1^T \end{bmatrix} \tag{10}$$

$$S = \begin{bmatrix} \sigma r \\ 1 \end{bmatrix} \tag{11}$$

where $1 = [1, 1, \ldots, 1]^T$ is a $p \times 1$ vector. The parameter σ is a small number (typically $1 \times 10^{-5}$) controlling how close the resulting abundances would sum to one. After the aforementioned procedure, the next step is to minimize the least squares error (LSE) by estimating the non-negative abundance values, which can be expressed as:

$$\text{minimize LSE } \{(N\alpha\text{-}S)^T(N\alpha\text{-}S)\} \text{ subject to } \alpha \geq 0 \tag{12}$$

Using the Lagrange multipliers, the following function can be defined.

$$J = \frac{1}{2}(N\alpha - S)^T(N\alpha - S) + \lambda(\alpha - c) \tag{13}$$

where $c$ is an unknown constant vector and each member of this vector is nonnegative to enforce the nonnegativity constraint.

$$\left(\frac{\partial J}{\partial \alpha}\right)\Big|_{\hat{\alpha}_{FCLS}} \text{ and } \alpha = c.$$

So, the obtained equation contains two unknowns: the abundance estimates and the Lagrange multipliers. Solving for these unknown results in

$$\hat{\alpha}_{FCLS} = (N^T N)^{-1} N^T S - (N^T N)^{-1}\lambda = \hat{\alpha}_{LS} - (N^T N)^{-1}\lambda \tag{14}$$

and

$$\lambda = N^T(S - N\hat{\alpha}_{FCLS}) \tag{15}$$

In (14) $\hat{\alpha}_{LS}$ is unconstrained least squares solution.

Iterating through (14) and (15) provides the numerical solution for the non-negativity constraints. To begin this iterative method, set all the Lagrange multipliers to zero and calculate the abundance using (14). Note that for first iteration unconstrained least squares solution will be used. At each iteration, those indices (Lagrange multipliers) corresponding positive abundance fractions are placed in the passive set P and the remainder abundance fractions are placed in the active set R. Iterate (14) and (15) until all Lagrange multipliers in the passive set are zero and all Lagrange multipliers in the active set are either zero or negative. At this point, the Kuhn-Tucker conditions are satisfied and an optimal mean-squared error solution for the unmixing of the image can be obtained.

## 2.5 *Nonnegatively Constrained Least Squares (NCLS)*

The idea of NCLS is to minimize the LSE by estimating the non-negative abundance values. In other words, to solve the following optimal problem

$$\text{minimize LSE } \{(r - M\alpha)^T(r - M\alpha)\} \text{ subject to } \alpha \geq 0 \tag{16}$$

As can be seen, the solution provided in (14) and (15) accounts for both the nonnegativity and sum-to-one constraints. So, the answer to the NCLS is straightforward. To handle the nonnegativity constraint, the modifications applied to the endmember matrix ($M$) and the pixel signatures ($r$) will be eliminated, i.e., the sum-to-one constraint will be ignored.

## 2.6 *Sum-to-one Constrained Least Squares (SCLS)*

Unlike FCLS and NCLS, the SCLS method produces a closed form solution. The SCLS solves the following optimization problem:

$$\text{minimize LSE}\{(\mathbf{r} - M\alpha)^T(\mathbf{r} - M\alpha)\} \text{ subject to } \Delta = \{\alpha \Big| \textstyle\sum_{j=1}^{p} \alpha_j = 1\} \tag{17}$$

The answer to this problem can be expressed as:

$$\hat{\alpha}_{SCLS}(r) = P^{\perp}_{M,l}\hat{\alpha}_{LS}(r) + (M^T M)^{-1}\mathbf{1}[\mathbf{1}^T(M^T M)^{-1}\mathbf{1}]^{-1} \tag{18}$$

where

$$P_{M,l}^{\perp} = I_{L\times L} - (M^T M)^{-1}\mathbf{1}[\mathbf{1}^T(M^T M)^{-1}\mathbf{1}]^{-1}\mathbf{1}^T \quad (19)$$

and $\mathrm{l} = [\mathrm{l}, \mathrm{l}, \ldots, \mathrm{l}]^T$ is a $p \times 1$ vector.

### 2.7 *Adaptive Matched Subspace Detector (AMSD)*

Another kind of subpixel target detection method is based on a statistical hypothesis test. Adaptive matched subspace detector (AMSD) is such an algorithm that formulates the target and background subspaces and uses the LMM and the generalized likelihood ratio test (GLRT) to separate a probable subpixel target (Manolakis et al. 2001). AMSD uses the LMM; however, the sum-to-one and nonnegativity constraints are not satisfied. To develop such algorithm, we establish the following hypothesis tests

$$\begin{aligned} H_0 &: r = Ba_{b,0} + n = Sa_{b,0} + w, w \sim \mathrm{N}(0, \sigma_{\mathrm{w}}^2 I) \\ H_1 &: r = Sa_s + Ba_{b,1} + n = Sa + w \end{aligned} \quad (20)$$

where $H_0$ is the null hypothesis (target absent) and $H_1$ is the alternate hypothesis (target present). In the AMSD, the noise is assumed as a zero mean normal distribution with covariance matrix $\sigma I$. Then, the hypotheses are expressed as

$$\begin{aligned} H_0 &: r \sim N(Ba_{b,0}, \sigma_0^2 I) \\ H_1 &: r \sim N(Sa_s + Ba_{b,1}, \sigma_1^2 I) \end{aligned} \quad (21)$$

In practice, noise variance $\sigma_w^2$ and abundance vector $a$ are unknown and should be estimated from image data using maximum likelihood estimation (MLE) for both the null and alternate hypotheses. Afterwards, taking the derivative of the MLEs with respect to each of the unknown parameters and setting them equal to zero, the MLE estimate of abundance and noise variance are found. After some algebraic manipulations and forming the GLRT, the AMSD detector will be obtained

$$D_{AMSD}(r) = \frac{r^T(P_B^{\perp} - P_Z^{\perp})r}{r^T P_Z^{\perp} r} \quad (22)$$

where $B$ contains the background signatures, and $Z$ is defined as the concatenation of the target and background signatures.

### 2.8 *Adaptive Cosine/Coherent Estimator (ACE)*

We may focus on modeling the background in a stochastic sense. Here, we think of finding a mean centred target in additive background noise. Therefore, the model we assume is

$$\begin{aligned} H_0 &: r \sim N(0, \sigma_0^2 \Gamma) \\ H_1 &: r \sim N(Sa_s, \sigma_1^2 \Gamma) \end{aligned} \quad (23)$$

where $\Gamma = \frac{1}{N}\sum_{i=1}^{N} r(i)r^T(i)$ is the MLE of the covariance matrix. $N$ is the number of background pixels in the image. If we assume that $N$ is very large, the covariance estimate from these likelihoods can be simplified to $\Gamma = \sum_{i=1}^{N} r(i)r^T(i)$, which is a standard assumption made in the literature. Referring to (20), the ACE algorithm sets $B = 0$, therefore the sum-to-one and nonnegativity constraints cannot be met either as they require a background subspace (Broadwater & Chellappa 2007).

The ACE detector can be expressed as

$$D_{ACE}(r) = \frac{r^T \hat{\Gamma}^{-1} S(S^T \Gamma^{-1} S) S^T \Gamma^{-1} r}{r^T \hat{\Gamma}^{-1} r} \quad (24)$$

Despite this seemingly simple background model, the ACE detector is one of the most powerful subpixel detectors available for HSI data (Manolakis & Shaw 2002).

## 3 Decision Fusion

As stated before, in this paper the fusion process is carried out on decision level, so this section gives an insight into the decision fusion methods. Fusion at decision level can be divided into three categories based on the type of the classifier outputs (Xu et al. 1992):

Type 1 (abstract level): given $L$ classifiers each classifier ($D_i$) outputs a class label. So, for

any input object ($x$) to be classified, the $L$ classifiers define a vector $s=[s_1,s_2,...,s_L]^T \in \Omega^L$ where $\Omega=\{\omega_1,\omega_2,...,\omega_c\}$ is the set of class labels. In the simplest fusion scheme, the ensemble chooses the class which receives the highest number of votes (majority voting).

Type 2 (rank level): the output of each $D_i$ is a subset of $\Omega$, with the alternatives ranked in order of plausibility of being correct label.

Type 3 (measurement level) each classifier outputs a c-dimensional vector of measurements $[d_{i,1},d_{i,2},\ldots,d_{i,c}]$, each value of this vector represents how likely a label is.

### 3.1 *Decision Fusion Approaches*

*Product rule*: This rule is good if the individual classifiers are independent, i.e., that the outcomes of $d_{i,j}$ for random $x$ are independent for fixed $i$ (classifier) and variable $j$ (class). This is hardly ever the case.

$$\mu_j(x)=\prod_i^L d_{i,j} \qquad ,j=1,...,c \tag{25}$$

The rule assumes noise free and reliable confidence estimates. It fails if these estimates may be accidentally zero or very small (Duin 2002).

*Mean rule*: In case the base classifiers contain independent noise behaviour, the errors in the confidences are averaged out by the summation (Duin 2002).

$$\mu_j(x)=\frac{1}{L}\sum_{i=1}^L d_{i,j} \qquad ,j=1,...,c \tag{26}$$

*Maximum rule*: At first glance this seems reasonable: select the classifier that is most confident of itself. However, this assumption immediately fails, if some classifiers are more overtrained than others. In that case they may be overconfident and thereby dominating the outcome, without having a better performance (Duin 2002).

$$\mu_j(x)=\max_{i=1}^L d_{i,j} \qquad ,j=1,...,c \tag{27}$$

*Minimum rule*: Like for the maximum rule, a good example of a situation in which this rule is really adequate is hard to find (Duin 2002).

$$\mu_j(x)=\min_{i=1}^L d_{i,j} \qquad ,j=1,...,c \tag{28}$$

*Median rule*: The mean and the median methods have approximately the same performance for normally distributed data but are different for the uniform distribution, the average being the better of the two (Kuncheva 2002).

$$\mu_j(x)=med_{i=1}^L d_{i,j} \qquad ,j=1,...,c \tag{29}$$

### 3.2 *Proposed Fusion Scheme*

In real world hyperspectal applications especially target detection, training data are limited or it is expensive to collect such data for all phenomena. In particular, when spectra of several targets are known and ground truth of only a few are available. So, this motivated us to design a scheme which makes use of available knowledge and is capable to generalize it for other targets.

The block diagram of the proposed fusion scheme is shown in Fig. 1. The proposed framework is as follows. For the detectors which require to know the spectra of all endmembers present in the scene, the HySime (Bioucas-Dias & Nascimento 2008) coupled with the vertex component analysis (VCA) (Nascimento & Bioucas-Dias 2005) are employed to extract such knowledge. In case of using CEM and ACE, the only required knowledge is the targets' spectra. However, due to sensitivity of the CEM detector to noise (Chang & Heinz 2000), which is the function of the sample correlation matrix rank, only a subset of eigenvectors is used to calculate the inverse of sample correlation matrix. The number of eigenvectors is determined through singular value decomposition (SVD) (Chang & Heinz 2000). After generating detection results for each target, the individual decisions are aggregated through the aforementioned decision fusion rules. As shown in (Kuncheva 2002), apart from the fusion methods, diversity among classifiers can provide higher classification accuracies. In case of target detection, corresponding to each decision fusion rule, the most optimum detectors can be chosen using a search strategy. In fact, we aim at improving detection performance through choosing

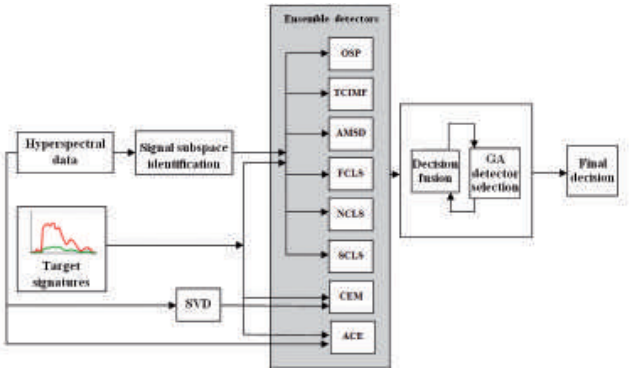



**Fig. 1:** Architecture of the proposed detector fusion scheme, SVD = singular value decomposition, OSP = orthogonal subspace projection, TCIMF = target constrained interference minimized filter, AMSD = adaptive matched subspace detector, FCLS = fully constrained least squares, NCLS = nonnegatively constrained least squares, SCLS = sum-to-one constrained least squares, CEM = constrained energy minimization, ACE = adaptive cosine/coherent estimator.

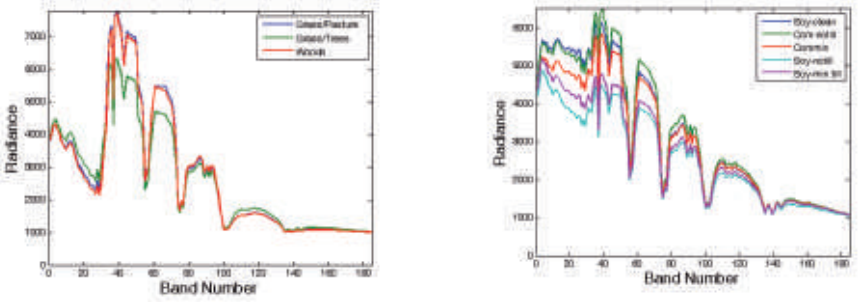
the most effective detectors for each decision rule. A possible way to reach this goal is to take advantage of GA. This approach may allocate distinct banks of detectors for each decision rule through GA.

## 4 Experimental Results and Discussion

### 4.1 *Dataset Description*

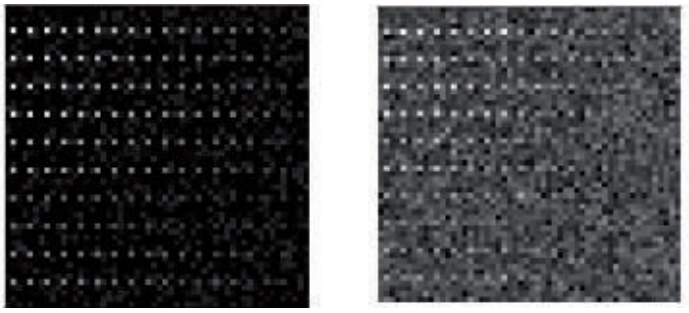
*Synthetic dataset*: To evaluate the subpixel detection performance, a synthetic image consisting of 54 × 54 mixed pixels is simulated. In order to generate the background image, three materials' spectra were selected and mixed according to a Dirichlet distribution. The spectra of both background and target pixels were obtained from Indian Pines dataset (Hsieh & Landgrebe 1998) (after removing noisy and water absorption bands). Fig. 2 shows the spectra of the target and background endmembers. The spectra of soy-clean, corn-notill, corn-mintill, soy-notill and soy-mintill are chosen as desired targets and the spectra of grass/pasture, grass/trees and woods classes are regarded as background endmembers. The synthetic image has 185 spectral bands.

Having generated the background image, a 10 × 18 matrix consisting of target pixels was implanted in the following manner. Two rows of the target matrix correspond to one target. The first column of this matrix replaced the entire background pixel, i.e., first column contains targets with abundance values of 1.




(a) (b)

**Fig. 2:** Spectral signatures of background and target endmembers, (a) background endmembers and (b) target endmembers.



(a) (b)

**Fig. 3:** Synthetic hyperspectral dataset, (a) 180 simulated target panels and (b) band 30 of the synthetic image with SNR 20:1.

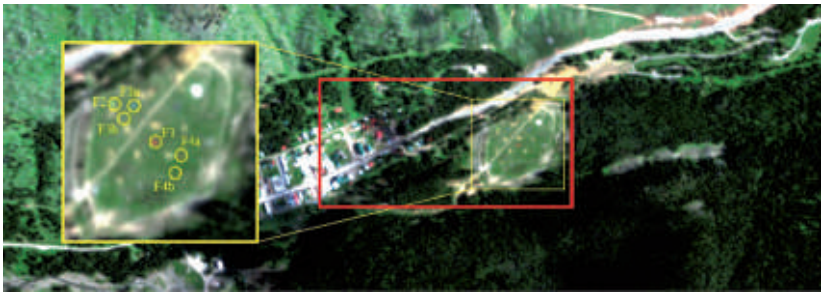


**Fig. 4:** True colour representation of the HyMap dataset; the red region represents the study area.

In the next column, the abundance values of targets decreased to .95 while the abundance values of background pixels increased to .05. By the same token, this pattern continues till the last column, wherein the abundance values of targets reduce to .15. After generating the synthetic image, it was corrupted by Gaussian additive noise to reach a signal-to-noise ratio (SNR) of 20:1. Fig. 3 shows band 30 of the synthetic image.

*Real-world hyperspectral dataset*: The second dataset employed in this work is a real hyperspectral reflectance image of size $120 \times 250$. This image has been extracted from a larger image acquired by the HyMap airborne hyperspectral sensor over Cook City, MT, U.S., on July 4, 2006. The data have a ground sample distance (GSD) of about 3 m and 126 bands in the VNIR-SWIR range. The scene encompasses several types of land cover classes. During acquisition, several targets with sizes comparable to or smaller than the GSD were placed in open grass fields (Tab. 1) and their reflectances were collected. A true colour image of the scene with highlighted target locations can be seen in Fig. 4 (Snyder et al. 2008).

## 4.2 *Organization of the Experiment*

The experiment was organized as follows:

*Signal subspace identification and endmember extraction*: In case of using real-world data, the background endmembers are not known *a priori*. However, some of the aforementioned detectors require this *a priori* knowledge (i.e., background endmembers). So, two key issues have to be addressed. One is the number of endmembers assumed to be present in the hyperspectral data, which is referred to as virtual dimensionality (VD) (Chang & Du 2004). Another is how to estimate these endmembers once the VD is determined. There are some algorithms to estimate these quantities. In this paper, the HySime and VCA were used to estimate the signal subspace dimension and the endmembers' spectra, respectively. Two reasons can justify the use of these two al-

**Tab. 1:** HyMap target characteristics.

| TargetName | Colour | Material | Units | Size |
|---|---|---|---|---|
| **F1** | Red | Cotton | 1 | 3m×3m |
| **F2** | Yellow | Nylon | 1 | 3m×3m |
| **F3** | Blue | Cotton | 2 | 2m×2m, 1m×1m |
| **F4** | Red | Nylon | 2 | 2m×2m, 1m×1m |

gorithms: 1) the HySime method is unsupervised and fully automatic (i.e., it does not depend on any tuning parameters) (Broadwater & Chellappa 2007), and 2) the VCA achieves state-of-the-art performance with a computational complexity between one and two orders of magnitude lower than the best algorithms (Nascimento & Bioucas-Dias 2005).

As stated in section 2, the ACE and CEM algorithms assume no background signatures, which removes the need to extract and identify the proper number of background endmembers. However, as noted in section 2, the CEM approach has a significant shortcoming. It is very sensitive to the noise. As was shown in (Chang & Heinz 2000), the noise sensitivity is closely related to the number of eigenvectors to be used to calculate $R_{L\times L}^{-1}$ and the intrinsic dimensionality of a hyperspectral image, which is usually less than the data dimensionality $L$. If the number of eigenvectors ($q$) is known *a priori* ($q$, for example), we can use SVD so that $R_{L\times L}$ can be reduced to $\tilde{R}_{L\times L}=\tilde{V}\Lambda\tilde{V}^T$, where $\tilde{V}=(\tilde{v}_1,\tilde{v}_2,...,\tilde{v}_q)$ is an eigenmatrix, $\tilde{v}_k$ is the $L$-dimensional vector corresponding to the $k$th eigenvalue $\lambda_k$, and $\Lambda=diag\{\lambda_1,\lambda_2,...,\lambda_q\}$ is a diagonal matrix with eigenvalues as diagonal elements. Using this eigen-decomposition, the inverse of $\tilde{R}_{L\times L}=\tilde{V}\Lambda\tilde{V}^T$ can be found by $\tilde{R}_{L\times L}^{-1}=\tilde{V}\Lambda^{-1}\tilde{V}^T$ (Harsanyi 1993). Criteria for choosing the number of basis vectors are mostly based on the percent of variability explained by the first $q$ vectors. Since the SVD is very efficient in capturing the directions (vectors) explaining most of the variability, a relatively small number of basis vectors tends to explain more than 99.99% of the overall variability (Ientilucci 2005). So, we use the SVD to compute the inverse of sample correlation matrix.

*Target detection*: After endmember extraction, target detection methods will be applied. Before decision fusion the outputs of detectors for each target are made comparable by scaling them to the [0-1] interval.

*Decision fusion*: By combining the individual outputs, we aim at a higher accuracy than that of the best classifier. There is a consensus among the researchers in classifier combination that the major factor for a better accuracy is the diversity in the classifier team and, so, the fusion method is of a secondary importance (Kuncheva 2002). In other words, diversity among the members of a team of classifiers is deemed to be a key issue in classifier combination (Kuncheva & Whitaker 2003). Thus, we suggest using a GA to design a multiple-classifier system corresponding to each of the fusion rules.

In GA, binary coding scheme is used. The chromosome length is determined based on the number of the detectors. The process starts by generating a population equal to the chromosome length based on Goldberg's rule of thumb (Goldberg 1989). For initialization all chromosome bits are set to one. In other words, GA starts using all detectors. Binary tournament selection is applied for parent selection. Two point crossover is used, where the crossover probability and the mutation rate are set to .8 and .03, respectively. As the stopping condition for the algorithm, the number of generations is set to 50.

*Detection performance*: For any given target detector or binary classifier, its performance is described via the receiver operating characteristic (ROC) curve, which plots detection rate versus false alarm rate for all possible thresholds (i.e., [1~0]). Generally, the area under the curve (AUC) is used as a metric to evaluate detection performance independent of threshold selection. Since we have more than two targets (or classes) here, the mean detection rate $R_D$ and mean false alarm rate $R_F$

can be defined by taking the mean of detection rate and false alarm rate over all targets as $\overline{R}_D = \sum_{j=1}^{p} w_j R_{D,j}$ and $\overline{R}_F = \sum_{j=1}^{p} w_j R_{F,j}$, respectively, where $w_j = \frac{N_j}{N_{1,\ldots,p}}$, $N_j$ is the number of $j$-th target pixels and $N_{1,\ldots,p}$ is the total number of all target pixels (Chang et al. 2001). Here, GA aims to maximize the AUC, i.e., the AUC is used as the fitness measure.

### 4.3 *Experimental Results on Synthetic Dataset*

As stated in the previous section, detection performance is described via the AUC. Tab. 2 shows the AUC of target detectors as well as the number of false alarm pixels for the detection probability of .9. It can be induced that in terms of AUC and number of false alarm (FA) pixels, the ACE and CEM algorithms showed the best performance. At first sight, it seems that fusing the results of these detectors with those of others may deteriorate the fusion output; however, fusion results, especially in case of mean and median rules, show the efficiency of the fusion scheme.

Although fusing detection results increases the detection performance, the choice of the detectors to take the results from is also of vital importance. So, in order to choose the most optimum subset of detectors for each fusion rule via GA, the AUC of the first two targets

**Tab. 2:** AUC and number of false alarm pixels (synthetic dataset).

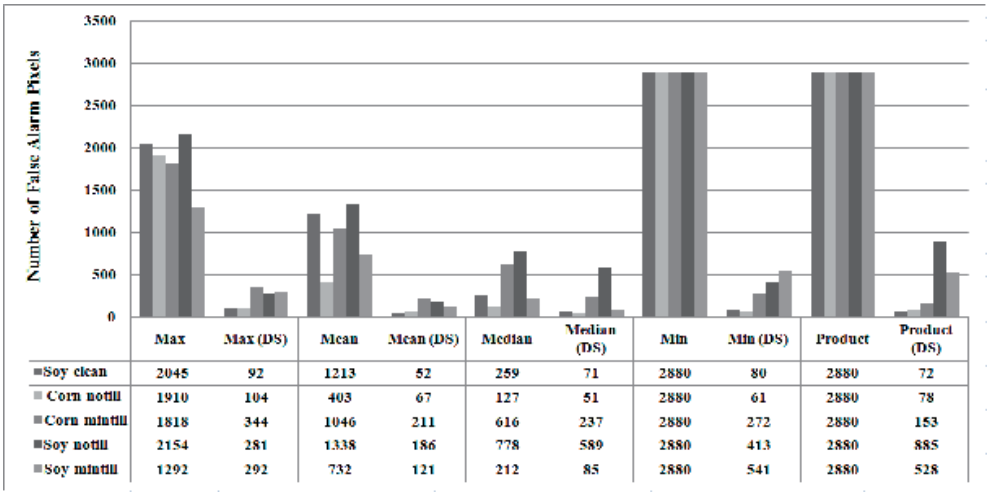
| Detector | Mean AUC | Number of False Alarm Pixels | | | | | Mean False Alarm Pixels |
|---|---|---|---|---|---|---|---|
| | | **Soy-clean** | **Corn-notill** | **Corn-mintill** | **Soy-notill** | **Soy-mintill** | |
| OSP | 0.643 | 1997 | 2175 | 2171 | 2617 | 2118 | 2216 |
| CEM | 0.968 | 92 | 104 | 344 | 281 | 292 | 223 |
| TCIMF | 0.634 | 2025 | 2167 | 2031 | 2536 | 2061 | 2164 |
| ACE | 0.958 | 81 | 61 | 200 | 508 | 812 | 332 |
| AMSD | 0.556 | 2880 | 2880 | 2880 | 2880 | 2880 | 2880 |
| FCLS | 0.875 | 103 | 175 | 2880 | 2880 | 2880 | 1784 |
| NCLS | 0.869 | 211 | 2880 | 2880 | 2880 | 2880 | 2346 |
| SCLS | 0.688 | 1703 | 2146 | 2192 | 2776 | 2059 | 2175 |
| Max (DS*) | 0.968 | 92 | 104 | 344 | 281 | 191 | 223 |
| Mean (DS) | 0.976 | 52 | 67 | 211 | 186 | 121 | 127 |
| Median (DS) | 0.974 | 71 | 51 | 237 | 589 | 85 | 207 |
| Min (DS) | 0.957 | 80 | 61 | 272 | 413 | 541 | 273 |
| Product (DS) | 0.954 | 72 | 78 | 153 | 885 | 528 | 343 |

*detector selection (via GA)

**Tab. 3:** Optimum detectors for each decision fusion rule (synthetic dataset).

| | OSP | CEM | TCIMF | ACE | AMSD | FCLS | NCLS | SCLS |
|---|---|---|---|---|---|---|---|---|
| **Max** | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 |
| **Mean** | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| **Median** | 1 | 1 | 0 | 1 | 0 | 1 | 1 | 1 |
| **Min** | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| **Product** | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 |

**Tab. 4:** Effect of optimum detector selection (synthetic dataset).




| | Max | Max (DS) | Mean | Mean (DS) | Median | Median (DS) | Min | Min (DS) | Product | Product (DS) |
|---|---|---|---|---|---|---|---|---|---|---|
| Soy clean | 2045 | 92 | 1213 | 52 | 259 | 71 | 2880 | 80 | 2880 | 72 |
| Corn notill | 1910 | 104 | 403 | 67 | 127 | 51 | 2880 | 61 | 2880 | 78 |
| Corn mintill | 1818 | 344 | 1046 | 211 | 616 | 237 | 2880 | 272 | 2880 | 153 |
| Soy notill | 2154 | 281 | 1338 | 186 | 778 | 589 | 2880 | 413 | 2880 | 885 |
| Soy mintill | 1292 | 292 | 732 | 121 | 212 | 85 | 2880 | 541 | 2880 | 528 |

(soy-clean and corn-notill) are chosen as the fitness measure. Final chromosomes corresponding to each decision fusion rule are presented in Tab. 3.

It is worth mentioning that the mean rule as the best decision fusion rule does not make use of ACE. Instead it uses the detection outputs of FCLS, which are notably worse than the ACE outputs. Tab. 4 demonstrates the effectiveness of choosing optimum detectors. It can be induced that detector selection strategy causes a dramatic drop in the number of FA pixels in all cases.

### 4.4 *Experimental Results on Real-World Data*

In the case of the HyMap dataset, a limited number of target pixels would prevent a reliable estimate of the detection probability and leads to a sparse ROC curve. Therefore, the detection performance is evaluated by finding the highest test statistics within the region of interest (ROI) and by counting the number of pixels in the image with a higher value. Hence, the resulting value is the number of false alarms arising from the detection of the given target (Matteoli et al. 2011). Therefore, in this case GA aims at minimizing the number of false alarm pixels instead of maximizing the AUC. Here, the sum of FA pixels of targets 1 and 3 is regarded as GA fitness function. As mentioned before, the HySime and VCA methods are used to estimate the signal subspace and endmembers, respectively. The HySime method is applied to the HyMap subset and the signal subspace dimension of 21 is obtained.

As mentioned, the CEM algorithm is sensitive to noise, and the noise sensitivity is closely related to the number of eigenvectors to be used to calculate $\tilde{R}_{L\times L}^{-1}$, so the number of eigenvectors is chosen such that 99.99% of the data variance is maintained.

Quantitative results on the real-world hyperspectral data, containing the number of FA pixels, are tabulated in Tab. 5. It can be observed that regardless of the physical constraints imposed on the FCLS and NCLS detectors, their detection performance is lower than the SCLS algorithm; however, from the material quantification point of view (i.e., estimating true abundance fraction) this might not be true. In the case of material quantification it is necessary to satisfy nonnegativity and sum-to-one constraints. But in the case of target detection the satisfaction of the aforementioned constraints is not necessary. As mentioned in Chang & Heinz (2000) this unconstrained disadvantage turns out to be an advantage in the detection of targets. It also can be inferred that similar to the previous case, the mean fusion rule outperforms the best detector (i.e., OSP).

The comparative results of conventional fusion rules and the proposed method are pre-

**Tab. 5:** Number of false alarm pixels (real-world data).

| Detector | Number of False Alarm Pixels | | | | Mean False Alarm Pixels |
|---|---|---|---|---|---|
| | **1st target** | **2nd target** | **3rd target** | **4th target** | |
| OSP | 2 | 0 | 71 | 23 | 24 |
| TCIMF | 0 | 0 | 1207 | 0 | 302 |
| AMSD | 5 | 0 | 347 | 280 | 158 |
| FCLS | 145 | 646 | 154 | 29990 | 7734 |
| NCLS | 377 | 572 | 342 | 29990 | 7820 |
| SCLS | 1 | 0 | 78 | 40 | 30 |
| ACE | 0 | 0 | 3945 | 0 | 986 |
| CEM | 126 | 447 | 607 | 285 | 366 |
| Max (DS) | 2 | 0 | 82 | 41 | 31 |
| Mean (DS) | 0 | 0 | 64 | 2 | 17 |
| Median (DS) | 0 | 0 | 156 | 6 | 41 |
| Min (DS) | 14 | 1 | 9 | 29990 | 7504 |
| Product (DS) | 3 | 0 | 13 | 29990 | 7502 |

**Tab. 6:** Effect of optimum detector selection (real-world data).

| Fusion rule | Number of False Alarm Pixels | | | | Mean False Alarm Pixels |
|---|---|---|---|---|---|
| | 1st target | 2nd target | 3rd target | 4th target | |
| Max | 5 | 3 | 191 | 5 | 51 |
| Max (DS*) | 2 | 0 | 82 | 41 | 31 |
| Mean | 0 | 0 | 158 | 4 | 41 |
| Mean (DS) | 0 | 0 | 64 | 2 | 17 |
| Median | 0 | 0 | 385 | 2 | 97 |
| Median (DS) | 0 | 0 | 156 | 6 | 41 |
| Min | 14 | 1 | 9 | 29990 | 7504 |
| Min (DS) | 14 | 1 | 9 | 29990 | 7504 |
| Product | 3 | 0 | 13 | 29990 | 7502 |
| Product (DS) | 3 | 0 | 13 | 29990 | 7502 |

*detector selection (via GA)

**Tab. 7:** Optimum detectors for each decision fusion rule (real-world data).

| | OSP | CEM | TCIMF | ACE | AMSD | FCLS | NCLS | SCLS |
|---|---|---|---|---|---|---|---|---|
| **Max** | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| **Mean** | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 1 |
| **Median** | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Min** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Product** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

sented in Tab. 6. The results substantiate the positive impact of detector selection on improving the fusion results. For instance, in the case of the mean rule while taking all detectors into account a mean false alarm rate of 41 is obtained. By selecting the optimum detectors, this number decreases from 41 to 17. However, in the case of min and product rules, detector selection does not affect the final results and all detectors are taken into account. This result is in agreement with what was stated in section 3. In fact, this poor performance is due to wrong assumptions made in these fusion methods.

Comparing the results of the proposed approach with conventional fusion rules in Tabs. 4 and 6 show a drop of the number of false alarm pixels. As stated before, the major factor to obtain a better fusion result is the diversity among classifiers, so it can be implied that the proposed approach fulfills the diversity condition.

Tab. 7 shows the output chromosomes of GA corresponding to each decision fusion rule. It can be seen that just as in the synthetic dataset experiment, the mean rule as the best fusion rule does not necessarily take advantage of all detectors with high performance. It does not take the AMSD results into account, but it chooses the CEM detector instead, which in this case contains higher FA pixels compared to AMSD.

## 5 Conclusions

In this paper, a decision fusion approach for improving hyperspectral subpixel target detection performance was proposed. To illustrate the efficiency of the proposed method, a synthetic and a real-world hyperspectral dataset were used. Since in conventional fusion rules the redundancy in the detectors' outputs may lead to error propagation, the success of each fusion rule is highly dependent on creating diversity using GA. The proposed method showed that decision fusion is an effective yet simple tool to better detection results. Furthermore, experimental results on both datasets showed that the mean fusion rule is a reliable and robust solution to improve detection perfomance. Although this research showed the success of fusion approach to improve detection performance, due to lack of training data in detection applications, further research on developing unsupervised methods to measure the diversity among detectors' outputs is required.

## Acknowledgement

We wish to express our appreciation to Dr. John Kerekes of the Rochester Institute of Technology for providing the HyMap dataset.

## References

Alexandre, L., Campilho, A. & Kamel, M., 2001: On combining classifiers using sum and product rules. – Pattern Recognition Letters **22** (12): 1283–1289.

Bioucas-Dias, J. & Nascimento, J., 2008: Hyperspectral Subspace Identification. – IEEE Transactions on Geoscience and Remote Sensing **46** (8): 2435–2445.

Bro, R. & Jong, S., 1997: A fast non-negativity-constrained least squares algorithm. – Journal of Chemometrics **11** (5): 393–401.

Broadwater, J. & Chellappa, J., 2007: Hybrid detectors for subpixel targets. – IEEE Trans. Geosience Remote Sensing **29** (11): 1891–1903.

Chang, C.I., 2005: Orthogonal Subspace Projection (OSP) Revisited: A Comprehensive Study and Analysis. – IEEE Transactions on Geoscience and Remote Sensing **43** (3): 502–518.

Chang, C.I. & Du, Q., 2004: Estimation of number of spectrally distinct signal sources in hyperspectral imagery. – IEEE Transactions on Geoscience and Remote Sensing **42** (3): 608–619.

Chang, C.I. & Heinz, D., 2000: Constrained Subpixel Target Detection for Remotely Sensed Imagery. – IEEE Transactions on Geoscience and Remote Sensing **38** (3): 1144–1159.

Chang, C.I., Chiang, S.S., Du, Q., Ren, H. & Ifarragaerri, A., 2001: An ROC analysis for subpixel detection. – IEEE 2001 International Geoscience and Remote Sensing Symposium: 2355–2357.

Dasarathy, B., 1994: Decision Fusion. – IEEE Computer Society Press.

Du, Q., Ren, H. & Chang, C.I., 2003: A Comparative Study for Orthogonal Subspace Projection and Constrained Energy Minimization. – IEEE Transactions on Geoscience and Remote Sensing **41** (6): 1525–1529.

Duin, R., 2002: The Combining Classifier: to Train or Not to Train? – Proceedings 16th International Conference on Pattern Recognition **2:** 765–770.

Goldberg, D.E., 1989: Sizing populations for serial and parallel genetic algorithms. – **3rd** International Conference on Genetic Algorithms: 70–79.

Harsanyi, J. & Chang, C.I., 1994: Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach. – IEEE Transactions on Geoscience and Remote Sensing **32** (4): 779–785.

Harsanyi, J.C., 1993: Detection and classification of subpixel spectral signatures in hyperspectral image sequences. – Ph.D. dissertation, Dept. Electrical Eng., Univ. Maryland, Baltimore County.

Heinz, D. & Chang, C.I., 2001: Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. – IEEE Transactions on Geoscience and Remote Sensing **39** (3): 529–545.

Heinz, D., Chang, C.I. & Althouse, M., 1999: Fully constrained least-squares based linear unmixing. – IEEE 1999 International Geoscience and Remote Sensing Symposium: 1401–1403.

Hsieh, P.F. & Landgrebe, D., 1998: Classification of high dimensional data. – Ph.D. dissertation, School of Electrical and Computer Eng., Purdue Univ., West Lafayette.

Ientilucci, E., 2005: Hyperspectral Sub-Pixel Target Detection Using Hybrid Algorithms and and Physics Based Modeling. – Ph.D. dissertation, Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York.

Kraut, S., Scharf, L.L. & Butler, R.W., 2005: The adaptive coherence estimator: A uniformly most-powerful-invariant adaptive detection statistic. – IEEE Transactions on Geoscience and Remote Sensing **53** (2): 427–438.

Kuncheva, L.I., 2002: A Theoretical Study on Six Classifier Fusion Strategies. – IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2): 281–286.

Kuncheva, L.I. & Whitaker, C., 2003: Measures of Diversity in Classifier Ensembles and Their Relationship with the Ensemble Accuracy. – Machine Learning **51** (2): 181–207.

Mangari, U., Samanta, S., Das, S. & Cho, P., 2010: A Survey of Decision Fusion and Feature Fusion Strategies for Pattern Classification. – IETE Technical Review **27** (4): 293–307.

Manolakis, D. & Shaw, G., 2002: Detection Algorithms for Hyperspectral Imaging Applications. – IEEE Signal Processing Magazine **19** (1): 29–43.

Manolakis, D., Marden, D. & Shaw, G.A., 2003: Hyperspectral image processing for automatic target detection applications. – Lincoln Lab. Journal **14** (1): 79–116.

Manolakis, D., Siracusa, C. & Shaw, G., 2001: Hyperspectral subpixel target detection using the linear mixing model. – IEEE Transactions on Geoscience and Remote Sensing **39** (7): 1392–1409.

Matteoli, S., Ientilucci, E.J. & Kerekes, J.P., 2011: Operational and performance considerations of radiative-transfer modeling in hyperspectral target detection. – IEEE Transactions on Geoscience and Remote Sensing **49** (4): 1343–1355.

Nascimento, J. & Bioucas-Dias, J., 2005: Vertex component analysis: a fast algorithm to unmix hyperspectral data. – IEEE Transactions on Geoscience and Remote Sensing **43** (4): 898–910.

Ren, H. & Chang, C.I., 2000: A target-constrained interference-minimized filter for subpixel detection in hyperspectral imagery. – IEEE 2000 International Geoscience and Remote Sensing Symposium.

Snyder, D., Kerekes, J., Fairweather, I., Crabtree, R., Shive, J. & Hager, S., 2008: Development of a web-based application to evaluate target finding algorithms. – IEEE 2008 International Geoscience and Remote Sensing Symposium: 915–918.

Xu, L., Krzyzak, A. & Suen, C., 1992: Methods of Combining Multiple Classifiers and Their Applicaton to Handwriting Recognition. – IEEE Transactions on Systems, Man and Cybernetics **22** (3): 380–384.

Zhang, L., Du, B. & Zhong, Y., 2010: Hybrid Detectors Based on Selective Endmembers. – IEEE Transactions on Geoscience and Remote Sensing **48** (6): 2633–2646.

Addresses of the Authors:

Hamed Gholizadeh & Mohammad Javad Valadan Zoej, Faculty of Geodesy and Geomatics, K.N. Toosi University of Technology, Tehran, P.O. Box 15875-4416, Iran, Tel: +98-21-88786212, Fax: +98-21-88786213; e-mail: h_gholizade@sina.kntu.ac.ir & valadanzouj@kntu.ac.ir.

Barat Mojaradi, School of Civil Engineering, Iran University of Science and Technology, Tehran, P.O. Box: 16765-163, Iran, Tel: +98-21-77240399, Fax: +98-21-77240398, e-mail: mojaradi@iust.ac.ir.

Manuskript eingereicht: November 2011
Angenommen: Januar 2012

PFG 2012 / 3, 0281 – 0298
Stuttgart, June 2012

Article

# Multi-Temporal Hyperspectral and Radar Remote Sensing for Estimating Winter Wheat Biomass in the North China Plain

Wolfgang Koppe, Martin L. Gnyp, Köln, Simon D. Hennig, Friedrichshafen, Fei Li, Yuxin Miao, Xinping Chen, Peking, China, Liangliang Jia, Shijiazhuang, China & Georg Bareth, Köln

**Keywords:** SAR, hyperspectral imaging, vegetation indices, multi-spectral, biomass

**Summary:** This paper illustrates the results obtained in the frame of experimental campaigns carried out on winter wheat fields in the North China Plain from March 2006 to June 2007. Investigations focused on the methodology of estimating biomass on a regional scale with hyperspectral (EO-1 Hyperion) and microwave data (Envisat ASAR). Special importance is drawn to the combined analysis of microwave and optical satellite data for crop monitoring. Since hyperspectral and synthetic aperture radar (SAR) sensors respond to crop characteristics differently, their complementary information content can support the estimation of crop conditions. During the regular field measurements, satellite data from jointing to ripening stages were acquired. Linear regression models between measured surface reflection as well as surface backscatter and wheat's standing biomass were established. For hyperspectral data, the normalized ratio index (NRI) based on 825 nm and 1225 nm wavebands was calculated from 2006 data as input for the regression model. In addition, Envisat ASAR VV polarization data were related to winter wheat crop parameters. Bivariate correlation results from this study indicate that both multi-temporal EO-1 Hyperion as well as Envisat ASAR data provide notable relationships with crop conditions. As expected, linear correlation of hyperspectral data performed slightly better for biomass estimation ($R^2 = 0.83$) than microwave data ($R^2 = 0.75$) for the 2006 field survey. Based on the results, hyperspectral Hyperion data seem to be more sensitive to crop conditions. Improvements for crop parameter estimation were achieved by combining hyperspectral indices and microwave backscatter into a multiple regression analysis as a function of crop parameters. Combined analysis was performed for biomass estimation ($R^2 = 0.90$) with notable improvements in prediction power.

**Zusammenfassung:** *Multi-temporale Hyperspektral- und Radarfernerkundung zur Ableitung von Biomasse des Winterweizens in der nordchinesischen Tiefebene.* Ziel der vorliegenden Studie ist die Betrachtung des Potentials multi-temporaler optischer und Radarfernerkundung zur Ableitung der Biomasse des Winterweizens auf regionaler Ebene. Hierzu wurden in der nordchinesischen Tiefebene in den Wachstumsperioden 2006 und 2007 umfangreiche Feldmessungen von Bestandsparametern während der Satellitenüberflüge durchgeführt. Die verwendeten Satellitendaten sind zum einen Hyperspektraldaten (EO-1 Hyperion) und zum anderen C-Band Radardaten (Envisat ASAR). Neben der separaten Auswertung von Hyperspektral- und SAR-Daten wurde weiterhin das Synergiepotential aus beiden Aufnahmeverfahren betrachtet. Mit Hilfe von linearen Regressionsmodellen zwischen Satellitendaten und Biomasse wurde die Sensitivität hyperspektraler Reflexion und Radarrückstreuung im Hinblick auf das Wachstum des Winterweizens untersucht. Für die Hyperspektraldaten erwies sich der normalized ratio Index (NRI) mit den Wellenlängenbereichen 825 nm und 1225 nm als sensitiv für die Ableitung von Biomasse. Das Modell wurde auf Basis von Daten der Wachstumsperiode 2006 entwickelt und auf die Wachstumsperiode 2007 zur Validierung angewendet. Weiterhin wurde die gemessene Biomasse mit der gleichpolarisierten (VV) C-Band Rückstreuung des Envisat ASAR Sensors linear in Beziehung gesetzt. Als Ergebnis zeigt sich ein deutlicher Zusammenhang zwischen Fernerkundungsdaten und Biomasse, wobei der Regressionskoeffizient deutlich höher für den NRI basierend auf Hyperspektraldaten ($R^2 = 0.83$) ausfällt, als der lineare Zusammenhang mit der Radarrückstreuung ($R^2 = 0.75$). Um den komplementären Informationsgehalt von Hyperspektral- und Radardaten zu nutzen, wurde ein multiples Regressionsmodell erstellt, welches eine Verbesserung der Biomasseschätzung ermöglicht ($R^2 = 0.90$).


www.schweizerbart.de
DOI: 10.1127/1432-8364/2012/0117
1432-8364/12/0117 $ 4.50

## 1 Introduction

China cereal acreage and production is one of the most important in the world, with a crop area of about 88 million ha and production estimated at 483 million tons in 2009, accounting for ca. 22% of total global production (FAO 2011). The North China Plain is one of the most important cereal production regions in China, accounting for almost 50% of China's winter wheat cultivation (National Bureau of Statistics of China 2010). In agricultural issues, timely monitoring of crop growth status at an early stage is important for in-season site specific crop management, detection of plant vitality as well as assessment of seasonal production at local and regional level (Miao et al. 2009, Laudien & Bareth 2006).

The sensitivity of the visible (VIS) and near infrared (NIR) reflectance to water content, pigment of the leaves, as well as leaf structure, permits determining indicators for crop conditions (Kumar et al. 2003). For the linkage of crop parameters with spectral reflectance measurements, a wide range of vegetation indices were developed (Zhao et al. 2007, Broge & Mortensen 2002). Vegetation indices obtained from spectral reflectance measurements are designed to enhance the vegetation cover signal while minimizing the response of various background materials (Schowengerdt 2007).

Numerous studies have successfully related vegetation indices calculated from the visible and near infrared bands of multispectral scanners with crop parameters such as crop vigour (Broge & Leblanc 2000), standing biomass and grain yield (Tucker 1979, Doralswamy et al. 2003). At higher vegetation densities, standard vegetation indices, such as simple ratio (SR) or normalized difference vegetation index (NDVI) are generally less accurate (Jongschaap & Schouten 2005) and tend to saturate (Haboudane et al. 2004, Mutanga & Skidmore 2004), which results in a limited prediction value. To overcome these limitations, hyperspectral vegetation indices for biomass and grain yield prediction have been tested, that are calculated based on all waveband combinations (Xavier et al. 2006, Hansen & Schjoerring 2003). Narrow band vegetation indices other than standard NDVI were successfully used for biomass estimation of winter wheat in the North China Plain (Koppe et al. 2010), which reduces the saturation effect. The same was reported by Mutanga & Skidmore (2004) for pasture biomass estimation. However, Darvishzadeh et al. (2008) and Jamer et al. (2003) demonstrated that biophysical parameters could be better estimated by multivariate methods such as partial least square regression because 2-channel vegetation indices make only use of a small subset of the available spectral information. Beside the advantages and robustness of vegetation indices for biomass estimation, a drawback is the necessity of reference data for model calibration. Alternative and more complex quantification approaches are physically based radiative transfer models (Cho et al. 2008). These models simulate the hyperspectral signature and remotely sensed data can be used for model recalibration (Richter et al. 2009).

Nonetheless, methods based on optical remote sensing are limited in monitoring applications due to cloud cover or daylight dependency. Thus, a monitoring system based only on optical remote sensing would be rather unreliable especially in higher latitudes or in the wet tropics. SAR overcomes the problems of daylight dependency and cloudiness by actively illuminating the earth surface in the wavelength range from a few centimeters to one meter (Brisco & Brown 1998). SAR sensors with their all weather capabilities are good remote sensing sources due to their frequent revisits and sensitivity to surface characteristics (Kugler et al. 2007).

Since the amount of energy backscattered towards the sensor strongly depends on surface roughness and dielectric properties, it is reasonable that SAR can be used for crop type classification, growth stage mapping and biomass monitoring (McNairn & Brisco 2004). For different applications, knowledge of the interaction of the surface characteristics with sensor configurations such as frequency, resolution, incidence angle and polarization is of importance (Inoue et al. 2002). For C-Band SAR measurements, many studies dealt successfully with prediction of standing biomass (Brisco & Brown 1998), but the interpretation of the SAR backscatter has proven to be complicated.

There are also investigations on wheat's biomass retrieval based on spaceborne C-band sensors (Mattia et al. 2003). The results from these studies showed that the backscattering of crops is a complex combination of acquisition parameters (polarization, incidence angle) as well as crop and cultivation characteristics (crop geometry, density, canopy and soil moisture). The combination of these parameters controls the interaction of the incoming electromagnetic wave with the crop canopy and the underlying soil layer.

In the past, quite a few experiments have been performed on wheat fields, either based on spaceborne SAR sensors or on ground-based scatterometers. Satalino et al. (2009) and Brown et al. (2003) acquired C-band spaceborne and scatterometer data over wheat fields and found that wheat biomass is strongly related to HH/VV backscatter during the whole growing season. The good performance of the HH/VV ratio is due to the differently attenuated vertically and horizontally polarized waves that propagate through a mainly vertical medium of wheat (Picard et al. 2003). McNairn et al. (2004) differentiated zones of productivity of wheat fields also using scatterometer data. They reasoned that zones of higher productivity had higher backscatter for linear polarizations, with the greatest contrast for HV.

To move towards an operational crop monitoring approach, it is necessary to mitigate the risk associated with reliance on a single source. To meet these requirements and to improve crop parameter estimation and discrimination, methodologies that integrate optical as well as SAR data were developed (Brisco & Brown 1995). McNairn et al. (2009) performed crop inventories based on multitemporal and multisensoral satellite data with classification improvements of 3 to 18% when adding SAR data to single optical data. Using Envisat ASAR and Landsat TM multispectral images, Liu et al. (2006) achieved a significant improvement for yield estimation by combining both sources in the prediction model.

The synergistic use of different sources was also successfully performed for the integration of hyperspectral optical data with SAR data, mainly for the enhancement of land cover classification based on an image fusion approach at the pixel level (Chen et al. 2003, Chang et al. 2004) and at the feature level (Held et al. 2003).

In the following study, hyperspectral EO-1 Hyperion data and Envisat ASAR data are combined at the feature level to improve crop parameter estimation in North China Plain.

The main objectives of this paper are: (1) to quantitatively describe C-VV SAR backscattering of winter wheat canopy; (2) to investigate the ability to predict crop standing biomass by Envisat ASAR and EO-1 Hyperion data; (3) to explore the potential of complementary use of SAR and hyperspectral data for mapping crop and field conditions at a regional level.

## 2 Material and Methods

### 2.1 *Combined Analysis of ASAR and Hyperion Data*

The plant parameter estimation for winter wheat consists of separate processing chains for SAR and hyperspectral data in a first step and the combined analysis at the feature level (Pohl & van Genderen 1998) as a second step. The generalized processing chain is illustrated in Fig. 1. The processing of ASAR and Hyperion data contains sensor specific pre-processing steps of the multi-temporal datasets. Subsequent to the pre-processing of the raw data, regression models with the ground truth data were established to estimate the predicting power of remotely sensed data for wheat's plant parameters. Correlation coefficients were calculated for individual hyperspectral indices and SAR backscatter to assess the overall information content of the remotely sensed time series. A further field campaign in 2007 with accompanying EO-1 Hyperion data acquisitions allowed a second validation of established hyperspectral regression models for 2006. Finally, the potential of integrating SAR and hyperspectral data is explored at the feature level for model improvements. Therefore, coefficients of multiple correlations were generated involving hyperspectral indices and SAR backscatter.

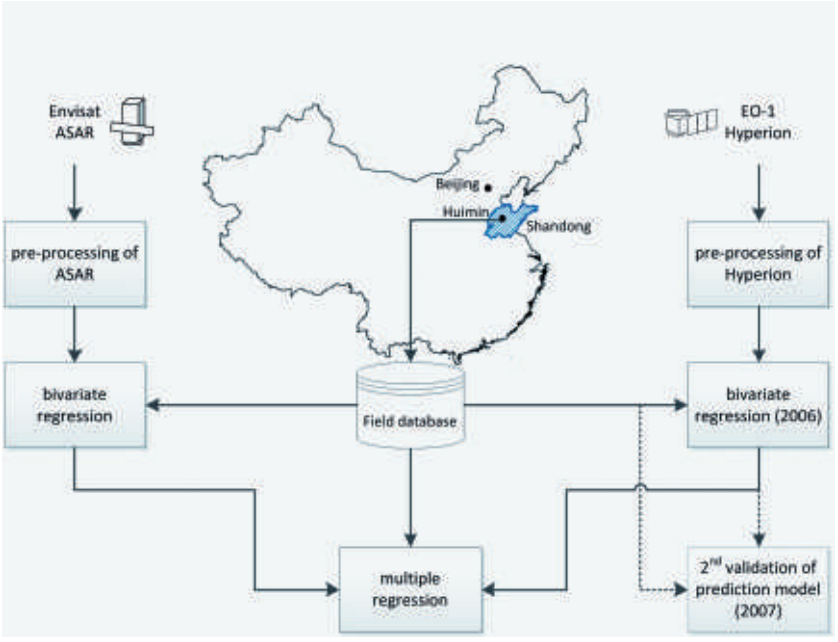



**Fig. 1:** Area of interest and processing flow for SAR- and hyperspectral derived crop condition estimation.

## 2.2 *Study Area*

The test site Huimin County is located in China in the northeast of the Shandong province at around 37.3° N latitude and 117.4° E longitude, which is part of the North China Plain. The area is characterized by a warm temperature sub-humid continental monsoon climate with a mean temperature of around 12.5 °C. The average yearly precipitation sum is approximately 600 mm with a maximum between June and August. The topography of the area is rather flat with an average elevation of around 20 m. In the south of the study site the hills rise up to 400 m.

More than two-thirds of the fields in the study area are cultivated in a crop rotation system with two harvests per year, winter wheat from autumn to June and summer-maize from June to autumn. The cultivation of winter wheat is only possible with irrigation in the dry period of spring.

The investigated crop fields are located around small villages close to the Yellow River in southern Huimin County (Fig. 2). In each year, four fields with an average of around 2.5 to 4.5 ha were selected for regression analysis with satellite data. Two of these four fields were the same in both years. All fields were managed by the farmers in their customary manner. Winter wheat cultivars Jimai20, Jimai21 and Weimai8 were sown from mid of September to beginning of October. Since the fields belong to different farmers, the amount of N-fertilization varies from farmer to farmer. The winter wheat was planted between September and October and harvesting took place between the beginning and mid of June in the following year.

## 2.3 *Ground Truth Data Acquisition*

Spectral reflectance and agronomic parameter measurements were taken throughout the growing period of winter wheat from March to June in 2006 and 2007. The measurements were carried out on a regular basis and were synchronized as far as possible with the satellite overpass of EO-1 and Envisat. Field measurements were carried in four villages, Xili, Xujia, Dongjie and Shizhang, whereas four

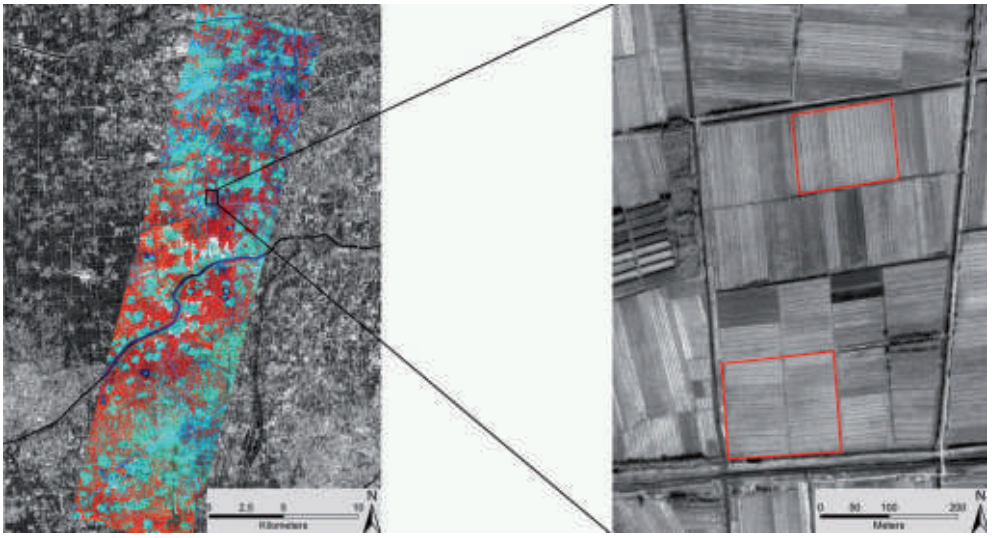


**Fig. 2:** Left: Envisat ASAR VV (10 May 2006) overlaid by EO-1 Hyperion (R: 823 nm, G: 1638 nm, B: 671 nm, April 19, 2006). Right: Two of the four fields used for ground truth measurements (Background: Ikonos pan image, June 20, 2006).

**Tab. 1:** Average soil chemical properties of the four fields.

| Chemical property | Mean | Min | Max | STD |
|---|---|---|---|---|
| Total N (g/kg) | 0.9 | 0.6 | 1.3 | 0.1 |
| Olsen-P (mg/kg) | 32.5 | 11.3 | 52.6 | 11.7 |
| Exchangeable-K (mg/kg) | 120.3 | 65.6 | 222.0 | 24.8 |
| Organic matter (g/kg) | 13.5 | 9.5 | 18.5 | 1.7 |

**Tab. 2:** Crop condition variables mean and standard deviation (STD) measured at different dates in 2006 and 2007.

| Date | Growth stage | n* | Biomass dry (kg/sqm) | | Plant height (cm) | | PWC (%) | |
|---|---|---|---|---|---|---|---|---|
| | | | Mean | STD | Mean | STD | Mean | STD |
| 25 March 2006 | Tillering | 65 | 0.06 | 0.03 | 11.4 | 1.9 | 73.2 | 1.5 |
| 13 April 2006 | Jointing | 58 | 0.21 | 0.05 | 36.8 | 2.8 | 82.3 | 0.6 |
| 19 April 2006 | Jointing | 63 | 0.48 | 0.05 | 45.0 | 4.8 | 83.9 | 1.4 |
| 29 April 2006 | Booting | 64 | 0.74 | 0.09 | 54.9 | 5.6 | 83.5 | 1.2 |
| 10 May 2006 | Heading | 59 | 0.96 | 0.07 | 74.5 | 5.5 | 76.3 | 1.7 |
| 11 April 2007 | Jointing | 67 | 0.43 | 0.04 | 41.4 | 4.8 | - | - |
| 21 April 2007 | Jointing | 67 | 0.62 | 0.07 | 64.6 | 6.0 | - | |
| 6 May 2007 | Booting | 62 | 1.05 | 0.08 | 93.4 | 5.8 | - | - |

* n – measured samples at the different ground truth acquisition days

large fields in total were selected. The fields are each about 2.5 to 4.5 ha in size and subdivided into smaller plots that were managed by different farmers. Two of these fields are shown in Fig. 2 and the average soil chemical properties of the four fields are listed in Tab. 1. In average 63 randomly selected points were sampled at the four fields (around 15 per field) at the different ground truth acquisition days (Tab. 2) to account for spatial variability. At these points, spectral reflection, plant height and destructive biomass measurements were performed. To transfer point measurements to surface data, a continuous surface from the set of points was created by inverse distance weighting method. Taking the different image characteristics of EO-1 Hyperion and Envisat ASAR into account, the sub-plots were prepared in a different way. For EO-1 Hyperion, regression analysis was performed on a pixel basis. After excluding mixed pixel on the field borders, 57 pixel in 2006 and 54 pixel in 2007 remained for further analysis. For Envisat ASAR data, the four fields were separated into homogenous parcels to overcome the speckle effect.

The canopy spectral reflectance measurements were obtained using high resolution spectrometers from Analytical Spectral Devices Inc. (ASD). In 2006, a handheld fieldspec (range: 325 nm to 1075 nm) was used and in 2007 a QualitySpec (range: 350 nm to 1800 nm). As a new spectrometer covering a wider spectral range was available in 2007, this QualitySpec was used for canopy spectral reflectance measurements. In-situ spectral reflectance measurements were acquired for calibration of satellite hyperspectral data. Contemporaneously with the reflectance measurements, agronomic data were collected for the fields as well. Aboveground biomass was taken destructively by cutting vegetation at ground level within an area of 30 cm by 30 cm. The samples were dried at 70 °C to constant weight. As a result of regular ground truth collection, mean and standard deviation (STD) of biomass, plant height and plant water content (PWC, for 2006) at different growth stages are shown in Tab. 2. Because of favourable weather conditions in 2007 the winter wheat's growing season started earlier.

### 2.4 *Pre-Processing of Hyperspectral and SAR Data*

During the first field campaign in 2006, Envisat ASAR data and EO-1 Hyperion data were

**Tab. 3:** Envisat ASAR and EO-1 Hyperion image acquisition dates and associated crop growth stages.

| Satellite | Date of acquisition | Growth stage | Mode | Path direction | Rel. Orbit | Local inc. angle (°) |
|---|---|---|---|---|---|---|
| Envisat ASAR | 25 Mar 2006 | Tillering | IS5 | Ascending | 168 | 39 |
| | 13 April 2006 | Jointing | IS6 | Ascending | 440 | 42 |
| | 29 April 2006 | Booting | IS6 | Ascending | 168 | 39 |
| | 10 May 2006 | Heading | IS6 | Descending | 318 | 40 |
| | 3 June 2006 | Ripening | IS6 | Ascending | 168 | 39 |
| EO-1 Hyperion | 19 April 2006 | Jointing | | | | |
| | 6 May 2006 | Heading | | | | |
| | 31 May 2006 | Ripening | | | | |
| | 11 April 2007 | Jointing | | | | |
| | 21 April 2007 | Booting | | | | |
| | 6 May 2007 | Heading | | | | |

acquired for the test area. For this study, Envisat ASAR IMP VV intensity precision images with a pixel spacing of 12.5 m and a nominal spatial resolution of 30 m were selected. The imaging modes of the ASAR data were acquired in IS5 and IS6 mode with an incidence angle range between 39° to 42°. To get a larger area covered and to shorten time between acquisitions, different relative orbits were chosen. The four test fields for the recent study are located in the overlapping area of all acquisitions.

The Envisat and Hyperion data were acquired at approximately the same time for the area of interest during the 2006 field campaign. The Envisat and Hyperion acquisitions of the multi-temporal stacks overlap at mid of April (jointing stage) and beginning of May (booting to heading stage). Hyperion data provided 242 bands within the spectral range from 0.4 to 2.5 µm. In addition to the growing season of 2006, three EO-1 Hyperion images were acquired during the growing season of 2007 to validate Hyperion image results from the previous year. The list of acquired Envisat ASAR and EO-1 Hyperion images is presented in Tab. 3.

## EO-1 Hyperion Data

The Hyperion sensor is mounted on the Earth Observing One (EO-1) Satellite platform that follows the World Reference System-2 (WRS-2) with a 16 day repeat cycle for nadir mode. It acquires data in a 4.5 km by 100 km footprint with 30 m resolution. Pre-processing of the satellite images from Hyperion is necessary to improve the quality for multi-temporal data analysis. The processing includes the correction of artifacts introduced by the sensor, atmospheric correction and geometric correction (Khurshid et al. 2006).

For Hyperion, a Flag-Mask indicating sensor and processing artifacts was delivered with the data product. During correction of sensor artifacts, 101 of the 242 hyperspectral bands were excluded using ENVI software (ITT Visual Information Solutions 2011).

Atmospheric correction is an important step when using satellite data for multi-temporal analysis and for relating satellite imagery to ground truth data, for instance with data from portable spectrometers. At sensor radiance of EO-1 Hyperion data are conditioned by (a) surface reflectance and (b) by atmospheric scatters, caused by water vapour and aerosols (Cairns 2003). The correction of atmospheric effects was carried out with the MODTRAN-based radiative transfer algorithm (Berk et al. 2000), which converts the at-sensor radiance to surface reflectance using ENVI's FLAASH module (ITT Visual Information Solutions 2011). As the standard atmosphere used in the radiative transfer algorithm does not exactly reflect conditions at the time of acquisition, it was necessary to calibrate initial reflectance with in-situ spectrometer measurements.

Geometric correction of satellite images is necessary for spatially related problems and for overlaying data from different sources. Good orthorectification results were achieved by using the sensor model, ground control points (GCPs) and a digital elevation model (SRTM). In the present work 25 GCPs, which were evenly distributed over the area of interest, were used for the orthorectification. A cubic convolution interpolation with a filter kernel of 3x3 was applied during resampling. The results were validated using 20 independent GCPs, and an overall RMSE of ~ 0.5 pixels (15 m) for the location accuracy after orthorectification was achieved. For subsequent regression analysis with ground truth data, the spectral profile for each pixel of the test fields of 2006 and 2007 was extracted.

## Envisat ASAR Data

Prior to data analysis, Envisat ASAR images were pre-processed in a four step approach. At first, the image DN were converted to averaged backscattering intensity (dB, squared amplitude) expressed in sigma nought. The normalization of the ASAR images facilitated the multitemporal backscatter analysis of the winter wheat. Based on the header file information and an elevation model, the ASAR images can be converted to sigma nought as follows (Rosich & Meadows 2004):

$$\sigma^0_{ij}[dB] = 10 * \log_{10}\left(\frac{DN^2_{ij}}{K}\sin(\alpha_{ij})\right) \qquad (1)$$

Where DN is the pixel intensity of the i, j pixel, K the absolute calibration constant and α the incidence angle of the i, j pixel. Finally, sigma nought values were transformed to logarithmic scale. After normalization, the images were co-registered. The first acquisition was defined as the master and subsequent images were treated as slaves and were co-registered to the master image. Next, speckle noise caused by interference of different elementary scatterers was filtered by a 3x3 gamma adaptive filter which showed acceptable results for the tradeoff between edge preservation and speckle reduction. The reduction of the noise level was evaluated by visual inspection and statistical measurement of effective number of looks (ENL) as suggested by Oliver & Quegan (2004):

$$ENL = \frac{\mu^2}{\sigma^2} \quad (2)$$

Where μ is mean value and σ the standard deviation of the measured area. ENL is obtained by calculating the mean and variance intensity over a homogenous area. The higher the value of the quotient the lower the speckle noise in the area. The gamma map filter with a 3x3 kernel yielded an ENL of 29 compared to around 15 of the noisy images.

In a last step, the co-registered and calibrated image stack was georeferenced using well distributed GCPs. The residual error was 0.6 pixel in range and 0.9 pixel in azimuth direction. The test parcels, for which ground truth measurement was performed, were buffered by a one-pixel zone in order to exclude pixels near the boundaries since they could contain information from neighbouring fields.

In order to investigate the temporal backscatter behaviour of wheat as well as for regression analysis, the four fields were separated into six parcels depending on management practice. The fields were the same as for EO-1 Hyperion data analysis in 2006. To overcome random noise (speckle), pixel values of each parcel (around 60 pixel per parcel) were averaged and the multi-temporal SAR signature was generated. Pre-processing procedure was used to prepare ASAR data for multitemporal regression analysis.

## 3 Results and Discussion

The synergetic use of hyperspectral and radar data for crop monitoring is useful because of its complementary information content. Hyperspectral data provide a high spectral resolution to enhance monitoring of plant biophysical characteristics (Kumar et al. 2003) and SAR data contribute surface texture and dielectric information (Woodhouse 2006) to the combined analysis.

### 3.1 *Temporal Backscatter Behaviour of Winter Wheat*

Averaged backscatter values of Envisat ASAR VV were extracted over crop fields in the test area. Fig. 3 shows the temporal variation of winter wheat and bare field backscattering coefficient $\sigma^0$ of ASAR VV expressed in decibel. In the box plot, the centre horizontal line marks the median of the sample and the length of each box shows the range of the central 50% of the sample. In general, as the crop grows the number of leaves and the stem height increases, resulting in a corresponding increase in ground cover. This causes an increase in volume backscattering due to the increase of canopy constituents of wheat.

At the beginning of the growing season in March, when stem height of wheat was about 9 to 13 cm, the VV backscatter of wheat fields is close to the backscatter from bare soil. The backscatter from ploughed fields is still higher, since soil surface roughness is higher than for the sowed fields. Similar observations were described by McNairn et al. (2009) during the beginning of the growing season. At this growing stage, the backscatter is mainly driven by soil moisture and roughness parameters. When the crop is in the jointing stage on April 13 (stem height between 30 and 40 cm), the backscatter is significantly lower than at the tillering stage (decrease of about 4.5 dB), indicating a strong attenuation of the soil's backscatter by the wheat plants. As the wheat continued to grow through the stages jointing to ripening, the observed backscatter gradually increased by about 3 dB. This suggests a change in the dominant scattering mechanism from soil and roughness backscattering

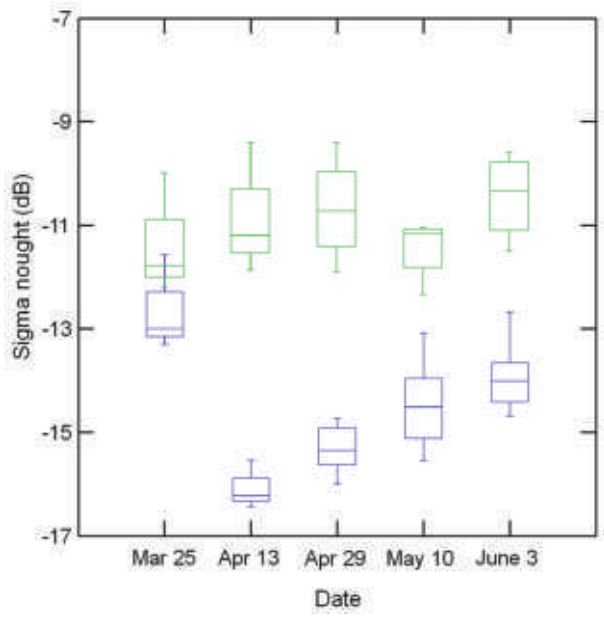



**Fig. 3:** Temporal C-VV backscatter characteristics of bare soil (green) and winter wheat (blue).

to canopy volume scattering, which is in coincidence with the findings of Mattia et al. (2003), who mentioned a change in scattering mechanism from soil to volume scattering for incidence angles greater than 40° of fully developed wheat. Also an integral equation method (IEM) simulation carried out by Stiles et al. (2000) verified a lower sensitivity of higher incidence angles of VV polarization to soil properties.

In comparison to our observation of an increase in backscatter that begins at a stem height of 35 cm for VV polarization, Karjalainen et al. (2008) reported a gradual increase in backscatter for cross-polarization images that does not begin until a stem height of 50 cm is reached.

At the end of the growing season, the water content of plants decreases, while the contribution of soil and surface roughness upon the SAR backscatter increases. As the last Envisat data acquisition was in the ripening stage of the plant, the increase in backscatter from mid of May (heading) to beginning of June (ripening) is also influenced by the soil component. Looking at the backscatter curve of bare fields, a similar increase as for the crop canopy is visible.

The results of field data and backscatter analysis show that the peak of volumetric moisture within the canopy (jointing) did not necessarily coincide with the peak in backscatter (Tab. 2 and Fig. 3). Besides volume scattering, the interactions between ground and stem for fully developed wheat seem to have an important contribution to the cumulative backscatter value, as also mentioned by Picard et al. (2003).

Considering the whole vegetation period, the range of backscatter of about 3 dB is lower than measured by Karjalainen et al. (2008), but comprehensible if findings by McNairn et al. (2004) are taken into account. They detected a higher sensitivity of HV polarization to crop condition than for VV polarization.

During the growing season, the volumetric soil moisture was relatively constant (between 15 and 30%), thus the change in backscatter of about 3 dB on average of the test parcels was most likely caused by the accumulation of aboveground biomass. Despite the usage of different orbits with slightly different local incidence angles, a dependence of the backscattered signal on incidence angle could not be observed. A strong impact on radar backscatter based on incidence angle differences of

about 5–6° was reported by Ban & Howarth (1998).

### 3.2 *Temporal Reflection Behaviour of Winter Wheat*

By using the results from regular in-situ spectral reflectance measurements with QualitySpec (ASD) in 2007, the reflectance characteristic of winter wheat during the growing season was compared to spectral measurements of EO-Hyperion. The energy reflected by plants is correlated with crop conditions such as growth stage and nutrient supply. Throughout the measurements from jointing to ripening stage, the reflectance behaviour alters due to structural changes as well as changes in foliar pigments (Kumar et al. 2003). Up to heading stage the reflectance in the red spectrum (0.65 to 0.7 µm) decreases, which is due to increasing absorption by foliar pigments in the red spectrum. Compared to this, reflectance in the near infrared (NIR) increases caused by the fast development of plants structural components (Thenkabail et al. 2000). The 2006 EO-1 Hyperion data is in good consistence with in-situ field spectrometer measurements (Koppe et al. 2010) due to good atmospheric conditions during the satellite overpass. Fig. 4 shows the spectral average of all fields determined by QualitySpec (red line) and EO-1 Hyperion (blue line) data for 2007. Based on the overlay it can be stated that atmospheric correction performed well; the course of both reflectance curves at a given date are similar in a certain range. Differences can be observed on April 11 and April 21 in the visible and near infrared parts of the spectrum. For these dates hazy conditions with different spatial intensities affected the acquisitions. As a consequence the high absorption (VIS) and maximum reflection (NIR) parts of the spectra are more damped. Especially in the VIS the spectra differ significantly. This part of the electromagnetic spectrum is particulary affected by hazy atmospheric conditions (Kumar et al. 2003), and atmospheric correction of the

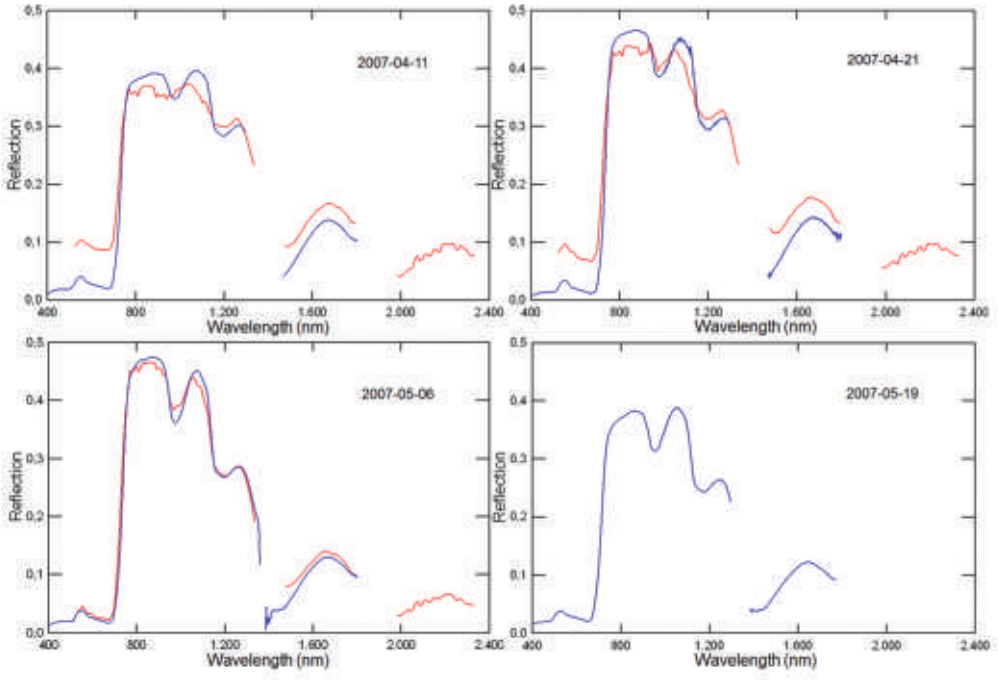



**Fig. 4:** Averaged reflection behaviour of winter wheat recorded by QualitySpec (blue) and EO-1 Hyperion (red) in the growing season of 2007.

scenes phased problems. For the last satellite acquisition at the beginning of May the measured spectra fits to the in-situ measurements.

The maximum reflectance difference between red and NIR is reached at the end of April (booting). With the beginning of May, the divergence of red and NIR reflectance decreases again. A slightly different behaviour of spectral reflectance is recorded between NIR (875 nm) and SWIR (1225 nm). The divergent trend between the two wavelengths continues up to the heading stage (beginning of May) which delays the saturation effect of crop parameter estimation at full canopy cover (Mutanga & Skidmore 2004). At the mid of May the reflection in NIR is strongly decreasing due to senescence (Fig. 4, lower right). The linear relationship of normalized ratio index (NRI) with biomass is lost at this time.

### 3.3 *Relating Plant Parameters to ASAR and Hyperion Images*

Tab. 4 shows the used input data and the relationships between the hyperspectral index (NRI) and ASAR backscatter (C-VV) with standing biomass. Furthermore, the root-mean-square error (RMSE) and relative error (RE) for the regression models are listed. The results of bivariate and multiple correlation analysis will be explained below.

### Relating Crop Condition to Envisat ASAR

The sensitivity of Envisat ASAR signals was analyzed as a function of wheat field characteristics. Based on the averaged backscatter values per parcel and the corresponding measured biomass (dry matter) values, a linear regression equation was derived. The bivariate correlation based on linear regression between SAR backscatter and standing biomass resulted in $R^2 = 0.75$ and RMSE = 0.24 kg/m² (Fig. 5 and Tab. 4). The equation of the regression is:

$$BIOM_{2006} = 1.36(VV_{2006}) - 16.3 \quad (3)$$

The coefficient of determination suggests that Envisat ASAR multitemporal imagery is sensitive to crop condition during the growing season of winter wheat.

In an agricultural environment, the monitoring and prediction of biomass and grain yield is one of the most important objectives. Fig. 5 shows that biomass is significantly positively correlated with C-VV backscatter over the growing season with a dynamic range of around 3 dB. At the different dates, the measured sub-plots vary between three and seven, caused by missing measurements at day of ground truth acquisition. For the relationship of C-VV backscatter to biomass, Mattia et al. (2003) also found an increase in back-

**Tab. 4:** Input data for model calculation and correlation coefficients between standing biomass and satellite data (2006), for 2007 correlation coefficient between predicted biomass and measured biomass.

| Year | Input data typ | Input acquisitions (date) | n* | Correlation coefficients (R2) | RMSE (kg/sqm) | RE (%) |
|---|---|---|---|---|---|---|
| 2006 | hyperspectral | Apr 19, May 6 | 114 | 0.83 | 0.11 | 16.1 |
| 2006 | SAR | Apr 13, Apr 29, 10 May, Jun 3 | 23 | 0.75 | 0.24 | 25.5 |
| 2006 | hyperspectral / SAR | 19 April, 6 May / Apr 13, Apr 29, May 10, Jun 3 | 92 | 0.9 | 0.81 | 12.2 |
| 2007 | hyperspectral | Apr 11, Apr 21, May 6 | 113 | 0.84 | 0.12 | 17.2 |
| | | Apr 11 | 41 | 0.19 | 0.18 | 45.5 |
| | | Apr 21 | 33 | 0.58 | 0.18 | 29.8 |
| | | May 6 | 39 | 0.42 | 0.22 | 18.2 |

* n – number of used samples

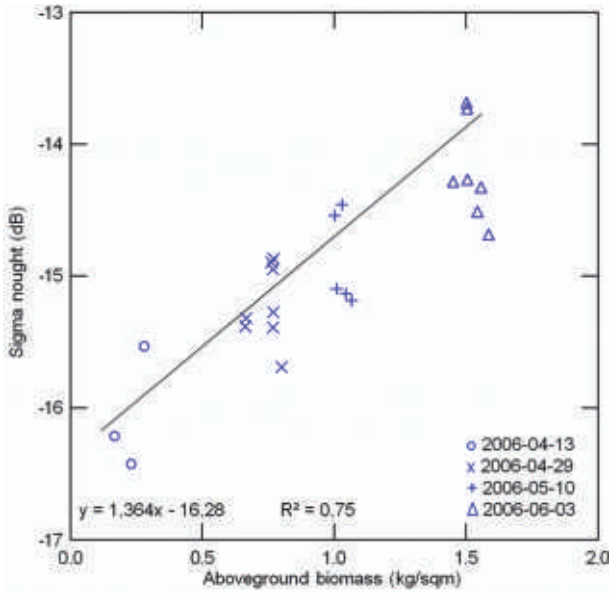



**Fig. 5:** C-VV backscatter related to aboveground biomass.

scatter following the increase in biomass, but only until the heading stage. After the heading stage, backscatter again decreases which was explained by a change in canopy geometry during heading and booting stage (McNairn et al. 2004). At the same phenological stages, Liu et al. (2006) reported a negative correlation of backscatter with biomass and explained this phenomenon also with changes in crop canopy. However, observations of Karjalainen et al. (2008) also show a congenerous increase of backscatter with an accumulation of wheat biomass until the heading stage as observed in this study. Beside the positive multi-temporal relationship between biomass and VV backscatter, it is also obviously that in this case microwave is not suitable for prediction of biomass at a given date (monotemporal). This may be due to less sub-plots which reduces the random sample and dynamic biomass range.

## Relating Crop Conditions to EO-1 Hyperion

Based on the studies of Schowengerdt (2007) and Thenkabail et al. (2000), all possible two-band combinations from the multi-temporal Hyperion data of 2006 (April 19 and May 6) were calculated according to (4). The last acquisition of Hyperion (May 30) was excluded from the analysis, because senesence has already started. The aim was to determine the best two NRI wavebands for wheat biomass estimaton in the North China Plain.

$$NRI_{(band1,band2)} = \frac{(\rho_{band1} - \rho_{band2})}{(\rho_{band1} + \rho_{band2})}$$
$$band1 \geq band2 \qquad (4)$$

For the total number of 9870 calculated NRIs, a correlation matrix between standing biomass and two-band vegetation indices was established. In this correlation matrix, wavelengths with high sensitivity to crop condition were detected. For estimation of standing biomass the waveband of 875 nm combined with 1225 nm proved to be most suitable (Koppe et al. 2010). Similar approaches for different types of vegetation cover showed that band combinations of the red edge (Zhao et al. 2007) as well as NIR or SWIR (Mutanga & Skidmore 2004, Xavier et al. 2006) provided a close relationship with LAI and aboveground biomass. These approaches performed much better than spectral bands used in standard vegetation indices. Based on the best waveband combination, a model for biomass esti-

mation was established which resulted in a coefficient of determination ($R^2$) of 0.83 and a RMSE of 0.11 kg/m$^2$:

$$BIOM_{2006} = 0.25(NRI_{2006}) - 0.01 \quad (5)$$

In order to prove the stability of the established hyperspectral model for EO-1 Hyperion data during the 2006 growing season (5) a validation analysis using EO-Hyperion data for the 2007 growing season was performed. For this, NRI was calculated based on the same waveband combination (875 nm and 1225 nm) that were used in 2006. The resulting NRI values of acquired hyperspectral images in 2007 (April 11, April 21 and May 6) were used to predict biomass at the acquisition dates. The applied equation is the one developed based on the 2006 data.

The validation result of the model is shown in the 1:1 plot in Fig. 6. The model developed for 2006 is able to predict also accumulated biomass in 2007, which is confirmed by a high coefficient of regression of the 1:1 plot ($R^2$ = 0.84). Furthermore, the result is clustered into three separate point clouds that correspond to the different acquisitions of Hyperion data. The acquisitions from April 11 and April 21 show a slight underestimation of the biomass in comparison to the last acquisition of May 6. This could be due to atmospheric distortions (haze) that affect the reflection from the surface. The difference between the used Hyperion wavebands (875 nm and 1225 nm) for the April scenes is lower than the four field measurements (Fig. 4). This lower waveband difference results in a lesser NRI that leads to an underestimation of the standing biomass in Fig. 6.

In spite of haze influence during the acquisition of EO-1 Hyperion data in April, the stability of the model established for 2006 and applied to 2007, can be regarded as applicable for winter wheat on regional scales in the North China Plain. Changing the scale from regional to local scale, a successful inter-year validation of the developed regression model using field spectrometer data from two years and different cultural conditions was performed by (Li et al. 2008). These observations suggest an across scale validity of hyperspectral crop parameter estimation models in the North China Plain.

## Synergy SAR – Hyperspectral

The correlation results based on a linear regression of wheat crop parameters against

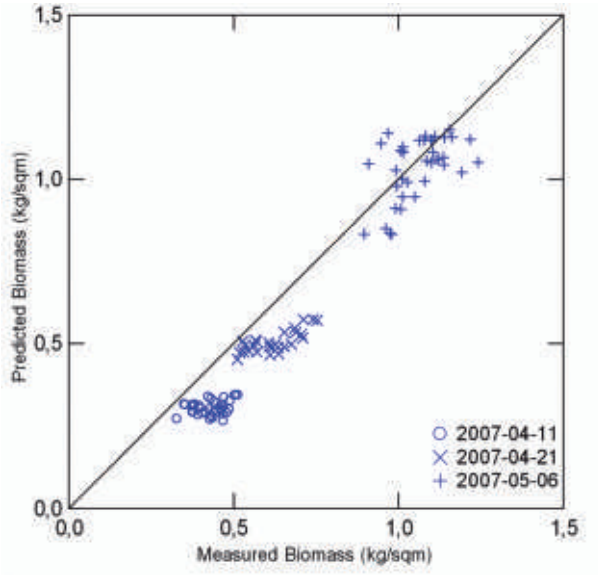



**Fig. 6:** Measured versus predicted aboveground biomass in 2007 using the regression model developed in (5).

ASAR VV backscatter and NRI calculated from Hyperion data in 2006 show, that crop parameters can be predicted by remotely sensed data from the different acquisition systems. Similar close relationships between crop parameters and satellite data has already been established by Thenkabail et al. (2004) and Xavier et al. (2006) for hyperspectral data and furthermore by McNairn et al. (2004) for SAR data.

Best results for bivariate correlation were achieved with narrow band vegetation indices derived from EO-1 Hyperion data ($R^2$ = 0.83). Lower coefficients of determination were achieved by using Envisat ASAR backscatter for correlation with crop parameters for the 2006 SAR campaign ($R^2$ = 0.75). If the models derived from SAR and hyperspectral data are combined based on multiplication, the biomass prediction could be improved to $R^2$ = 0.90. The applied equation is:

$$BIOM_{(Hyp/ASAR)} = 2.73\,(NRI) + 0.16\,(VV) + 2.68 \quad (kg\ m^{-2}\ \ R^2 = 0.90) \tag{6}$$

where $BIOM_{(Hyp/ASAR)}$ is the combined biomass estimate from optical and SAR image analyses, NRI is the normalized ratio index derived from (4) and VV is the ASAR C-VV backscatter. Beside the multiple analysis based on multiplication, there are other types of combination that were not addressed. In this study, the multiple analysis should only demonstrate the improvement of prediction power based on multiple data sources. Similar improvements for yield prediction were achieved by combining time series from SAR and optical data (Liu et al. 2006).

## 4 Conclusion

This paper describes multitemporal hyperspectral and C-band radar data processing for monitoring winter wheat growth in the North China Plain. Addressing the objectives of this paper stated in the introduction, the following conclusions may be drawn from the research:

1. To achieve high prediction accuracy of wheat's crop parameters based on a single source (ASAR or Hyperion) and combined analysis of SAR and optical data, it is important to select suitable dates for satellite data acquisition. Based on the satellite multitemporal and multisource data analysis as well as on analysis of field spectrometer data (Li et al. 2008), certain acquisition periods of SAR and hyperspectral data for winter wheat in the North China Plain seem to be convenient. For the acquisition of Hyperion data, the optimized acquisition window is from the beginning of April after tillering to the heading stage before saturation effect of reflectance occurs and flowering begins. In comparison to multispectral imagery, hyperspectral data can improve the performance early and late in the season (Xavier et al. 2006), which enlarges the acquisition window (Strachan et al. 2002). The use of hyperspectral narrow band vegetation indices can reduce saturation effects at the end of the growing season, which was confirmed for winter wheat by Koppe et al. (2010). For SAR data, observations suggest that a relationship between backscatter and crop parameters can be established between jointing and heading/flowering. Almost similar optimal temporal range from tillering to heading stage for crop parameter estimation is reported by Mattia et al. (2003) and Picard et al. (2003).
2. Multitemporal Envisat ASAR VV data at an incidence angle range from 39° to 42° were analyzed as a function of wheat phenological stage. The results show a significant sensitivity to canopy developing stage and a general increasing trend in backscattering with winter wheat growth. However, it is worth mentioning, that there are unsolved problems limiting the value of the results achieved by Envisat ASAR data. First, the resolution of Envisat ASAR is very low compared to the field size, thus only large fields can be included in the analysis. Secondly, the analysis was performed with a small number of fields in a relatively homogenous agricultural environment. Within field and inter-field variation of crop parameters of a specific acquisition date was too low. During the vegetation period, soil moisture remained relatively constant (at least from one acquisition to another), consequently changes in backscatter are due to

crop growth and less influenced by varying soil moisture.

3. Envisat ASAR C-VV data show certain sensitivity to aboveground biomass. Bivariate regression analysis resulted in a coefficient of determination of about 0.75 for biomass. The simplified relationship between backscatter and crop condition is valid from jointing to heading stage. Before jointing and after heading stage, sophisticated models are necessary to separate backscatter contributions to retrieve canopy parameters. For the relationship of hyperspectral data to crop conditions, narrow band normalized ratio indices based on NIR (875 nm) and SWIR (1225 nm) were calculated for the 2006 growing cycle. As expected, hyperspectral indices show a much higher sensitivity to winter wheat conditions than C-VV radar data. Coefficient of determination was 0.83 for aboveground biomass for multi-temporal approach. To test the validity of the hyperspectral bivariate model, the established model of 2006 was applied to 2007 data to predict aboveground biomass. The predicted vs. measured 1:1 plot resulted in a high coefficient of determination ($R^2$ = 0.84), which proves inter-year validity of hyperspectral prediction power. Compared to multi-temporal approach, single date biomass prediction is more difficult as coefficent of determination is between 0.19 and 0.58 and RMSE is high (Tab. 4). This is caused by lower variablity of aboveground biomass across all fields at a given date.
4. In general, optical and SAR data provide complementary information from a vegetated surface. Hyperspectral sensors record surface reflectance in a wide range of the electromagnetic spectrum which provides the opportunity to extract information about crop canopy parameters and processes at pixel level. SAR microwave penetrate the vegetation and backscatter provide information from inside the canopy. Because of inherent speckle, backscatter values must be spatially averaged for regression analysis. Both kinds of information could be analyzed separately or combined on a feature level. In a final investigation, crop parameters were related to combined C-VV backscatter and hyperspectral indices by means of a multiple regression model. For the relationship with biomass and plant height, regression models with coefficients of determination of 0.90 for biomass were established. This is an improvement of around 9% for aboveground biomass in comparison to single source hyperspectral regression model. While performance improvement is not that much of a combined model, the synergism of using complementary systems in monitoring winter wheat is obvious. Improvements are more likely to be found in terms of filling acquisition gaps of optical data due to cloud cover and providing additional information by SAR.

## References


Ban, Y. & Howarth, P.J., 1998: Orbital effects on ERS-1 SAR temporal backscatter profiles of agricultural crops. – International Journal of Remote Sensing **19:** 3465–3470.

Berk, A., Anderson, G.P., Acharya, P.K., Chetwynd, J.H., Bernstein, L.S., Shettle, E.P., Matthew, M.W. & Adler-Golden, S.M., 2000: MODTRAN4 User's Manual. – Hanscom AFB, MA: Air Force Res. Lab.: 1–93.

Brisco, B. & Brown, R.J., 1995: Multidate SAR/TM synergism for crop classification in western Canada. – Photogrammetric Engineering and Remote Sensing **61:** 1009–1014.

Brisco, B. & Brown, R.J., 1998: Agricultural applications with radar. – Henderson, F.M. & Lewis, A.J. (eds.): Principles and Applications of Imaging Radar, John Wiley and Sons, New York.

Broge, N.H. & Leblanc, E., 2000: Comparing predicting power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. – Remote Sensing of Environment **76:** 156–172.

Broge, N.H. & Mortensen, J.V., 2002: Deriving green crop area index and canopy chlorophyll density of winter wheat from spectral reflectance data. – Remote Sensing of Environment **81:** 45–57.

Brown, S.C.M., Quegan, S., Morrison, K., Bennett, J.C. & Cookmartin, G., 2003: High-Resolution Measurements of Scattering in Wheat Canopies Implications for Crop Parameter Retrieval. – IEEE Transactions on Geoscience and Remote Sensing **41:** 1602–1610.

Cairns, B., 2003: Special Issue Papers – Atmospheric Correction and its Application to an Analysis of Hyperion Data. – IEEE Transactions on Geoscience and Remote Sensing **41** (6): 1232–1245.

Chang, Y.L., Han, C.C., Ren, H., Chen, C.T., Chen, K.S. & Fan, K.C., 2004: Data fusion of hyperspectral and SAR images. – Optical engineering **43:** 1787–1797.

Cho, M.A., Skidmore, A.K. & Atzberger, C., 2008: Towards red-edge positions less sensitive to canopy biophysical parameters for leaf chlorophyll estimation using properties optique spectrales des feuilles (PROSPECT) and scattering by arbitrarily inclined leaves (SAILH) simulated data. – International Journal of Remote Sensing **29** (8): 2241–2255.

Chen, C.M., Hepner, G.F. & Forster, R.R., 2003: Fusion of hyperspectral and radar data using the IHS transformation to enhance urban surface features. – ISPRS Journal of Photogrammetry and Remote Sensing **59:** 310–322.

Darvishzadeh, R., Skidmore, A.K., Schlerf, M., Atzberger, C., Corsia, F. & Choa, M., 2008: LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. – ISPRS Journal of Photogrammetry and Remote Sensing **63** (4): 409–426.

Doralswamy, P.C., Moulin, S., Cook, P.W. & Stern, A., 2003: Crop yield assessment from remote sensing. – Photogrammetric engineering & remote sensing **74:** 665–674.

FAO, 2011: FAO statistical Yearbook, 2011: fao.org/economic/ess/ess-publications/ess-yearbook/ess-yearbook2010/en/ (9.11.2011).

Haboudane, D., Miller, J.R., Pattey, E., Zarco-Tejada, P.J. & Strachan, I.B., 2004: Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. – Remote Sensing of Environment **90:** 337–352.

Hansen, P.M. & Schjoerring, J.K., 2003: Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least square regression. – Remote Sensing of Environment **86:** 542–553.

Held, A., Ticehurst, C., Lymburner, L. & Williams, N., 2003: High resolution mapping of tropical mangrove ecosystems using hyperspectral and radar remote sensing. – International Journal of Remote Sensing **24:** 2739–2759.

Inoue, Y., Kurosu, T., Maeno, H., Uratsuka, S., Kozu., T., Dabrowska-Zielinska, K. & Qi, J., 2002: Season-long daily measurements of multifrequency (Ka, Ku, X, C, and L) and full-polarization backscatter signatures over paddy rice field and their relationship with biological variables. – Remote Sensing of Environment **81:** 194–204.

Jamer, T., Kötz, B. & Atzberger, C., 2003: Spektroradiometrische Ableitung biophysikalischer Vegetationsparameter von Weizenbeständen: Vergleichende Untersuchung verschiedener empirisch-statistischer Verfahren. – PFG **2003** (1): 43–50.

ITT Visual Information Solutions, 2011: exelisvis.com/ (15.8.2011).

Jongschaap, R.W.W. & Schouten, L.S.M., 2005: Predicting wheat production at regional scale by integration of remote sensing data with a simulation model. – Agronomy for sustainable development **25:** 481–489.

Karjalainen, M., Kaartinen, H. & Hyyppä, J., 2008: Agricultural Monitoring Using Envisat Alternating Polarization SAR Images. – Photogrammetric Engineering & Remote Sensing **74:** 117–124.

Khurshid, K., Staenz, K., Sun, L., Neville, R., White, H.P., Bannari, A., Champagne, C.M. & Hitchcock, R., 2006: Preprocessing of EO-1 Hyperion data. – Canadian Journal of Remote Sensing **32:** 84–97.

Koppe, W., Li, F., Gnyp, M.L., Miao, Y., Jia, L., Chen, X., Zhang, F. & Bareth, G., 2010: Evaluating Multispectral and Hyperspectral Satellite Remote Sensing Data for Estimating Winter Wheat Growth Parameters at Regional Scale in the North China Plain. – PFG **2010** (3): 167–178.

Kugler, Z., De Grove, T., Brakenridge, T. & Elaine, A., 2007: Towards a Near-Real Time Global Flood Detection System (GFDS). – 10th International Symposium on Physical Measurements and Signatures in Remote Sensing ISPMSRS07, GITV BV (Publ.), Davos, Switzerland, Lemmer – JRC37044.

Kumar, L., Schmidt, K., Dury, S. & Skidmore, A., 2003: Imaging spectrometry and vegetation science. – van der Meer, F.D. & de Jong, S.M. (eds.): Imaging Spectrometry, Dordrecht, The Netherlands.

Laudien, R. & Bareth, G., 2006: Multitemporal Hyperspectral Data Analysis for Regional Detection of Plant Diseases by using a Tractor- and an Airborne-based Spectrometer. – PFG **2006** (3): 217–228.

Li, F., Gnyp, M.L., Jia, L., Miao, Y., Yu, Z., Koppe, W., Bareth, G., Chen, X. & Zhang, F., 2008: Estimating N status of winter wheat using a handheld spectrometer in the North China Plain. – Field Crops Research **106:** 77–85.

Liu, L., Wang, J., Bao, Y., Huang, W., Ma, Z. & Zhao, C., 2006: Predicting winter wheat condi-

tion, grain yield and protein content using multitemporal Envisat ASAR and Landsat TM satellite images. – International Journal of Remote Sensing **27:** 737–753.

Mattia, F., Le Toan, T., Picard, G., Posa, F.I., D'Alessio, A., Notarnicola, C., Gatti, A.M., Rinaldi, M. & Satalino, G., 2003: Multitemporal C-Band Radar Measurements on Wheat Fields. – IEEE Transactions on Geoscience and Remote Sensing **41** (7): 1551–1560.

McNairn, H. & Brisco, B., 2004: The application of C-band polarimetric SAR for agriculture: a review. – Canadian Journal of Remote Sensing **30:** 525–542.

McNairn, H., Hochheim, K. & Rabe, N., 2004: Applying plarimetric radar imagery for mapping the productivity of wheat crops. – Canadian Journal of Remote Sensing **30:** 517–524.

McNairn, H., Champagne, C., Shang, J., Holmstrom, D. & Reichert, G., 2009: Integration of optical and synthetic aperture radar (SAR) imagery for delivering operational annual crop inventories. – ISPRS Journal of Photogrammetry and Remote Sensing **64:** 434–449.

Miao, Y., Mulla, D.J., Randall, G.W., Vetsch, J.A. & Vintila, R., 2009: Combining chlorophyll meter readings and high spatial resolution remote sensing images for in-season site-specific nitrogen management of corn. – Precision Agriculture **10:** 45–62.

Mutanga, O. & Skidmore, A.K., 2004: Narrow band vegetation indices overcome the saturation problem in biomass estimation. – International Journal of Remote Sensing **25**: 3999–4014.

National Bureau of Statistics of China, 2010: China agriculture yearbook. – 258 p., China Agriculture Press.

Oliver, C. & Quegan, S. (ed.), 2004: Understanding Synthetic Aperture Radar Images. – SciTech, NC 27613.

Picard, G., Le Toan, T. & Mattia, F., 2003: Understanding C-Band Radar Backscatter From Wheat Canopy Using a Multiple-Scattering Coherent Model. – IEEE Transactions on Geoscience and Remote Sensing **41:** 1583–1591.

Pohl, C. & van Genderen, J.L., 1998: Multisensor image fusion in remote sensing concepts, methods and applications. – International Journal of Remote Sensing **19:** 823–854.

Richter, K., Atzberger, C., Vuolo, F., Weihs, P. & D'Urso, G., 2009: Experimental assessment of the Sentinel-2 band setting for RTM-based LAI retrieval of sugar beet and maize. – Canadian Journal of Remote Sensing **35** (3): 230–247.

Rosich, B. & Meadows, P., 2004: Absolute Calibration of ASAR Level 1 Products. Technical Note. – http://earth.esa.int/ (5.1.2010).

Satalino, G., Mattia, F., Le Toan, T. & Rinaldi, M., 2009: Wheat Crop Mapping by Using ASAR AP data. – IEEE Transactions on Geoscience and Remote Sensing **47:** 527–530.

Schowengerdt, R.A., 2007: Remote sensing: models and methods for image processing. – 3rd ed., Elsevier Academic Press, Amsterdam, The Netherlands.

Stiles, J., Sarabandi, K. & Ulaby, F., 2000: Electromagnetic scattering from grassland – Part II: Measurement and modeling results. – IEEE Transactions on Geoscience and Remote Sensing **38:** 349–356.

Strachan, I.B., Pattey, E. & Boisvert, J.B., 2002: Impact of nitrogen environmental conditions on corn as detected by hyperspectral reflectance. – Remote Sensing of Environment **80:** 213–224.

Thenkabail, P.S., Smith, R.B. & De Pauw, E., 2000: Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. – Remote Sensing of Environment **71:** 158–182.

Thenkabail, P.S., Enclona, E.A., Ashton, M.S. & van der Meer, B., 2004: Accuracy assessment of hyperspectral waveband performance for vegetation analysis application. – Remote Sensing of Environment **91:** 354–376.

Tucker, C.J., 1979: Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. – Remote Sensing of Environment **8:** 127–150.

Woodhouse, I.H. (ed.), 2006: Introduction to Microwave Remote Sensing. – Taylor & Francis, Boca Raton, New York, USA.

Xavier, A.C., Rudorff, B.F.T., Moreira, M.A., Alvarenga, B.S., de Freitas, J.G. & Salomon, M.V., 2006: Hyperspectral field reflectance measurements to estimate wheat grain yield and plant height. – Science Agriculture **63:** 130–138.

Zhao, D., Huang, L., Li, J. & Qi, J., 2007: A comparative analysis of broadband and narrowband derived vegetation indices in predicting LAI and CCD of a cotton canopy. – ISPRS Journal of Photogrammetry & Remote Sensing **62:** 25–33.

Addresses of the Authors:

Wolfgang Koppe, Astrium GEO-Information Services, D-88039 Friedrichshafen, Tel.: +49-7545-84226, e-mail: wolfgang.koppe@astrium.eads.net or Institute of Geography, GIS & RS Group, University of Cologne, D-50923 Köln.

Simon Hennig, Astrium Geoinformation Services, D-88039 Friedrichshafen, Tel.: +49-7545-82995, e-mail: simon.hennig@astrium.eads.net

Dr. Fei Li, College of Resources and Environmental Sciences, China Agricultural University, CN-

100094, Beijing and College of Ecology and Environmental Science, Inner Mongolia Agricultural University, CN-010019, Hohhot, Tel.: +86-471-4307376, e-mail: cau_lifei@163.com

Martin Gnyp, Institute of Geography, GIS & RS Group, University of Cologne, D-50923 Köln, Tel.: +49-221-470-6620, e-mail: mgnyp1@uni-koeln.de

Dr. Yuxin Miao, Agro-Informatics and Sustainable Development Group, College of Resources and Environmental Sciences, China Agricultural University, CN-100193 Beijing, e-mail: ymiao@cau.edu.cn

Dr. Liangliang Jia, Institute of Agricultural Resources & Environment, Hebei Academy of Agriculture and Forestry Sciences, CN-050051 Shijiazhuang, e-mail: jiall@cau.edu.cn

Prof. Dr. Xinping Chen, Department of Plant Nutrition, College of Resources & Environmental Sciences, China Agricultural University, CN-100094 Beijing, e-mail: chenxp@cau.edu.cn

Prof. Dr. Georg Bareth, Institute of Geography, GIS & RS Group, University of Cologne, D-50923 Köln, Tel.: +49-221-470-6552, Fax: +49-221-470-1638, e-mail: g.bareth@uni-koeln.de

Manuskript eingereicht: Dezember 2011
Angenommen: März 2012

# Hochschulnachrichten

## Leibniz Universität Hannover, Dissertation Petra Helmholz

Frau Dipl.-Ing. Petra Helmholz wurde am 19. 1. 2012 an der Fakultät für Bauingenieurwesen und Geodäsie der Leibniz Universität Hannover mit der Arbeit „*Verifikation von Ackerland- und Grünlandobjekten eines topographischen Datensatzes mit monotemporalen Bildern*“ zum Dr.-Ing. promoviert.

Referent: Prof. Dr.-Ing. habil. Christian Heipke, Leibniz Universität Hannover. Koreferenten: Prof. Dr.-Ing. Markus Gerke, University of Twente, PD Dr.techn. Franz Rottensteiner, Leibniz Universität Hannover. Gutachter: Prof. Clive Fraser, University of Melbourne.

**Kurzfassung:**
In dieser Arbeit wird ein neuer Ansatz zur automatischen Qualitätsbewertung bestehender 2D-Vektordaten vorgestellt, die in Geoinformationssystemen (GIS) gespeichert sind. Die notwendigen Informationen für die Qualitätskontrolle werden mit Hilfe von Bildanalyseverfahren aus aktuellen Fernerkundungsdaten gewonnen. Im Gegensatz zu anderen Verfahren liegt der Schwerpunkt in dieser Arbeit auf der Qualitätsbewertung von GIS-Ackerland- und Grünlandobjekten, die von besonderem Interesse sind, da sie wesentlich zur Nahrungsversorgung der Bevölkerung beitragen. Acker- und Grünland bedecken gleichzeitig den größten Anteil an der Gesamtfläche Deutschlands mit über 50%. Diese Arbeit fokussiert auf die Verifikation als Teil der Qualitätsbewertung.

Das vorgestellte semi-automatische System hat die Verifikation unter Verwendung orthorektifizierter monotemporaler Luft- und Satellitenbilder mit einer geometrischen Auflösung von 0,5 bis 1 m zum Ziel. Spezialisiert ist das System auf GIS, die dem Inhalt und Detaillierungsgrad einer topographischen Karte mittleren Maßstabs entsprechen. Ziel der Verifikation ist es, den Anteil der inkorrekten Objekte im GIS auf höchstens 5 % zu senken. Zeitgleich soll der manuelle Aufwand, der für den menschlichen Bearbeiter zur Qualitätskontrolle notwendig ist, durch die Verwendung des Systems minimiert werden. Dies wird erreicht, indem ein Teil der GIS-Objekte in einem Verifikationsverfahren automatisch verifiziert wird und daher vom Bearbeiter nicht mehr betrachtet werden muss.

Das entwickelte Verifikationsverfahren ist zweistufig. Zunächst wird im ersten Schritt die gemeinsame Klasse Acker-/Grünland von anderen Klassen wie Siedlung, Industrie und Wald mittels einer pixelbasierten Klassifikation getrennt. Wird ein GIS-Acker-/Grünland-Objekt einer anderen Klasse zugewiesen, so wird dieses vom System abgelehnt und muss durch den menschlichen Bearbeiter manuell verifiziert werden. Im zweiten Schritt werden alle anderen GIS-Ackerland- und Grünlandobjekte mittels einer objektbasierten Klassifikation entweder der Klasse Ackerland oder Grünland zugewiesen. Ist eine Klassifikation nicht möglich, wird das GIS-Objekt einer Zurückweisungsklasse zugewiesen. Durch die Kombination eines pixel- und eines objektbasierten Ansatzes sollen die Vorteile beider Verfahren genutzt werden. Nachdem alle GIS-Objekte einer Klasse zugewiesen wurden, wird das Ergebnis mit der Eintragung im GIS verglichen. Stimmt die ermittelte Klasse mit der im GIS überein, wird das GIS-Objekt als richtig angenommen, wenn nicht, wird es verworfen, und eine manuelle Kontrolle des GIS-Objektes durch den menschlichen Bearbeiter ist notwendig.

Der Schwerpunkt der Arbeit liegt auf dem zweiten Schritt. Da in GIS-Ackerland- und Grünlandobjekten häufig mehrere Bewirtschaftungseinheiten (Schläge) vorhanden sind, wird zunächst eine Segmentierung mittels Wasserscheidentransformation durchgeführt. Ziel ist, dass ein Segment einem Schlag entspricht. Die Zugehörigkeit eines Segmentes zur Klasse Ackerland oder Grünland wird anschließend durch eine objektbasierte Klassifikation mittels spektraler, textureller, struktureller und geometrischer Merkmale bestimmt. Als Klassifikator dient das Verfahren der Support Vector Machines. Alle Segmente, die zu klein sind, um klassifiziert zu werden,


DOI: 10.1127/1432-8364/2012/0118 1432-8364/12/0118 $ 4.00

werden einer Zurückweisungsklasse zugewiesen. Sind alle Segmente eines GIS-Objektes einer Klasse zugewiesen, wird die Flächensumme je Klasse berechnet. Ist die Flächensumme der Zurückweisungsklasse größer als die Summe der Flächen, die klassifiziert werden konnten, gilt das GIS-Objekt als nicht klassifizierbar und wird der Zurückweisungsklasse zugeordnet. Anderenfalls wird das GIS-Objekt der Klasse zugeordnet, deren Flächensumme die Mindestkartierfläche überschreitet. Wird von beiden oder von keiner Klasse die Mindestkartierfläche erreicht, wird das GIS-Objekt ebenfalls der Zurückweisungsklasse zugeordnet.

Das automatische Verifikationsverfahren wurde unter verschiedenen Gesichtspunkten an Hand von Beispielen des ATKIS Basis-DLM und eines Schlagkatasters getestet, die insgesamt 3313 GIS-Ackerland- und Grünlandobjekte beinhalten und eine Fläche von über 302 km$^2$ bedecken. Zunächst wurde bei der separaten Betrachtung des ersten Schrittes festgestellt, dass dieser für die Verifikation von GIS-Ackerland- und Grünlandobjekten geeignet ist. Ein ähnliches Ergebnis wurde bei der separaten Betrachtung des zweiten Schrittes erreicht. Nähere Untersuchungen des zweiten Schrittes ergaben, dass der Erfolg der Verifikation vom Zeitpunkt der Aufnahme des Bildes abhängt. Bei der Betrachtung der Merkmalsgruppen für die objektbasierte Klassifikation wurde festgestellt, dass die besten Ergebnisse immer dann erreicht werden können, wenn spektrale und strukturelle Merkmale beteiligt sind. Insgesamt konnte bei der Kombination beider Schritte eine erfolgreiche Verifikation von GIS-Ackerland- und Grünlandobjekten durchgeführt werden. Der Anteil der inkorrekten GIS-Objekte nach dem Verifikationsprozess beträgt weniger als 5% und der Bearbeiter spart mindestens die Hälfte der Zeit bei der Qualitätskontrolle unter Verwendung des Systems verglichen zum Arbeitsaufwand ohne Verwendung des Verifikationssystems.

Die Dissertation ist in der Schriftenreihe „Wissenschaftliche Arbeiten der Fachrichtung Geodäsie und Geoinformatik der Leibniz Universität Hannover“ (ISSN 0174-1454) als Heft Nr. 299 erschienen.

## Karlsruher Institut für Technologie, Dissertation Ulrike Sturm-Hentschel

Frau Dipl.-Geoökol. Ulrike Sturm-Hentschel promovierte am 8. 2. 2012 an der Fakultät für Bauingenieur-, Geo- und Umweltwissenschaften des Karlsruher Instituts für Technologie (KIT) mit der Arbeit über „*Fernerkundungsgestützte Prozessanalyse im Küstenraum Benins*“ zur Dr. rer. nat.

Referent: Prof. Dr. Joachim Vogt, KIT. Koreferenten: Prof. Dr.-Ing. Dr. h.c. Hans-Peter Bähr, i.R., KIT, und Prof. Dr.-Ing. Stefan Hinz, KIT.

**Kurzfassung:**
Küstenräume stellen weltweit die am stärksten besiedelten Gebiete dar. Zahlreiche Prozesse laufen dort mit hoher Dynamik ab und führen zu räumlichen Veränderungen und Nutzungskonflikten. In Entwicklungsländern geschieht dies zumeist ohne staatliche planerische Interventionen oder gar integrierte Managementkonzepte. Im Untersuchungsgebiet – dem Küstenraum des westafrikanischen Landes Benin – wurden im Rahmen der Arbeit Siedlungsentwicklungen, Veränderungen landwirtschaftlicher Nutzflächen und Küstenerosion als besonders bedeutsame Prozesse herausgegriffen und hinsichtlich ihrer räumlichen Entwicklung und Interaktionen detailliert untersucht.

Gerade in Küstenräumen ist, beispielsweise aufgrund des hohen Nutzungsdruckes und der zu erwartenden mittel- und langfristig Klimaänderungen, vorausschauendes Handeln erforderlich, um eine angepasste Entwicklung in Gang zu setzen und Risiken infolge von Übernutzung zu vermindern. Dazu bedarf es verlässlicher raumbezogener Datengrundlagen, an denen es jedoch in Entwicklungsländern oftmals mangelt. Der Fernerkundung kommt daher eine besondere Aufgabe bei der Erfassung räumlicher Veränderungen als Indikatoren raumrelevanter Prozesse zu, da sich die Aufnahmebedingungen am Boden oft schwierig gestalten und für große Räume zu aufwändig sind. Zudem lassen sich mit Hilfe fernerkundlicher Methoden Prozesse auch in der Retrospektive aufdecken. Kontext bilden in tropisch-subtropischen Küstenräumen – wie dem Untersuchungsgebiet – auf der einen

Seite kleinräumige Landnutzungsstrukturen, die oftmals undeutliche Abgrenzungen aufweisen, und auf der anderen Seite heterogene, zeitlich und räumlich lückenhafte Fernerkundungsdatensätze, die den Einsatz algorithmengestützter Ansätze nur begrenzt möglich machen.

In der Arbeit wird ein methodischer Ansatz vorgestellt, der es erlaubt, unter diesen speziellen und doch für Entwicklungsländer typischen Bedingungen raumverändernde Prozesse zu analysieren. Dazu werden beispielhaft Teilräume der Küstenzone Benin untersucht, die einem extremen Nutzungsdruck und -wandel ausgesetzt sind.

Für die Untersuchungsräume liegen optische Fernerkundungsdaten unterschiedlicher räumlicher und spektraler Auflösung vor (PAN-Luftbild-Kontaktabzüge, RGB-Luftbildscans, QuickBird- und SPOT-Satellitendaten für zwei bzw. drei Zeitpunkte innerhalb des Zeitraums 1995–2007). Diese wurden gestützt durch Untersuchungen am Boden ausgewertet. Der Schwerpunkt der Bildauswertung lag auf der visuellen Interpretation. So können zum einen über die heterogenen Daten hinweg vergleichbare Ergebnisse erarbeitet und zum anderen bestmögliche Aussagen über die Verwertbarkeit von Fernerkundungsdaten zur Prozessanalyse getroffen werden. Des Weiteren stellt der durch visuelle Interpretation manuell erstellte Datensatz eine Referenz für algorithmengestützte Klassifizierungsverfahren dar. Am Beispiel von Gebäude- und Küstenflächen wurde ein Verfahren der segmentbasierten Klassifizierung erarbeitet. Für eine quantitative räumliche GIS-Analyse der Einzelzeitpunkte und deren Veränderung wurden die Ergebnisse der visuellen Interpretation herangezogen.

In allen untersuchten Teilräumen zeigt sich hohe Siedlungsaktivität, die basierend auf detektierten Gebäuden hinsichtlich Siedlungsverdichtung, -erweiterung und Zersiedlung sowie die Entstehung neuer Siedlungsbereiche räumlich und quantitativ erfasst wurde. Siedlungsveränderungen wirken direkt und indirekt auf die Verteilung anderer Landbedeckungs- und -nutzungsklassen. Bezüglich landwirtschaftlicher Nutzflächen, die aufgrund ihrer sozio-ökonomischen Relevanz untersucht wurden, konnte Verdrängung innerhalb und im Umkreis von Siedlungen festgestellt und deren Verlagerung u.a. in zuvor ungenutzte bzw. extensiv genutzte Bereiche aufgezeigt werden. Küstenerosion, die als dritter relevanter raumverändernder Prozess identifiziert wurde, vermindert die Flächenverfügbarkeit in Teilbereichen des Küstenraumes zusätzlich. Zur Analyse der Küstenerosion wurden die Küstenlinienindikatoren Wasserlinie, Nass-Trocken-Linie sowie Abbruchkante als Näherungen der Küstenlinie, bzw. die durch sie begrenzten Flächen extrahiert und ihre Verschiebung durchgängig quantifiziert.

Ein besonderes Merkmal der vorliegenden Arbeit ist, dass trotz eingeschränkter Bilddatenlage detaillierte Analysen raumverändernder Prozesse durchgeführt werden konnten und in ihrer räumlichen, sozio-ökonomischen sowie ökologischen Relevanz diskutiert wurden. Für den Küstenraum Benins liegen damit erstmalig quantitative flächenhafte Daten bezüglich Siedlungsaktivität, Landwirtschaftsveränderung und Küstenerosion vor. Die Ergebnisse liefern eine wichtige Basis für (integrierte) Planungsmaßnahmen und können für zukünftige Prognosen herangezogen werden. Mit der beschriebenen Methodik können die Datenreihen und die Prozessanalyse vervollständigt, zeitlich fortgeführt und für andere Gebiete erweitert werden.

Die Arbeit erscheint in der Bibliothek des Karlsruher Instituts für Technologie und kann unter www.bibliothek.kit.edu herunter geladen werden.

## Technische Universität München, Dissertation Stefan Gernhardt

Herr Dipl.-Ing. Stefan Matthias Gernhardt wurde am 26. 9. 2011 an der Fakultät für Bauingenieur- und Vermessungswesen der Technischen Universität München mit der Arbeit „*High Precision 3D Localization and Motion Analysis of Persistent Scatterers using Meter-Resolution Radar Satellite Data*“ zum Dr.-Ing. promoviert.

Referent: Prof. Dr.-Ing. habil. Richard Bamler, Technische Universität München. Koreferenten: Prof. Dr.-Ing. habil. Stefan Hinz,

Karlsruher Institut für Technologie, Prof. Dr.-Ing. Franz J. Meyer, University of Alaska Fairbanks, USA.

**Kurzfassung:**
Hochaufgelöste Radardaten, wie sie beispielsweise von dem deutschen Fernerkundungssatellit TerraSAR-X bereitgestellt werden können, eröffnen neue Möglichkeiten für die so genannte Persistent Scatterer Interferometrie (PSI). Diese ermöglicht v.a. in urbanen Regionen nahezu flächenhafte Messungen der Oberflächendeformation im Bereich von wenigen Millimetern im Jahr. Die Vorteile dieser neuen Klasse von Radardaten für diese Methode werden in der Arbeit untersucht und bewertet. Durch die sehr hohe Auflösung ist – im Vergleich zu Daten bisheriger synthetischer Apertur Radar (SAR) Satelliten, wie z. B. ERS oder ENVISAT – ein sehr großer Anstieg der zur Verfügung stehenden Punktstreuer (sog. Persistent Scatterers, PS) zu erwarten. Daher wird zunächst der Einfluss unterschiedlicher Aufnahmeparameter hinsichtlich des Auftretens der PS untersucht. Anhand von mehreren Beispielen wird gezeigt, dass die Geometrie der Aufnahmekonfiguration sowie die Anordnung und Beschaffenheit der Objekte in der Szene vor der Datenbestellung berücksichtigt werden müssen. Diese Parameter beeinflussen die lokalen PS Punktdichten, welche je nach geometrischer Konstellation zwischen 40.000 und 100.000 PS/km$^2$ variieren. Die Erkenntnisse sind hilfreich für eine optimale Wahl der Aufnahmeparameter bezüglich der Anforderungen der Deformationsmessung.

Im Rahmen der Arbeit werden mehrere Deformationseffekte an Einzelgebäuden in zwei Testgebieten (Berlin und München) untersucht. Des Weiteren wird der Vorteil der hohen PS Flächenabdeckung für eine Einzelobjektüberwachung demonstriert und die hohe Genauigkeit der linearen sowie periodischen Bewegungsschätzung deutlich gemacht. Die Verfügbarkeit von mehreren Datenstapeln des gleichen Gebiets ermöglicht eine Interpretation der Bewegungsrichtungen in horizontaler und vertikaler Richtung. Am Berliner Hauptbahnhof liegen beispielsweise sehr interessante periodische Bewegungsmuster vor, welche auf eine thermische Ausdehnung der Stahlkonstruktion zurückgeführt werden können. Eine Untersuchung der Konstruktion vor Ort zeigt die baulichen Details, die die Bewegung mit den beobachteten Amplituden erklären. Diese erlauben eine thermische Deformation ohne zu einer Zerstörung der Anlage zu führen.

Die theoretische relative Lokalisierungsgenauigkeit der Punktstreuer wird anhand einer Studie von gleichmäßig angeordneten PS an Fassaden überprüft. Diese regelmäßigen Muster entstehen durch die Anordnung von Fensterreihen in den jeweiligen Stockwerken von Gebäuden, welche zu den entsprechenden Signaturen in den Radarbildern führen. Die Standardabweichungen der relativen Positionen solcher Punkte in einer Reihe bestätigen die theoretisch abgeleiteten Werte von 2 cm in "Range"-, 4 cm in "Azimuth"- und 60–140 cm in "Elevation"-Richtung. Die Positionen in Elevationsrichtung sind dabei aufgrund der beschränkten Variation der Satellitenbahnen um Faktor 20–30 schlechter bestimmbar als in Azimuth oder Range.

Auf die 3D-Lokalisation aufbauend wird eine neue Methode zur Fusion von PS Punktwolken präsentiert – mit dem Ziel, die abgebildeten Objekte mit möglichst vollständigen 3D Punktwolken zu repräsentieren. Die Ausgangsdaten werden dabei durch unabhängige PSI Prozessierungen inklusive Geokodierung der Punkte von unterschiedlichen Datenstapeln erzeugt. Der Algorithmus erlaubt eine Zusammenführung von Ergebnissen aus Daten, welche aus beliebigen Flugrichtungen aufgenommen sind, d. h. es wird vor allem eine Fusion von PS aus Stapeln von aufsteigenden und absteigenden Bahnen erzielt. Dadurch kann der Informationsgehalt der Ergebnisse deutlich gesteigert werden. Die unbekannten Höhenzuschläge der Referenzpunkte werden dabei ohne zusätzliche Referenzdaten aus dem jeweiligen Gebiet bestimmt. Der Algorithmus basiert auf einer kleinsten-Quadrate Ausgleichung, in der korrespondierende Punktpaare aus beiden Punktwolken zur Minimierung der jeweiligen Distanzen herangezogen werden. Die selektierten PS werden dabei als identisch angenommen. Nach der Fusion sind bei den meisten Gebäuden PS auf allen Flächen vorhanden. Daher ist eine Einzelobjektüberwachung realisierbar. Zusätzlich

ermöglicht die Fusion von Stapeln entgegengesetzter Flugrichtung eine Trennung der horizontalen und vertikalen Bewegungsanteile. Ein geeigneter, unkomplizierter Algorithmus sowie Ergebnisse eines Testgebiets werden gezeigt. Die Ergebnisse der Interpretation von Bewegungsschätzungen unterschiedlicher Stapel können dadurch validiert und exakt berechnet werden. Mehrere Beispiele verdeutlichen den Gewinn an Information zu Deformationen in städtischen Gebieten.

Die Dissertation ist in der Kollektion „Elektronische Prüfungsarbeiten" der Fakultät für Bauingenieur- und Vermessungswesen als Online-Veröffentlichung erschienen (nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss-20110926-1079592-1-1). Gleichzeitig ist die Arbeit in der Reihe C der Deutschen Geodätischen Kommission bei der Bayerischen Akademie der Wissenschaften (ISBN 978-3-7696-5084-6) unter der Nr. 672 online veröffentlicht (www.dgk.badw.de).

# Berichte von Veranstaltungen

## 5th International Computer Vision Summer School, 11.–16. Juli 2011, Sampieri, Italien

In Sampieri, einem kleinen Fischerdorf an der Südspitze Siziliens, fand vom 11.–16. Juli 2011 nun schon zum fünften Mal die „*International Computer Vision Summer School*" (ICVSS) statt. Sie wird gemeinsam von der Universität Catania und der Universität Cambridge organisiert. Veranstaltungsort war eine weitläufige Hotelanlage, die über ein eigenes Tagungszentrum mit großem Auditorium verfügt. Die Summer School richtet sich an Doktoranden und Wissenschaftler aus aller Welt, die in angenehmer Umgebung eine Woche lang ihr Wissen im Bereich „*Computer Vision*" vertiefen möchten.

Inhaltlich stand die Summer School unter dem Motto „*Registration, Recognition and Reconstruction in Images and Video*". Die von internationalen Experten gehaltenen Vorträge deckten ein breites Spektrum an Themen ab, die von „*3D Reconstruction*" über „*Image Composition and Video Summarization*", „*Tracking*", „*Video Surveillance*", „*Action Recognition*", „*Scene Understanding*" und „*Computational Photography*" bis hin zu „*Sparse Modeling*" und „*Convex and Nonlinear Optimization*" reichten. Zusätzlich zu den einzelnen Fachgebieten präsentierten die Vortragenden auch den aktuellen Stand ihrer eigenen Forschung. Zum Auftakt der Vortragsreihe hielt Steven Seitz (University of Washington) einen sehr interessanten Überblicksvortrag zum Thema „*Landmarks in 3D*

Teilnehmer der Sommerschule

*Computer Vision*“, in welchem er die wichtigsten technologischen und methodischen Entwicklungen vom Stereokomparator bis zur Kinect zusammenfasste.

Neben der Vortragsreihe gab es noch zwei Poster Sessions auf denen etwa 80 Teilnehmer die Möglichkeit nutzten, ihr eigenes Forschungsthema vorzustellen und mit den restlichen Teilnehmern zu diskutieren. Das beste Poster kam von Andreas Wendel (TU Graz) und wurde mit einem Preis ausgezeichnet. Weiterer Bestandteil des Programms war eine „*Reading Group*“, für welche man sich in ein bestimmtes Thema („*Visual Texture*“, „*Visual Feature*“, „*Shape*“, „*Context*“, „*Tradeoffs*“, „*Visual Communication vs. Visual Recognition*“) einlesen konnte mit dem Ziel, die Entwicklung einer bestimmten Idee von ihrer ersten Veröffentlichung bis in die heutige Zeit zu verfolgen und in einem Kurzvortrag zu präsentieren. Der Preis für den besten Beitrag ging an Richard Stebbing (University of Oxford).

Als Ausgleich zum umfangreichen fachlichen Programm fand am Abend des dritten Tages ein Ausflug nach Ragusa Ibla statt. Nach einer Führung durch die Stadt, die zum UNESCO-Weltkulturerbe zählt, wurde zu einem umfangreichen, traditionell sizilianischen Menü geladen, umrahmt von Darbietungen einer lokalen Volkstanzgruppe.

Abschließend lässt sich festhalten, dass die „*International Computer Vision Summer School*“ eine gut organisierte, fachlich hervorragende Veranstaltung in angenehmer Umgebung ist. Vor allem der anregende Austausch zwischen den Teilnehmern und die Möglichkeit, einen Überblick über das weite Spektrum der „*Computer Vision*“ und aktuelle Forschungsarbeiten zu erhalten, machen die Teilnahme lohnenswert. Der Erfolg der Veranstaltung spiegelt sich auch in den Teilnehmerzahlen wider. Auf die ca. 100 verfügbaren Plätze gab es etwa 250 Anmeldungen aus allen Teilen der Welt.

Weitere Informationen, auch über die kommende ICVSS, sind unter www.dmi.unict.it/icvss verfügbar.

Florian Schmidt, Karlsruhe
Martin Weinmann, Karlsruhe

## 11. Oldenburger 3D-Tage, 1.–2. Februar 2012, Oldenburg

Wie jedes Jahr Anfang Februar luden Thomas Luhmann und seine Mitarbeiter vom Institut für Angewandte Photogrammetrie und Geoinformatik der Jade Hochschule zu den Oldenburger 3D-Tagen ein. Die Organisation der 11. Veranstaltung lag nun schon zum zehnten Mal in den Händen von Christina Müller unter der Leitung von Thomas Luhmann und Thomas Kersten (HCU Hamburg), dem Leiter des Arbeitskreises *Optische 3D-Messtechnik* der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation. Diese ist ebenso ein Kooperationspartner der Veranstaltung wie das Institut für Mess- und Auswertetechnik der Jade Hochschule. Außerdem unterstützte das Institut für Innovations-Transfer in Emden die Durchführung.

Nach dem Jubiläum im letzten Jahr war 2012 wieder der bewährte zeitliche Rahmen von zwei Tagen gegeben. Mittwoch früh um neun Uhr ging es los mit der Eröffnungsveranstaltung; wer sich schon am Dienstag in Oldenburg einfand, hatte Gelegenheit beim Vorabendtreffen im *Schwan* sich in gemütlicher Runde auszutauschen und einzustimmen. Nach den Worten von Thomas Luhmann, der als geschäftsführender Direktor die Veranstaltung eröffnete, folgten Grußworte vom Präsidenten der Jade Hochschule, Elmar Schreiber und Frau Johanna Wanka, der Niedersächsischen Ministerin für Wissenschaft und Kultur. Der anschließende Eröffnungsvortrag wurde in diesem Jahr von Konrad Schindler von der ETH Zürich gehalten, der einen Ausblick in die Photogrammetrie im Jahr 2025 wagte.

In der Folge wurden in insgesamt zehn Sessions zu den Themenschwerpunkten *Laserscanning*, *Oberflächenerfassung*, *Licht und Projektion*, *Prüfung und Kalibrierung*, *Dynamische Prozesse*, *Sensoren und Systeme*, *6DOF und Messkonfiguration* sowie *Laserscanning und Photogrammetrie: Kombinierte Verfahren* insgesamt 56 Vorträge gehalten. Dazu kamen noch zwei Ausstellerforen, in denen neueste Produktentwicklungen zum Thema Hard- und Software vorgestellt wurden, und eine Studierendensession, in der Studie-

Christina Müller, Hochschulpräsident Elmar Schreiber, Ministerin Johanna Wanka und Thomas Luhmann.

Konrad Schindler (ETH Zürich) beim Eröffnungsvortrag.

rende der Jade Hochschule Projekt- und Abschlussarbeiten präsentierten.

Ein bekanntes und wiederkehrendes Thema war die Prüfung und Kalibrierung von terrestrischen Laserscannern. Neben mehreren Vorträgen dazu am ersten Nachmittag und intensiven Diskussionen später am Abend traf sich am zweiten Tag der Arbeitskreis „*Prüfung und Kalibrierung von terrestrischen Laserscannern*“. Hier wurde über die Ergebnisse diskutiert, die in einem ersten Workshop zum Thema Scanneruntersuchung 2011 an der HCU Hamburg erreicht wurden sowie über weitere Möglichkeiten zur Einrichtung von Testfeldern und Prüfeinrichtungen. Eventuell ergibt sich die Möglichkeit, im nächsten Jahr eine Messkampagne an der HTW Dresden zu realisieren, aber auch andere Forschungsinstitutionen wie z.B. das Fraunhofer Institut in Rostock, das sich mit TLS im Schiffbau beschäftigt, stellen ihre Prüfeinrichtungen zur Verfügung. Einigkeit herrschte über die Aussage, dass die neueste Generation von Laserscannern heutzutage so genau arbeitet, dass es für Prüfungen immer schwieriger wird, entsprechende Referenzen einzurichten. Abschließend berichtete Frau Sabine Szyler vom Fraunhofer Institut für Fabrikbetrieb und -automatisierung IFF aus Magdeburg über den Industriearbeitskreis „*Laserscanning und Virtual Reality im Anlagenbau*“. Dieser Arbeitskreis beschäftigt sich mit Themen wie z.B. einem standardisierten Workflow von Laserscanprojekten und Kosten-Nutzen-Analysen der Laserscantechnologie im Anlagenbau. Der Arbeitskreis besteht seit 2010, weitere Informationen hierzu finden sich unter www.fasa-ev.de. Interessierte aus der Messtechnik sind im Arbeitskreis herzlich willkommen.

Ein zweites hochaktuelles Thema war die automatische Generierung von Punktwolken aus Bildverbänden. Durch rasant gestiegene Rechnerleistungen in den letzten Jahren werden Algorithmen aus der Computer Vision wie Structure from Motion oder Dense Image Matching zum Werkzeug für Jedermann. Die Software ist häufig Open Source oder Low-Cost, so dass sich heute jeder in der Erstellung von hochaufgelösten 3D-Modellen versuchen kann. Bestehen bleibt die Frage nach der Qualität solcher Modelle. Mehrere Vorträge zeigten nicht nur das Potenzial dieser Algorithmen auf, sondern auch die Grenzen, die hinsichtlich der Genauigkeit noch bestehen. Geodäten und Photogrammeter sind aufgefordert, das entsprechende Fachwissen einzubringen, um diese Modelle auch für Messzwecke zuverlässig einsetzbar zu machen.

Die Vorträge in dem Vortragsblock *Oberflächenerfassung* zeigten deutlich, wo noch Bedarf bei der Parametrisierung von 3D-Modellierungen und bei der hochpräzisen Konturerfassung existiert. Weiterführende Ergebnisse insbesondere bei der Erfassung nichtkooperativer Oberflächen zeigten jedoch das hohe Potenzial, welches in der optischen 3D Messtechnik genutzt werden kann. Neue Sensoren wie z.B. die inzwischen leistungsgesteigerten Time-of-Flight Sensoren, die zwar so heißen, aber mit dem Verfahren der Phasendifferenz arbeiten, wurden intensiv untersucht und mit herkömmlichen Sensoren verglichen.

Impressionen von der Fachausstellung der Oldenburger 3D-Tage 2012.

Zaubertrick bei der Abendveranstaltung mit Mr. Pat-Trick und Kristin Schreyer, HCU Hamburg (alle Fotos: Piet Meyer, IAPG Oldenburg).

Auch hier sind in Zukunft Entwicklungen zu erwarten, die Verbesserungen in Detailauflösung und in der Punktgenauigkeit bewirken sollten.

Für Anwender im Bereich der Bauaufnahme und Dokumentation von archäologischen Szenen wurden interessante Software- und Hardwaresysteme vorgestellt. Bei der Erfassung von 2D-Aufmaßen wird die Echtzeitfähigkeit von Bildaufnahmesystemen zur teilautomatisierten Bestandsaufnahme und Kontrolle genutzt. Leistungsfähige 3D-Drucker ermöglichen später die maßstäblich verkleinerte physikalische Ausgabe der erfassten Szenen. Die schnelle Laser-Projektion statistischer Muster sowie die Anwendung hochfrequenter Speckleprojektion für die photogrammetrische 3D-Objektvermessung sind als absolutes Highlight anzusehen. Durch diese Messsysteme wird die 3D-Vermessung von nicht-texturierten Oberflächen auch mit hohen Bildraten (um 200 fps) möglich. Beiträge aus dem Bereich der Sensorprüfung und Kalibrierung zeigten, dass die Kalibrierung von optischen Messsystemen gerade in Grenzbereichen zwingend erforderlich ist, sei es bei der Nutzung von Endoskopkameras in Fahrzeugsicherheitsversuchen oder bei der Nutzung von Überweitwinkel- und Fisheye-Objektiven in Stereokamerasystemen.

Neben dem Fachprogramm boten die Oldenburger 3D-Tage 2012 wie immer ausreichend Raum für Diskussionen und Kontaktpflege, insbesondere in den großzügig geplanten Kaffee- und Mittagspausen konnten die Exponate der insgesamt 25 Firmenaussteller intensiv besichtigt und Vorführungen der unterschiedlichen Systeme diskutiert werden.

Die 230 Teilnehmer in diesem Jahr gliederten sich auf vier unterschiedliche Bereiche auf, die Hochschulen stellten mit 45% den größten Anteil (davon wiederum 25% Studierende), gefolgt von den Dienstleistungsanbietern mit 22%. Anwender schließen sich mit 21% und Systemherstellern mit 12% an. Das traditionelle Grünkohlessen war wie immer gut besucht, bot reichlich an Speis und Trank mit anregenden Getränken und Gesprächen und Unterhaltung, untereinander oder durch Mr. Pat-Trick, der mit seinen Zaubertricks zu allerlei Spekulationen anregte.

Die Vortragsfolien der Präsentationen können im Internet unter www.jade-hs.de/3dtage im Programm 2012 heruntergeladen werden. Auf die schriftlichen Beiträge muss wie immer noch ein wenig gewartet werden, der Tagungsband 2012 erscheint wie jedes Jahr Anfang Mai.

Dem Team um Thomas Luhmann und Christina Müller sei für die wieder einmal hervorragende Organisation der 3D-Tage herzlichst gedankt. Die 12. Oldenburger 3D-Tage sind für den 13. und 14. Februar 2013 geplant.

Maren Lindstaedt, Hamburg, und
Manfred Wiggenhagen, Hannover

# Veranstaltungsankündigung

## 1. Gemeinsames Arbeitskreis-Treffen des AK „Fernerkundung“ der DGfG und des AK „Interpretation von Fernerkundungsdaten“ der DGPF in Bochum am 4./5.10.2012

Erstmalig werden die Arbeitskreise „Fernerkundung“ der Deutschen Gesellschaft für Geographie (DGfG und „Interpretation von Fernerkundungsdaten“ der Deutschen Gesellschaft für Photogrammetrie, Fernerkundung und Geoinformation (DGPF) ein gemeinsames Arbeitskreistreffen in Bochum in der Zeit vom 4. bis 5. Oktober 2012 abhalten. Das Motto der Veranstaltung lautet: „Fernerkundung – von der wissenschaftlichen Entwicklung zur Praxisreife“. Die zweitägige Veranstaltung beginnt mit zwei praxisnahen Workshops einerseits zur objektorientierten Klassifikation und andererseits zur Change Detection Analyse. Neben regulären Vorträgen werden zwei Keynote-Präsentationen angeboten und eine interaktive Postersession. Das Abendprogramm greift die Bergbaugeschichte im Ruhrgebiet auf.

Folgende Fristen sind vorgesehen:
– Einreichung von Abstracts: 31.5.2012
– Begutachtung der Abstacts: 15.6.2012
– Endgültiges Programm: 15.6.2012
– Registrierung der Autoren: 15.7.2012
– Einreichung “extended abstract“: 31.8.2012
– Registrierung Frühbucher: 15.7.2012

Teilnehmergebühren bei Anmeldung/Bezahlung bis 14.7.2012
– Vollzahler € 60.–
– Studierende/Doktoranden € 30.–
Registrierung nach 15.7.2012: Aufschlag € 10.– (nur vor Ort)

Bankverbindung: Comdirekt Bank, BLZ 20041155, Kontonr. 4683967 (Kontoinhaber: Carsten Jürgens, Verwendungszweck: „AK2012“)

Wir würden uns über eine rege Teilnahme sehr freuen und ermuntern insbesondere den wissenschaftlichen Nachwuchs zur Einreichung von Beiträgen. Die Abstracts sollen eine Länge von ca. 1 Seite (1200 – 1500 Zeichen) aufweisen.

Stellvertretend für die beiden Arbeitskreise,
Carsten Jürgens & Horst Weichelt

Kontakt:
Prof. Dr. Carsten Jürgens
Ruhr-Universität Bochum
Geographisches Institut
AG Geomatik, Professur für Geo-Fernerkundung
Universitätsstrasse 150
44801 Bochum
Tel. 0234-32-23376
carsten.juergens@rub.de
www.ak-fernerkundung.de

# Veranstaltungskalender

## 2012

7.–12. Juni: **WHISPERS 2012**, Workshop on Hyperspectral Imaging and Signal Processing, in **Shanghai**, China.

14.–15. Juni: 8. **GIS-Ausbildungstagung** in **Potsdam**. gis.gfz-potsdam.de/index.php?id=238

4.–6. Juli: **AGIT** 2012 - Symposium und Fachmesse für Angewandte Geoinformatik in **Salzburg**, Österreich. agit.at

22.–27. Juli: **IGARSS 2012**, International Geoscience and Remote Sensing Symposium, in **München**. igarss12.org

24. August–3. September: **XXII ISPRS Congress** 2012 in **Melbourne**, Australien. isprs2012-melbourne.org

4.–5. Oktober: 1. Gemeinsames Arbeitskreis-Treffen des **AK „Fernerkundung" der DGfG** und des **AK „Interpretation von Fernerkundungsdaten" der DGPF** in **Bochum**. ak-fernerkundung.de

7.–13. Oktober: **ECCV 2012**, European Conference on Computer Vision, in **Florenz**, Italien. eccv2012.unifi.it

9.–11. Oktober: **INTERGEO 2012** und **60. Deutscher Kartographentag** und **3. Europäischer Kongress der CLGE** (Comité de Liaison des Géomètres Européens) in **Hannover**. intergeo.de

16.–18. Oktober: 9th **Symposium on Location-Based Services** in **München**. lbs2012.tum.de

## 2013

17. –23. Februar: 17. **Internationale Geodätische Woche** in **Obergurgl**, Österreich. uibk.ac.at/vermessung/veranstaltung/obergurgl.html

27. Februar–1. März: **Dreiländertagung der DGPF, OVG und SGPF** in **Freiburg**. dgpf.de

30. April–2. Mai: 8th **International Symposium on Mobile Mapping Technology** 2013 in **Tainan**, Taiwan. conf.ncku.edu.tw/mmt2013/

25.–30. August: 26th **International Cartographic Conference (ICC)** in **Dresden**. icc2013.org/

2.–6. September: XXIVth **CIPA Heritage Documentation Symposium** in **Straßburg**, Frankreich. cipa.icomos.org

9.–13. September: 54. **Photogrammetrische Woche** in **Stuttgart**. ifp.uni-stuttgart.de/phowo

8.–15. Dezember: **ICCV 2013**, International Conference on Computer Vision, **Sydney**, Australien. iccv2013.org

# Korporative Mitglieder

**Firmen**

AEROWEST GmbH
AICON 3D Systems GmbH
aphos Leipzig AG
Becker GeoInfo GmbH
Bernhard Harzer Verlag GmbH
Blom Deutschland GmbH
Brockmann Consult
bsf swissphoto GmbH
Büro Immekus
CGI Systems GmbH
con terra GmbH
DB Netz AG
DELPHI IMM GmbH
Deutsches Bergbau-Museum
EFTAS Fernerkundung Technologietransfer GmbH
ESG Elektroniksystem- und Logistik-GmbH
Esri Deutschland GmbH
EUROPEAN SPACE IMAGING
Eurosense GmbH
fokus GmbH
g.on experience gmbh
GAF GmbH
GeoCart Herten GmbH
GeoContent GmbH
Geoinform. & Photogr. Engin. Dr. Kruck & Co. GbR
geoplana Ingenieurgesellschaft mbH
GEOSYSTEMS GmbH
GGS - Büro für Geotechnik, Geoinformatik, Service
Hansa Luftbild AG
IGI - Ingenieur-Gesellschaft für Interfaces mbH
ILV Ing.-büro für Luftbildausw. und Vermessung
Imetric 3D GmbH
Infoterra GmbH
INVERS - Industrievermessung & Systeme
ITT Visual Information Solutions Germany
J. Linsinger ZT-GmbH
Jena-Optronik GmbH
Leica Geosystems GmbH
Luftbilddatenbank-Würzburg
Messbildstelle GmbH
Microsoft Photogrammetry
MILAN Geoservice GmbH
M.O.S.S. Computer Grafik Systeme GmbH
PHOENICS GmbH
PMS - Photo Mess Systeme AG
RIEGL Laser Measurement Systems GmbH
RWE Power AG, Geobasisdaten/Markscheidewesen
technet GmbH
TRIGIS Vermessung + Geoinformatik GmbH
Trimble Germany GmbH
trimetric 3D Service GmbH
Wichmann, VDE Verlag GmbH
Z/I Imaging Ltd.

**Behörden**

Amt für Geoinformationswesen der Bundeswehr
Bayerische Landesanstalt für Wald und Forstwirtschaft
Bundesamt für Kartographie und Geodäsie
Bundesmin. für Ernäh., Landw. u. Verbraucherschutz
Hess. LA für Bodenmanagement und Geoinformation
Innenministerium NRW, Gruppe Vermessungswesen
Inst. für Umwelt- und Zukunftsforschung
LA für Geoinformation und Landentwicklung, BW
LA für Vermessung und Geoinformation, Bayern
LB Geoinformation und Vermessung, Hamburg
LB f. Küstenschutz, Nationalpark u. Meeresschutz, SH
Landesvermessung und Geobasisinformation Niedersachsen
Märkischer Kreis, Vermessungs- und Katasteramt
Regierungspräs. Tübingen, Abt. 8 Forstdirektion
Regionalverband Ruhr
Staatsbetrieb Sachsenforst Pirna
Stadt Bocholt, Fachbereich 31
Stadt Düsseldorf, Vermessungs- und Katasteramt
Stadt Köln, Amt für Liegensch., Verm. und Kataster
Stadt Wuppertal, Verm., Katasteramt und Geodaten
Thüringer LA für Vermessung und Geoinformation

**Hochschulen**

BTU Cottbus, Lehrstuhl für Vermessungskunde
FH Frankfurt a.M., FB 1, Studiengang Geoinformation
FH Mainz, Inst. f. Raumbez. Inform.- und Messtechnik
Jade Hochschule, Inst. für Angew. Photogr. und Geoinformatik
HCU HafenCity Universität Hamburg, Geomatik
HfT Stuttgart, Vermessung und Geoinformatik
HS Bochum, FB Vermessung und Geoinformatik
HS Karlsruhe, Fakultät für Geomatik
HTW Dresden, FB Vermessungswesen/Kartographie
LUH Hannover, Inst. für Kartogr. und Geoinformatik
LUH Hannover, Inst. für Photogrammetrie und GeoInformation
MLU Halle, FG Geofernerkundung
Ruhr-Uni Bochum, Geographisches Institut
RWTH Aachen, Geodätisches Institut
TU Bergak. Freiberg, Inst. f. Markscheidewesen u. Geodäsie
TU Berlin, Computer Vision & Remote Sensing
TU Berlin, Inst. f. Geodäsie und Geoinformationstechnik
TU Braunschweig, Inst. für Geodäsie und Photogrammetrie
TU Clausthal, Inst. für Geotechnik und Markscheidewesen
TU Darmstadt, Inst. für Photogrammetrie und Kartographie
TU Dresden, Inst. für Photogrammetrie und Fernerkundung
TU München, FG Photogrammetrie und Fernerkundung
TU Wien, Inst. für Photogrammetrie und Fernerundung
Uni Bonn, Inst. für Photogrammetrie
Uni Göttingen, Inst. für Waldinventur und Waldwachstum
Uni Heidelberg, IWR Interdis. Zentr. f. Wiss. Rechnen
Uni Kassel, FB Ökologische Agrarwissenschaften
Uni Kiel, Geographisches Institut
Uni Stuttgart, Inst. für Photogrammetrie
Uni Würzburg, Geographisches Institut
Uni zu Köln, Geographisches Institut