Optimizing Object, Atmosphere, and Sensor Parameters in Thermal Hyperspectral Imagery

Jörgen Ahlberg

Journal Article

N.B.: When citing this work, cite the original article.

©2016 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Jörgen Ahlberg, Optimizing Object, Atmosphere, and Sensor Parameters in Thermal Hyperspectral Imagery, IEEE Transactions on Geoscience and Remote Sensing, 2017. 55(2), pp.658-670.

<http://dx.doi.org/10.1109/TGRS.2016.2613685>

Postprint available at: Linköping University Electronic Press

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-134597>

Optimizing object, atmosphere, and sensor parameters in thermal hyperspectral imagery

Jörgen Ahlberg

Abstract

We address the problem of estimating atmosphere parameters (temperature, water vapor content) from data captured by an airborne thermal hyperspectral imager, and propose a method based on linear and non-linear optimization. The method is used for estimation of the parameters (temperature and emissivity) of the observed object, as well as for sensor gain under certain restrictions. The method is analyzed with respect to sensitivity to noise and number of spectral bands. Simulations with synthetic signatures are performed to validate the analysis, showing that estimation can be performed with as few as 10–20 spectral bands at moderate noise levels. The proposed method is also extended to exploit additional knowledge, for example measurements of atmospheric parameters and sensor noise. Additionally, we show how to extend the method in order to improve spectral calibration.

Index Terms

remote sensing, infrared imaging, infrared spectroscopy, optimization methods

I. INTRODUCTION

Hyperspectral electro-optical sensors (cameras) sample the incoming radiation at many, sometimes several hundreds of, spectral bands. Each pixel thus forms a vector of measurements from the different bands. This vector, the observed spectral signature, contains information of the material(s) present in the scene, and can be exploited for detection, classification, and recognition.

At some wavelengths, the influence of the atmosphere between the observed object and the sensor is significant, and it is often therefore necessary to perform atmospheric correction before further processing [1]. The effects of the atmosphere are typically known to some degree; there are standard simulation tools that provide good approximations. This paper treats methods for refining these approximations from observed spectral signatures of unknown objects, i.e., in-scene atmospheric estimation. The refinement, in form of an optimization process, includes estimating the emissivity and temperature of the observed objects under the graybody assumption.

A. Setting and problem formulation

A specific setting is considered in this paper, which leads to assumptions in the treated estimation problems. In that setting, a reconnaissance aircraft carries a nadir-looking sensor, and we want to analyze the observed objects

J. Ahlberg was with FOI, the Swedish Defence Research Agency, SE-581 11 Linköping, Sweden. He is now with the Computer Vision Laboratory, Dept. of Electrical Engineering, Linköping University, SE-581 83 Linköping, Sweden. E-mail: jorgen.ahlberg@liu.se.

on the ground. The sensor is a hyperspectral sensor in the long-wave infrared (LWIR) domain. Knowing the time, date, and location, we can obtain humidity and temperature profiles, aerosol and carbon dioxide contents in standard atmospheric models. However, the variability of some of these parameters (humidity and temperature) can drastically influence the atmosphere's radiative properties.

In order to retrieve the spectral signature of an observed object, we need to solve, or approximate, the temperature-emissivity separation (TES) problem [2]. A first step in the TES process is to correct for the atmospheric influence on the observation. The immediate goal of this study is therefore to estimate relevant atmosphere parameters. This is however not completely separable from the TES problem, and, in fact, for certain special cases, the TES problem is solved simultaneously. The secondary goal is to establish sensor requirements, i.e., how many spectral bands are needed for a certain noise level.

B. Related work

Despite a large body of research on atmospheric estimation and correction, and its impact on detection [1], there is little work focusing on minimal sensor requirements. Also, the methods in the literature require either high-resolution spectral data, or the use of spectral features in other wavelengths than the ones we study here [3], [4].

Well-known methods for atmospheric correction include AAC [5] and ISAC [6], [7]. Extensions include ARTEMIS [8], where a database of atmospheric transmissions is used and a smooth emissivity spectrum is retrieved. While many atmospheric correction techniques, including ours, rely on atmospheric modelling software to estimate atmospheric parameters, Smith et al. [9] apply a data-only scheme. Other extensions to attack special cases of the TES problem include [10], [11], [12]. Surface temperature retrieval methods for high-emissivity objects are treated in [13], including a review of the nine available methods in the literature. The type of parameterized models for the atmosphere that we use here is also used by Fox et al. [14] and Chandra et al [15]. Note that the literature on atmospheric correction in the visual and near infrared wavelengths is much more plentiful than for LWIR.

C. Contribution

The contribution of this paper is fourfold. First, we formulate an optimization procedure that can be used for finding various unknown atmosphere, object, and sensor parameters in a hyperspectral setting. Second, we formalize and derive the models for the different cases. Third, we extend the method to exploit a priori information, and, finally, we show how to extend the method for adjusting spectral calibration.

D. Outline

Section II explains how we model the observed objects, the atmosphere, and the sensor. In order to adapt the models to observed data, their parameters are estimated using the optimization procedures described in Section III. These optimization procedures are analyzed with respect to sensitivity in Section IV. Simulations are performed to examine the effect of varying the number of spectral bands and the amount of noise in Section V. Exploitation of a priori data in the form of estimates/measurements of some of those parameters is described in Section VI. A real

TABLE I
NOTATION

Symbol	Meaning	Unit
a, x, ϕ	Scalars	
$\mathbf{a}, \mathbf{x}, \boldsymbol{\phi}$	Vectors	
$\mathbf{A}, \mathbf{X}, \boldsymbol{\Phi}$	Matrices (except $\mathbf{L}$)	
$\mathbf{I}$	The identity matrix	
λ	Wavelength	μm
t	Temperature	K
$\varepsilon, \boldsymbol{\varepsilon}, \mathbf{E}$	Emissivity	
$\tau, \boldsymbol{\tau}, \mathbf{T}$	Atmospheric transmission	
$L, \mathbf{L}$	Radiance	μflick
$b, \mathbf{b}$	Blackbody radiance	μflick
$L^o, \mathbf{L}^o$	Object radiance	μflick
$\hat{L}^o, \hat{\mathbf{L}}^o$	Estimated object radiance	μflick
$L^s, \mathbf{L}^s$	At-sensor radiance	μflick
$L^\uparrow, \mathbf{L}^\uparrow$	Up-welling radiance	μflick
$L^\downarrow, \mathbf{L}^\downarrow$	Down-welling irradiance	μflick
Note: 1 flick = 1 $\frac{W}{\text{sr} \cdot \text{cm}^2 \cdot \mu\text{m}}$		

world experiment is described in Section VII and conclusions are drawn in Section VIII. Appendices contain the derived models used in Sections III and VI.

Table I summarizes the mathematical notation used in this paper.

II. MODELLING

In this section, we describe how we model the three important entities present in our scenario: the observed object, the atmosphere, and the sensor. The type of models used are non-controversial, even if other modeling methods, such as subspace models [15] and databases [8], exist.

A. Overview

We want to estimate and study the thermal radiance $L^o(\lambda)$ from an object at various wavelengths λ . In a laboratory setting, where the sensor is close to the object, the problem would be considerably less complicated. In the setting treated here, the sensor is airborne high above the observed objects, and we thus need to take atmosphere effects into account. We model the atmosphere by a function $a(\cdot)$, and the at-sensor radiance L^s is thus a function $L^s(\lambda) = a(L^o(\lambda))$.

Disregarding any spatial effects, the sensor is modeled as a spectral filter for each spectral band. The spectral filtering is the function that samples a continuous function $L(\lambda)$ and outputs a measurement vector $\mathbf{z}$. The sensor is modeled by a function $\mathbf{s}(\cdot)$, and our observed (measured) vector $\mathbf{z}$ is modeled as $\mathbf{z} \approx \mathbf{s}(L^s(\lambda))$. The wavelength λ will sometimes be omitted below in order to simplify the notation.

In practice, we perform all computations on wavelength-discrete vectors. The resolution (dimensionality) of these vectors are either determined by the sensor's spectral resolution (N bands/dimensions) or by the high-resolution simulations performed by MODTRAN[®] [16] (M bands/dimensions). These simulations give high-resolution transmission/radiance vectors (every cm^{-1}) that for our purposes serve well as (approximately) continuous functions, i.e., $M \gg N$. We will thus need discrete sensor and atmosphere functions, as detailed below.

B. Object radiance model

The radiation at the wavelength λ from an object is the sum of the emitted and the reflected radiation, i.e.,

$$L^o(\lambda) = \varepsilon(\lambda) \cdot b(t, \lambda) + (1 - \varepsilon(\lambda)) \cdot L^\downarrow(\lambda) \quad (1)$$

where $0 \leq \varepsilon(\lambda) \leq 1$ and $b(t, \lambda)$ is the radiation from a blackbody with temperature t according to Planck's spectral radiance function [18]

$$b(t, \lambda) = \frac{2hc^2 \cdot 10^4}{\lambda^5 (\exp(\frac{hc \cdot 10^6}{k\lambda t}) - 1)}. \quad (2)$$

Note that directional effects are neglected here, i.e., we the model assumptions include that the irradiance and/or the object surface is 100% diffuse.

The corresponding general discrete model is

$$\mathbf{L}_{t,\varepsilon}^o = \mathbf{E}\mathbf{b}_t + (\mathbf{I} - \mathbf{E})\mathbf{L}^\downarrow, \quad (3)$$

where $\mathbf{E} = \text{diag}(\varepsilon(\lambda_1) \dots \varepsilon(\lambda_M))$, $\mathbf{L}^\downarrow$ is an irradiance vector simulated in MODTRAN[®], and $\mathbf{b}_t$ is the blackbody radiance vector

$$\mathbf{b}_t = [b(t, \lambda_1) \dots b(t, \lambda_M)]^T. \quad (4)$$

We can see the object radiance $\mathbf{L}_{t,\varepsilon}^o$ as being parameterized by $M + 1$ parameters; the emissivities in each wavelength $\lambda_1, \dots, \lambda_M$ and the temperature t . In order to reduce the dimensionality of the search space in the optimization problem described in Section III, we seek parameterizations with lower dimensionality. Possible parameterizations include:

- Emissivity smoothness assumption. In the literature, it is sometimes assumed, e.g., in [19], that the object's emissivity spectrum is a smooth function. This is in various ways exploited to reduce the degrees of freedom for the emissivity, for example by minimizing the residual between an estimated emissivity $\varepsilon(\lambda)$ and a smoothed version $\varepsilon^s(\lambda) = (g * \varepsilon)(\lambda)$, where $g(\lambda)$ is a suitable smoothing filter kernel. Similarly, the emissivity could be parameterized using a few smooth basis functions. This assumption is useful when studying unknown scenes containing, e.g., vegetation, but will not be used in this paper.
- Graybody assumption. Assuming constant emissivity over the relevant wavelengths, the object radiance is

$$\mathbf{L}_{t,\varepsilon}^o = \varepsilon \mathbf{b}_t + (1 - \varepsilon) \mathbf{L}^\downarrow. \quad (5)$$

That is, the object radiance model is controlled by two parameters (ε and t) and the simulated irradiance.

- Near blackbody assumption. If the emissivity is close to one and/or that the down-welling radiance is small compared to the object radiance, then $(1 - \varepsilon)L^\downarrow \approx 0$ and we can use the model

$$\mathbf{L}_{t,\varepsilon}^o = \varepsilon \mathbf{b}_t \quad (6)$$

controlled by two parameters only.

- Blackbody assumption. Assuming the observed object is a blackbody, the object radiance vector will simply be

$$\mathbf{L}_t^o = \mathbf{b}_t. \quad (7)$$

This assumption is useful when known blackbody reference objects are present in the observed scene. The object radiance is then completely defined by its temperature.

We will in the following use the graybody model and thus let the object radiance be parameterized by one emissivity value ε and one temperature t . When it comes to the optimization procedures (Section III), the near-blackbody and blackbody cases as well are treated in the appendix for completeness.

C. Atmosphere model

The atmospheric transmission is modeled as a multiplicative transmission in each wavelength, and since the atmosphere is also a source of radiation in itself, the model of the observation is

$$a(L^o(\lambda)) = \tau(\lambda) \cdot L^o(\lambda) + L^\uparrow(\lambda), \quad (8)$$

where $L^\uparrow$ is the upwelling path radiance. Spatial distortion due to turbulence as well as scattering is disregarded.

Assuming that the atmospheric transmission $\tau_d(\lambda)$ *without* water vapor and the transmission of water vapor $\tau_v(\lambda)$ are known, we can parameterize the atmospheric transmission (for each wavelength) as

$$\tau_w = \tau_d \cdot \tau_v^w, \quad (9)$$

where the *water vapor content scaling* w is the relative amount of water vapor, i.e., $w = \frac{v}{v_0}$, where v is the actual amount of water vapor and v_0 is the reference amount of water vapor used in the MODTRAN[®] simulations (see Section II-D). Note that the equality (9) becomes an approximation in the discrete case. As long as w is close to one, the approximation is close to equality (see also Appendix A).

The up-welling atmosphere radiance depends on the atmospheric transmission and the temperature profile $\alpha(\nu)$, i.e., the air temperature at each altitude ν . Here, we simplify the temperature profile to a constant α (see discussion below) and express the radiance as

$$L^\uparrow(w, \alpha) = (1 - \tau_w) \cdot b(\alpha). \quad (10)$$

The *discrete atmosphere model* is thus

$$\mathbf{a}_{w,\alpha}(\mathbf{L}^o) = \mathbf{T}_w \mathbf{L}^o + \mathbf{L}^\uparrow \quad (11)$$

$$= \mathbf{T}_w \mathbf{L}^o + (\mathbf{I} - \mathbf{T}_w) \mathbf{b}_\alpha, \quad (12)$$

where

$$\mathbf{T}_w = \text{diag}(\boldsymbol{\tau}_d) \text{diag}(\boldsymbol{\tau}_v^w) \quad (13)$$

and $\boldsymbol{\tau}_d$ and $\boldsymbol{\tau}_v^w$ are the simulated M -dimensional atmosphere transmissions vectors.

Obviously, the atmosphere temperature can vary quite a lot with the altitude, and we can handle that in, at least, three ways:

- 1) Approximate $\alpha(\nu)$ with a constant, α , i.e the average temperature.
- 2) Model the air as multiple layers, each with its own temperature α_i (and possibly humidity w_i).
- 3) Average, at each wavelength, the blackbody curve over a (possibly weighted) temperature interval around an average parameter α .

We chose to follow the first approach, because it was considered the most pedagogical. The principle of the optimization described in Section III remains the same in all three cases. The second approach adds one or more degrees of freedom to the optimization, which makes the procedure more sensitive to noise, and thus not necessarily improving the results. The third approach replaces the blackbody function in (12) with a somewhat more complicated function but leaving everything else unchanged. Depending on the altitude of the sensor, this choice should be revisited.

D. Atmosphere simulation

As mentioned above, the atmosphere is simulated in MODTRAN[®]. For these simulations, weather forecasts and standard atmosphere models are used. The resulting transmission $a(\lambda)$ and radiance $L^\downarrow$ are then parameterized as previously explained (Section II-C).

Thus, in practice, we typically have a fair idea of the atmosphere's behavior, which we use as a starting point for refining the atmosphere parameters.

The atmospheric transmission used in the experiments in this paper is a simulation of the transmission from ground level to one kilometer above at midday of a Scandinavian summer's day. From this, the up-welling radiance can be computed according to (10). The down-welling radiance $L^\downarrow$ is simulated for the same time and location. All three are illustrated in Fig. 1. The simulation is performed for each cm^{-1} .

Fig. 1. Simulated atmospheric transmission, up-welling radiance, and down-welling radiance.

E. Sensor model

Recalling that we do not make any spatial considerations here, the sensor can be modeled purely as a spectral function $\mathbf{s} : \mathcal{L}^2 \rightarrow \mathbb{R}^N$. The sensor inputs an observed radiation $L^s(\lambda)$ (a continuous function of λ) and outputs a discrete vector $\mathbf{z}$ of N measurements. Each measurement z_n is a sample of $L^s(\lambda)$, which is sampled using a response function $r_n(\lambda)$ centered around λ_n . The response function $r_n(\lambda)$ is typically modeled as a Gaussian function with a certain width (full width half magnitude, FWHM) δ_n , i.e.,

$$r_n(\lambda) = f_{\text{Gauss}}(\lambda_n, \frac{\delta_n^2}{8 \ln 2}), \quad (14)$$

where $f_{\text{Gauss}}(\mu, \sigma^2)$ is the Gaussian probability density function with mean μ and variance σ^2 .

The sampling is thus expressed by a function $p : \mathcal{L}^2 \rightarrow \mathbb{R}^N$ such that

$$p(L^s(\lambda)) = [p_1(L^s(\lambda)) \dots p_N(L^s(\lambda))]^T, \quad (15)$$

and the sampling function for each spectral band is given by

$$p_n(L^s(\lambda)) = \int r_n(\lambda) L^s(\lambda) d\lambda. \quad (16)$$

Unfortunately, the sensor is not perfectly uniform or even linear. A common model is that there is a gain $g_{i,j}$ and an offset $m_{i,j}$ value for each sensor element (i, j) [17], and thus the output from the sensor element is

$$k_{i,j} \cdot p_n(L^s(\lambda)) + m_{i,j}. \quad (17)$$

We will in the following assume that the sensor is calibrated¹ so that the offsets equal zero and the gains are equal, i.e., $k_{i,j} = g \forall i, j$.

Thus, we model the measurements as

$$z_n = \mathbf{s}_n(L^s(\lambda)) = g \cdot p_n(L^s(\lambda)), \quad (18)$$

where p_n is given by (16).

The *discrete sensor function* $\mathbf{s} : \mathbb{R}^M \rightarrow \mathbb{R}^N$ samples high-resolution vectors using the Gaussian spectral filter, and our discrete sensor model, is

$$\mathbf{s}(\mathbf{L}^s) = g \mathbf{P}_{\delta_n, \lambda_n} \mathbf{L}^s, \quad (19)$$

where $\mathbf{P}$ is an $N \times M$ matrix determined by $\{\delta_n, \lambda_n\}$. We will in the following assume that $\{\delta_n, \lambda_n\}$ are known and omit them from our set of parameters. However, fear not, we will meet them again in Section VII.

F. Summarizing the models

Summarizing the models above, the model vector $\mathbf{y}$ for the observation is

$$\mathbf{y}(\mathbf{x}) = g \mathbf{P} (\mathbf{T}_w \mathbf{L}_{t,\varepsilon}^o + (\mathbf{I} - \mathbf{T}_w) \mathbf{b}_\alpha) \quad (20)$$

¹Determining the center wavelengths λ_n is referred to as spectral calibration and determining $k_{i,j}$ and $m_{i,j}$ (and/or g) as radiometric calibration (and non uniformity correction).

where $\mathbf{T}_w$ is given by (13) and $\mathbf{L}_{t,\varepsilon}^o$ is given by (3). τ_d , τ_v , and $\mathbf{L}^\downarrow$ are computed using MODTRAN[®]. We regard the model vector as a function of the parameter vector $\mathbf{x} = [g, w, \alpha, t, \varepsilon]$. The sensor is characterized by the parameters g , $\{\delta_n\}$, and $\{\lambda_n\}$ (the latter two specifying the sampling matrix $\mathbf{P}$). The atmosphere is parameterized by w and α . The object radiance is characterized by $\varepsilon(\mathbf{u})$, $t(\mathbf{u})$, and $L^\downarrow(\lambda)$ for a certain pixel coordinate $\mathbf{u}$ (in contrast to the sensor and atmosphere parameters, that can be assumed to be constant over a larger area, as discussed in Section III-B).

In practice, the measured vector $\mathbf{z}$ will not equal $\mathbf{y}$ exactly. We have model errors since few of the assumptions above are fully true, and we have noise added by the sensor. We will model the noise as a being zero-mean, additive, white, and Gaussian. Thus, our measurement will be $\mathbf{z} = \mathbf{y} + \mathbf{r}$, where $\mathbf{r} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathbf{r}})$.

Finding the optimal parameters is a non-linear optimization problem in the three parameters w , t , and α , and with g and ε as linear parameters in a sub-problem. These issues will be treated in the next section.

While the models here are generative, i.e., they explain how to synthesize a model vector $\mathbf{y}$ of the observation, they can also be used for atmospheric correction as described in Appendix A.

III. OPTIMIZATION PROCEDURE

In order to find the model parameters described in the previous section, we minimize the residual between the model and the observation. Minimizing the residual between an observation and a parameterized object-atmosphere-sensor model is an optimization problem in several variables. The problem is non-linear in some of the parameters and linear in some.

A. One pixel to rule them all

We start by studying a single measurement, that is, one pixel only. Thus, with our model vector $\mathbf{y}$ that depends on a parameter vector $\mathbf{x} = [g, \alpha, w, t, \varepsilon]$ and the observation vector $\mathbf{z}$, we estimate the parameters by minimizing

$$f(\mathbf{z}, \mathbf{x}) = \|\mathbf{r}(\mathbf{z}, \mathbf{x})\|^2 = \|\mathbf{z} - \mathbf{y}(\mathbf{x})\|^2 \quad (21)$$

with respect to $\mathbf{x}$. The objective function f is non-linear and a suitable optimization algorithm should be applied to find the optimal $\mathbf{x}$.

The problem is simplified by observing that ε and g are linear parameters that can be solved for in a linear sub-problem given values on t , α , and w . Thus, we split the parameter vector in a linear part $\mathbf{x}_l = [g, \varepsilon]$ and non-linear part $\mathbf{x}_{nl} = [\alpha, w, t]$. The non-linear parameters are optimized iteratively, and for each iteration the linear parameters are solved for. For the non-linear problem, we use Sequential Quadratic Programming (SQP) [20], chosen because it does not require analytical gradients or Hessians. SQP iteratively approximates the optimization problem as a quadratic problem, thus solving a sequence of quadratic programming problems. Implementations of SQP are available in several packages, we use the MATLAB[®] Optimization Toolbox [21].

For the linear sub-problem, we solve in a least-squared sense, and our notation for a constrained linear least-squares problem is to find

$$\arg \min_{\mathbf{x}} \|\mathbf{C}\mathbf{x}_l - \mathbf{d}\|^2 \quad (22)$$

constrained by $\mathbf{A}\mathbf{x}_l \leq \mathbf{b}$.

The specific linear sub-problem to be solved depends on our assumptions about the object radiance (blackbody, near blackbody, graybody) and whether we know or need to estimate the sensor gain. The resulting linear sub-problems (i.e., the derivations of $\mathbf{C}$, $\mathbf{d}$, $\mathbf{A}$, and $\mathbf{b}$) for several cases are detailed in Appendix B.

Given some basic knowledge about the world, we can easily set some constraints on the non-linear parameters $[t, \alpha, w]$ as well.

B. For a handful of pixels

When we have multiple (K) observations (pixels) from the same scene, the atmosphere parameters do not change much between the pixels, and also the sensor gain should be approximately constant during the data acquisition. Assuming that these parameters are constant over the scene, we can formulate a joint optimization problem where we estimate emissivities ($\varepsilon_1, \dots, \varepsilon_K$) and object temperatures ($t_1, \dots, t_K$) for each pixel but sensor gain, air temperature, and water vapor content scaling only once per scene, i.e.,

$$\arg \min_{\mathbf{X}} \|\mathbf{Z} - \mathbf{Y}(\mathbf{X})\|^2. \quad (23)$$

where

$$\mathbf{X} = [t_1 \dots t_K, \varepsilon_1, \dots, \varepsilon_K, g, w, \alpha] \quad (24)$$

$$\mathbf{Z} = [\mathbf{z}_1^T \dots \mathbf{z}_K^T]^T \quad (25)$$

$$\mathbf{Y} = [\mathbf{y}_1^T \dots \mathbf{y}_K^T]^T, \quad (26)$$

and $\mathbf{y}_k = \mathbf{y}(t_k, \varepsilon_k, g, w, \alpha)$. The multiple pixel version of (20) is thus

$$\mathbf{Y}(\mathbf{X}) = g\mathbf{P}_K \begin{bmatrix} \mathbf{T}_w \mathbf{L}_{t_1, \varepsilon_1}^o + (\mathbf{I} - \mathbf{T}_w) \mathbf{b}_\alpha \\ \vdots \\ \mathbf{T}_w \mathbf{L}_{t_K, \varepsilon_K}^o + (\mathbf{I} - \mathbf{T}_w) \mathbf{b}_\alpha \end{bmatrix}, \quad (27)$$

where $\mathbf{P}_K$ is the $NK \times NK$ matrix

$$\mathbf{P}_K = \begin{bmatrix} \mathbf{P} & & \\ & \ddots & \\ & & \mathbf{P} \end{bmatrix} \quad (28)$$

The non-linear problem

$$\arg \min_{\mathbf{X}_{nl}} \|\mathbf{Z} - \mathbf{Y}(\mathbf{X}_{nl})\|^2 \quad (29)$$

is essentially the same as above, however with $K+2$ parameters, which makes it a computationally heavy problem.

C. For a few pixels more

When the scene is large, it is not feasible to perform the joint estimation using all pixels simultaneously. This is due to two reasons. First, the computational burden will be huge, since the dimensionality of the problem grows linearly with the number of pixels, and the complexity grows exponentially with the dimensionality. Second, the

scene might contain many observations of materials that are not graybodies, meaning that the model error will be large.

Although there are several approaches on detecting high-emissivity (near blackbody) objects [22], the topic is not extensively treated here. Instead, a straight-forward approach is proposed as follows. The input is an initial estimate of the atmosphere's transmission and temperature.

- 1) Select a random set of observations (K pixels).
- 2) Fit a graybody curve to each of these observations, i.e., run the optimization described in Section III-A using the initial atmosphere estimates as starting point.
- 3) Select the observations with small enough fit errors. In practice, this thresholding is relatively simple.
- 4) Among the remaining observations, select J observations with as different (estimated) temperatures as possible (motivated in Section IV-A).
- 5) Use the J selected observations to globally estimate the atmosphere parameters using the procedure from Section III-B.

Since K is typically much smaller than the total number of pixels, this procedure can be run several times to get an estimate of the variance of the parameter estimates.

In the LWIR band, many types of vegetation are close enough to being graybodies to give small fit errors, making this method suitable for rural scenes or scenes including open water. A relatively small set of J such pixels can be used to find the atmosphere and sensor model parameters, whereafter the spectra of the remaining (non-graybody) pixels can be retrieved.

Evaluation of robust estimation approaches should be done in the future.

IV. SENSITIVITY ANALYSIS

The different parameters are not equally prone to estimation errors, and we investigate their sensitivity by generating a model vector $\mathbf{y}(\mathbf{x}_0)$ and then perturbing the parameters one by one, thus generating $\mathbf{y}(\mathbf{x}_{i,j})$ where x_i is perturbed j steps. For each parameter, we compute an error curve e telling the mean square error over the residual vector, i.e.,

$$e_j = \frac{1}{N} \|\mathbf{y}(\mathbf{x}_0) - \mathbf{y}(\mathbf{x}_{i,j})\|^2. \quad (30)$$

We do this for a few different object temperatures, since the sensitivity is highly variable with the object-air temperature difference. The results are plotted in Fig. 2. Since water content and temperature cannot be directly compared, the judgment is somewhat subjective. In practice, the object temperature estimation is relatively robust, while the water content scaling is sensitive to noise, especially when the temperatures of the object and the air are close. For equal object-air temperature, the water content estimate is random even without noise.

A. Sensitivity of atmosphere parameters

As seen in the previous subsection, the humidity estimate is very error-prone. Instead of plotting the error curve when perturbing the water content parameter, the error surface when perturbing both atmosphere parameters is

Fig. 2. Residuals when perturbing the model parameters from the true values. The true parameter values are $\alpha = 280$ K, $\varepsilon = 0.9$, $g = 1$, $w = 1$. The true object temperature is varied: Solid line: 280 K, dash-dotted: 290 K, dashed: 300 K, dotted: 310 K.

illustrated in Fig. 3. Note the different directions of the "valleys". In the middle graph, the object and air have equal temperatures, and the error will depend on the air temperature only. In this case we will still get a relatively robust estimate of the atmospheric temperature. However, for a larger object-air temperature difference, we will be able to estimate the water content, but the estimate is strongly correlated with the air temperature estimate. That is, an error in the air temperature estimation can easily be compensated by an error in the water content scaling estimate. Alternatively, an error in the water content estimate of 10% will lead to a bias of 2 K in the temperature estimate (this behavior changes slightly for a graybody).

The conclusions from this are:

- The estimates of all parameters except air temperature become more robust with a larger object-atmosphere temperature difference.
- If several observations are available, observations with as different object temperatures as possible should be selected. This is the reason for the selection process in Section III-C.
- If one of the atmosphere parameters can be measured, and this measurement can be exploited in the estimation process, the estimate of the other will be much more reliable. We return to this in Section VI.

B. Sensitivity of object and sensor parameters

The sensor and object parameters are all clearly correlated, as is illustrated by the error surfaces in Fig. 4. It is clear that we here have the same problem as mentioned earlier, that an error in one parameter can be compensated with an error in another one. Intuitively, this is expected, since all three parameters ε , t , and g have the major effect of increasing signal energy at increasing parameter levels.

Fig. 3. The isocurves of the error at different atmosphere and object temperatures. At the x axis is the estimated water content scaling, at the y axis is the estimated atmospheric temperature.

Fig. 4. The isocurves of the error when perturbing (pairwise) emissivity, object temperature, and sensor gain.

V. SIMULATIONS

A factor that influences the quality is the number of spectral bands used. In this section, we simulate a large number of observations in order to see how the performance degrades with increasing noise and fewer spectral bands.

A. Test methodology

The parameter estimations have been evaluated by simulations in the following way.

- Eight different noise levels have been selected, with the variances corresponding to no noise, 0, 20, 40, 50, 60, 70, and 80 dB [μflicks^2]. Expressed differently, variances between 0 and 80 dB translate to standard deviations between 1 and 10000 μflicks . To put this in a context, increasing the temperature of a blackbody from 300 to 301 K increases the emitted radiance with approximately 18 μflick at 8 μm and 12 μflick at 12 μm .
- For each noise level, 200 random spectral signatures have been created in high spectral resolution (one sample per cm^{-1}). The signatures were created by picking the parameters from uniform distributions in the following intervals: $t \in [280 \ 300]$, $\alpha \in [270 \ 290]$, $\varepsilon \in [0.9 \ 1.0]$, $w \in [0.9 \ 1.1]$, and $g \in [0.9 \ 1.1]$.
- Each of the signatures have been distorted by noise (additive white Gaussian noise).

- Nine different sensor models have been created, with 3, 5, 7, 9, 12, 15, 20, 30, and 40 spectral bands respectively. The response functions are Gaussians with $\text{FWHM} = 2 \cdot \Delta\lambda$, where $\Delta\lambda$ is the distance between each band's center wavelength. The band centers are evenly spread out between $8 + \Delta\lambda$ and $11.5 - \Delta\lambda \mu\text{m}$. Note that the sensors with fewer bands will be less noisy.
- Each of the sensors sampled each of the $200 \cdot 8$ signatures, and the average absolute errors of the parameters were measured.

This was repeated for different scenarios, as described below.

B. Simulations using synthetic signatures

We study four different cases, where g and w are known or unknown. From the analysis in the previous section, an unknown water content is expected to give a larger error in the air temperature estimate, and an unknown sensor gain is expected to give larger errors in the estimated of the object parameters (t and ϵ).

1) *Case 1: Known gain and water content:* In this simulation, w and g were set to a known value (1.0). The results are illustrated in Fig. 6(a), where the quality (or, rather, the error) of the estimated parameters is plotted against the number of spectral bands for different levels of noise. Two things are clear: (1) When the number of bands is too small, the estimation process is not provided with enough information, and (2) when the amount of noise is too large, the estimation apparently breaks down, and adding more bands does not help. For moderate noise levels (around 20 dB and less), the estimation, as expected, performs better for a larger number of bands. For these noise levels, 9 and 12 bands is enough to get an average error of less than 1K in the air and object temperature estimate respectively, and more than 15 bands does not give any significant improvement.

2) *Case 2: Known gain, unknown water content:* In this simulation, the true w was randomly picked but g fixed (and known). The results, see Fig. 6(b), are not much different from case 1, with the exception of the water content estimate that is quite bad even for a large number of bands with small amounts of noise (between 0.13 and 0.16 for all noise levels and numbers of bands). The estimate of the air temperature is worse than for the case with known water content – around five bands more are required for the same results.

3) *Case 3: Unknown gain, known water content:* In this simulation, the true g was randomly picked and w fixed (and known). As expected (see Section IV-B), the quality of the estimates of the object temperature t is reduced, especially when the number of bands is less than 10. The plots are found in Fig. 6(c).

4) *Case 4: Unknown gain and water content:* In this simulation, the true g and w were both randomly picked. As expected, the results are approximately the same as in case 3 (except for the water content parameter). The plots are found in Fig. 6(d).

C. Simulations using real signatures

In order to evaluate a more realistic scenario, the simulations above have also been made using real signatures of vegetation. These signatures are laboratory-measured hemispherical reflection of grass, green leaves of birch and aspen, and dry/brown leaves of birch. The leaves are measured on both front and back sides. Instead of randomly

choosing an emissivity constant, one of these seven signatures have been selected in the creation of each signature. The results are plotted in Fig. 5, and we make the following observations:

- The water content estimate is, as expected, error prone (average error of 0.14–0.16 regardless of noise/number of bands).
- The gain estimate is acceptable (average error less than 0.05) for noise levels of 20 dB and less and 10 bands or more.
- The air temperature estimate is good (average error 2 K or less) as long as the number of bands is more than 5 and the noise is less than 40 dB.
- At a noise level around 50 dB the estimation of air temperature breaks down, and also gets worse with increasing number of bands.

Fig. 5. Estimation errors depending on the noise level and the number of bands. The observed objects are randomly selected vegetation components with temperature observed through an atmosphere with random temperature. Top (Case 1): Water content and gain are fixed to known values. Bottom (Case 4): Water content and gain random (and unknown).

D. Summarizing the simulation results

If the observed signal is disturbed by noise, as in all practical cases, the atmospheric estimation is degraded. Another factor that influences the quality is the number of bands used. We simulate a large number of observations in order to see how the performance degrades with increasing noise and fewer spectral bands. From the analysis in the Section IV, an unknown water content is expected to give a larger error in the air temperature estimate, and an unknown sensor gain is expected to give larger errors in the estimates of the object parameters (t and ε). The simulations confirm the analysis from the previous section, and also indicate how the estimation quality varies with the number of bands.

- 10 bands or more seem to be needed for accurate estimates.
- Using more than 20 bands does not improve the estimation quality significantly.
- The water content is hard to estimate, even when using a large number of bands with small noise.
- If the sensor gain is unknown, the estimation quality of the object parameters is reduced.

VI. EXPLOITING A PRIORI KNOWLEDGE

From the sensitivity analysis, it is clear that stabilization of one or more of the parameters (especially the humidity parameter) would improve the performance. In a real situation, we often have access to in situ measurements of the air temperature and water content. Even if exact measurements cannot be done (measuring, e.g., the total water vapor content is hard), expected values can be derived from measurements. In the scheme above, this information is used only as a starting point for the optimization procedure. We show here how to exploit the knowledge of a measured value and its uncertainty by reformulating the optimization problem.

Instead of minimizing the residual, we find the parameters that minimize the deviation from the expected values (derived from measurements), weighted by the uncertainty (in form of a standard deviation). Moreover, assuming that our model of the observation holds, the residual $\mathbf{r} = \mathbf{z} - \mathbf{y}(\mathbf{x})$ is mainly due to sensor noise, and we input the noise level in the same way. The new expression to minimize, replacing (21), is thus

$$\frac{\sum r_n^2}{\sigma_r^2} + \left(\frac{w - m_w}{\sigma_w} \right)^2 + \left(\frac{\alpha - m_\alpha}{\sigma_\alpha} \right)^2 + \left(\frac{g - m_g}{\sigma_g} \right)^2, \quad (31)$$

where m_w, m_α , and m_g are our expectations of water content, air temperature, and sensor gain. σ_w, σ_α , and σ_g are the corresponding standard deviations, and σ_r^2 is the sensor noise level. We do not take into account any measurements of object emissivity and temperature, since these will not be available in our scenario

2.

As before, we solve for the non-linear parameters (α, w, t) in an iterative search procedure where we evaluate (31) in each step. In each iteration we solve for $\mathbf{x}_l$, which, in contrast to the problem in Section III is not a linear, but a quadratic problem on the form

$$\min_{\mathbf{x}} \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{f}^T \mathbf{x} \text{ such that } \mathbf{A} \mathbf{x} \leq \mathbf{b} \quad (32)$$

²In more general terms, we assume that the vector $\boldsymbol{\xi}^T = [\mathbf{r}^T \ w \ \alpha \ g]$ is a stochastic vector with an elliptic distribution centered around $\mathbf{m}_{\boldsymbol{\xi}}^T = [0 \ \dots \ 0 \ m_w \ m_\alpha \ m_g]$ and with covariance $\boldsymbol{\Sigma} = \text{diag}[\sigma_r^2 \ \dots \ \sigma_r^2 \ \sigma_w^2 \ \sigma_\alpha^2 \ \sigma_g^2]$ (with the Gaussian distribution, i.e., $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{m}_{\boldsymbol{\xi}}, \boldsymbol{\Sigma}_{\boldsymbol{\xi}})$ as a familiar example) and find the parameters that maximize the likelihood by minimizing $(\boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}})^T \boldsymbol{\Sigma}_{\boldsymbol{\xi}}^{-1} (\boldsymbol{\xi} - \mathbf{m}_{\boldsymbol{\xi}})$.

(a) Case 1: The water content scaling factor w as well as the gain g are assumed to be exactly known. (b) Case 2: The water content scaling w is unknown. The gain g is assumed to be exactly known. (c) Case 3: The water content scaling w is assumed to be exactly known. The gain g is unknown. (d) Case 4: The water content scaling w as well as the gain g is unknown.

Fig. 6. Estimation errors depending on the noise level and the number of bands. The observed objects are graybodies with random emissivities and temperature observed through an atmosphere with random temperature.

that we solve using an active set method [23]. As previously, the non-linear optimization is performed using SQP (see Section III), sampling the objective function to estimate the gradient and Hessian. Possibly, the search would be faster and more accurate using analytical expressions for these (switching to, e.g., a trust-region method). However, these functions are, frankly speaking, nasty to derive, and might also be so complex that the dense sampling is faster anyway, hence we keep the SQP method. Details and derivations of $\mathbf{H}$ and $\mathbf{f}$ for the different cases are given in Appendix C.

VII. EXPERIMENT

A. Data collection

Data was collected by FOI, the Swedish Defence Research Agency, at a rural trial site in Sweden. The setting was as follows: An aircraft flying at an altitude of one kilometer and equipped with an LWIR hyperspectral camera with 256 spectral bands acquired images over a rural area. Various objects of interest were placed in the area. Since most of these objects were of military nature, the imagery is classified. However, of special interest in the scope of this paper are the pools of water also present, with temperatures measured by thermometers. Also, the air temperature and humidity at the ground was measured.

B. Spectral calibration

The initial attempts to fit the models described in this paper to the measured data failed drastically. Visual inspection revealed that the spectral calibration of the sensor was erroneous, with the shorter wavelength spectral bands distributed more densely than the longer wavelength bands. Instead of each band center being separated by 17 nm, the first (shorter wavelengths) band centers were separated with around 13 nm and the last (longer wavelengths) with around 20 nm. The offsets from the true values were visually estimated (by comparing to curves such as the simulations in Fig. 1) to be up to 50 nm, enough to destroy the optimization process.

The solution was to parameterize the band centers. Instead of assuming that the spectral bands were centered around $\lambda_n = p_1 \cdot n + p_0$ for known values of $\{p_1, p_0\}$, we used the model

$$\lambda_n = p_2 \cdot n^2 + p_1 \cdot n + p_0, \quad (33)$$

and then included the parameters $\{p_i\}$ in the optimization. That is, the parameter vector $\mathbf{x}$ was extended to

$$\mathbf{x} = [g, w, \alpha, t, \varepsilon, p_2, p_1, p_0]. \quad (34)$$

Thus, the discrete sensor sampling matrix $\mathbf{P}$ in (19) has to be recomputed in each iteration of the optimization. However, no other changes needs to be done to the method, which also demonstrates its versatility. The drawback is of course that adding three degrees of freedom to the optimization makes it more sensitive to noise.

C. Results

For the real world experiment, the results are more of a qualitative nature since not all true values for the parameters are known (that is why the simulations above were done). For evaluation, spectra from the mentioned

pools of water were used. The temperature estimate was within the 2 K indicated by the simulations, however with a positive bias of >1 K. This might, speculative, be an effect of the spectral calibration. The atmosphere temperature estimate also converged to reasonable values; the true average air temperature between the ground and the sensor is not known. The spectral calibration was evaluated by visual inspection by comparing simulated and measured spectra; for the lower bands ($<10 \mu\text{m}$) the calibration seems to very exact. For the more noisy signals with less spikes above $10 \mu\text{m}$, the results are less certain.

VIII. CONCLUSION

We have presented a method for atmospheric estimation in hyperspectral LWIR data. The method also involves the estimation of object parameters (temperature and emissivity) under the restriction that the emissivity is constant. A procedure for finding a subset of observation where this restriction (approximately) holds is described. The method is analyzed with respect to its sensitivity to noise and number of spectral bands. We have shown how a priori information from physical measurements can be exploited in the estimation process.

In conclusion, the method works well under favorable conditions. Provided that the noise level is low enough, air and object temperatures can be estimated with a precision of around 2 K, using the information from less than 10 spectral bands. If the water vapor content of the atmosphere is unknown and/or the radiometric calibration of the sensor is inaccurate, the precision is of less quality.

Experiments on real data measured using an airborne hyperspectral camera confirm the predictions from the simulations. However, the nature of the experiments did not allow exact quantitative results.

ACKNOWLEDGMENTS

The author thankfully acknowledges the advice from Ingmar G. Renhorn as well as the support with MODTRAN[®] from Rolf Persson, both former employees of FOI, the Swedish Defence Research Agency. The data collection mentioned in Section VII was performed by FOI and financed by the Swedish Armed Forces Research and Technology Programme. This work was funded by the Swedish Research Council through the project EMC².

APPENDIX A

ATMOSPHERIC CORRECTION

When the atmosphere parameters are found, we can use them (and the resulting transmission and radiance) for atmospheric correction. Atmospheric correction (or compensation) is the process of computing (an estimate of) the object radiance from the measurement vector.

For the sake of simplicity, we assume that the sensor is spectrally and radiometrically calibrated so that λ_n are known, $k_{i,j} = 1$, and $m_{i,j} = 0$. We measure a spectral vector $\mathbf{z}$, and, if the parameters g , w , α are known, we can estimate the object radiance vector in sensor resolution

$$\mathbf{L}^o = p(L^o) = g^{-1}\mathbf{s}(L^o), \quad (35)$$

where $p(\cdot)$ is the sensor's sampling function given by (16) and our sensor model is $\mathbf{s}(L) = gp(L)$.

Since $\mathbf{y} = \mathbf{s}(a(L^o))$ we have $L^o = a^{-1}(\mathbf{s}^{-1}(\mathbf{y}))$. This is not very useful since $\mathbf{s}$ does not have an inverse, i.e., we cannot reconstruct the continuous function $L(\lambda)$ from the samples $\mathbf{y} = \mathbf{s}(L(\lambda))$. However, using the approximations

$$p(L_1 \cdot L_2) \approx p(L_1) \cdot p(L_2) \quad (36)$$

$$p(L_1 + L_2) \approx p(L_1) + p(L_2) \quad (37)$$

we have

$$p(L^o(\lambda)) = p(a^{-1}(L^s(\lambda))) \quad (38)$$

$$= p(\tau(\lambda)^{-1} \cdot (L^s(\lambda) - L^\uparrow(\lambda))) \quad (39)$$

$$\approx p(\tau(\lambda)^{-1}) \cdot (p(L^s(\lambda)) - p(L^\uparrow(\lambda))). \quad (40)$$

These approximations might not look very intuitive, but are nevertheless used implicitly when discretizing this type of functions.

Inserting

$$p(\tau(\lambda)^{-1}) \approx \mathbf{P}\mathbf{T}^{-1} \quad (41)$$

$$gp(L^s) = \mathbf{y} \quad (42)$$

$$p(L^\uparrow(\lambda)) \approx \mathbf{P}\mathbf{L}^\uparrow \quad (43)$$

in (40) we get

$$\mathbf{L}^o \approx (\mathbf{P}\mathbf{T}^{-1})(g^{-1}\mathbf{y} - \mathbf{P}\mathbf{L}^\uparrow) \quad (44)$$

$$\approx (\mathbf{P}\mathbf{T}^{-1})(g^{-1}\mathbf{z} - \mathbf{P}\mathbf{L}^\uparrow) + \mathbf{r}'. \quad (45)$$

where $\mathbf{r}' = \mathbf{P}\mathbf{T}^{-1}g^{-1}\mathbf{r}$. Assuming $\mathbf{r}$ to have mean zero, our best estimate of the object radiance (i.e., the atmospherically corrected radiance) is

$$\hat{\mathbf{L}}^o(\mathbf{z}, g, w, \alpha) = (\mathbf{P}\mathbf{T}_w^{-1})(g^{-1}\mathbf{z} - \mathbf{P}\mathbf{L}_{w,\alpha}^\uparrow). \quad (46)$$

with variance

$$\Sigma_{\mathbf{r}}' = g^{-2}\mathbf{P}(\mathbf{T}^{-2}\Sigma_{\mathbf{r}}\mathbf{P}^T). \quad (47)$$

Note that for wavelengths where $\tau(\lambda)$ is close (or equal) to zero the estimate is unreliable (or undefined).

APPENDIX B

THE LINEAR SUBPROBLEMS

In the single pixel case we regard each pixel separately (or we have only one observation). For the linear subproblem, we have a constrained linear least-squares problem that we write

$$\arg \min_{\mathbf{x}} \|\mathbf{C}\mathbf{x}_l - \mathbf{d}\|^2 \quad (48)$$

constrained by $\mathbf{A}\mathbf{x}_l \leq \mathbf{b}$.

The specific problem to be solved depends on our assumptions about the object radiance (blackbody, near blackbody, graybody) and whether we know or need to estimate the sensor gain. These problems are defined by the matrices/vectors $\mathbf{C}$, $\mathbf{d}$, $\mathbf{A}$, and $\mathbf{b}$ as below.

A. Graybody assumption, unknown gain

The model is given by

$$\mathbf{y} = g \mathbf{P} (\mathbf{T}_w (\varepsilon \mathbf{b}_t + (1 - \varepsilon) \mathbf{L}^\downarrow) + \mathbf{L}_{w,\alpha}^\uparrow), \quad (49)$$

and given observation $\mathbf{z}$, find

$$\arg \min_{\mathbf{x}} \|\mathbf{z} - \mathbf{y}(\mathbf{x})\|^2. \quad (50)$$

The problem is linear in ε and g , so that given values on $\mathbf{x}_{nl} = [t, \alpha, w]$, $\{\mathbf{T}_w, \mathbf{L}_{w,\alpha}^\uparrow\}$ are known and the model can be written as

$$\mathbf{y} = \mathbf{P}(\mathbf{T}_w \mathbf{L}^\downarrow + \mathbf{L}_{w,\alpha}^\uparrow)g + \mathbf{P}\mathbf{T}_w(\mathbf{b}_t - \mathbf{L}^\downarrow)g\varepsilon \quad (51)$$

$$= \mathbf{C}_0 g + \mathbf{C}_1 g \varepsilon. \quad (52)$$

ε and g are constrained by $0 \leq g_{lb} < g \leq g_{ub}$ and $0 \leq \varepsilon_{lb} < \varepsilon \leq \varepsilon_{ub} \leq 1$, and thus given by the solution to the least-squares problem formulated above with

$$\mathbf{C} = \mathbf{P} [\mathbf{T}_w \mathbf{L}^\downarrow + \mathbf{L}_{w,\alpha}^\uparrow \mid \mathbf{T}_w(\mathbf{b}_t - \mathbf{L}^\downarrow)] \quad (53)$$

$$\mathbf{x} = [g \mid g\varepsilon]^T \quad (54)$$

$$\mathbf{d} = \mathbf{z} \quad (55)$$

$$\mathbf{A} = \begin{bmatrix} \varepsilon_{lb} & -1 \\ -\varepsilon_{ub} & 1 \\ -1 & 0 \\ 1 & 0 \end{bmatrix} \quad (56)$$

$$\mathbf{b} = \begin{bmatrix} 0 \\ 0 \\ -g_{lb} \\ g_{ub} \end{bmatrix}. \quad (57)$$

B. Graybody assumption, known gain

If the sensor is radiometrically calibrated, g needs not to be estimated. Without loss of generality we can let $g = 1$, and thus

$$\mathbf{y} = \mathbf{C}_0 + \mathbf{C}_1 \varepsilon. \quad (58)$$

and the linear sub-problem is

$$\mathbf{C} = \mathbf{P} \mathbf{T}_w (\mathbf{b}_t - \mathbf{L}^\downarrow) \quad (59)$$

$$\mathbf{x} = \varepsilon \quad (60)$$

$$\mathbf{d} = \mathbf{z} - \mathbf{P}(\mathbf{T}_w \mathbf{L}^\downarrow - \mathbf{L}_{w,\alpha}^\uparrow) \quad (61)$$

$$\mathbf{A} = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad (62)$$

$$\mathbf{b} = \begin{bmatrix} -\varepsilon_{lb} \\ \varepsilon_{ub} \end{bmatrix} \quad (63)$$

$$(64)$$

C. Near-blackbody assumption

The near-blackbody assumption is simply a special case of the graybody assumption, where we set $\mathbf{L}^\downarrow = \mathbf{0}$ and a suitable lower bound ε_{lb} on the emissivity, for example 0.98. The linear sub-problem becomes

$$\mathbf{C} = \mathbf{P} [\mathbf{L}_{w,\alpha}^\uparrow \mid \mathbf{T} \mathbf{b}_t] \quad (65)$$

$$\mathbf{x} = [g \mid g\varepsilon]^T \quad (66)$$

$$\mathbf{d} = \mathbf{z} \quad (67)$$

$$\mathbf{A} = \begin{bmatrix} -1 & 0 \\ \varepsilon_{lb} & -1 \\ -1 & 1 \end{bmatrix} \quad (68)$$

$$\mathbf{b} = \mathbf{0}. \quad (69)$$

D. Blackbody assumption

Assuming that the object is a blackbody, the linear problem is of course even more simplified, boiling down to

$$\mathbf{C} = \mathbf{P}(\mathbf{L}_{w,\alpha}^\uparrow + \mathbf{T}_w \mathbf{b}_t) \quad (70)$$

$$\mathbf{x} = g \quad (71)$$

$$\mathbf{d} = \mathbf{z} \quad (72)$$

$$\mathbf{A} = -1 \quad (73)$$

$$\mathbf{b} = 0. \quad (74)$$

APPENDIX C

THE QUADRATIC SUBPROBLEMS

As described in Section VI, we should minimize

$$\frac{\sum r_i^2}{\sigma_r^2} + \left(\frac{w - m_w}{\sigma_w} \right)^2 + \left(\frac{\alpha - m_\alpha}{\sigma_\alpha} \right)^2 + \left(\frac{g - m_g}{\sigma_g} \right)^2 \quad (75)$$

in order to take estimates of humidity, atmosphere temperature and sensor gain into account. Inserting (20), i.e.,

$$\mathbf{y}(\mathbf{x}) = g \mathbf{P} (\mathbf{T}_w \mathbf{L}_{t,\varepsilon}^o + (\mathbf{I} - \mathbf{T}_w) \mathbf{b}_\alpha) \quad (76)$$

into (75) we arrive, after some manipulation, at

$$\begin{aligned} \|\mathbf{r}(\mathbf{z}, \mathbf{x})\|^2 = & \frac{1}{2\sigma_g^2} \mathbf{x}_l^T \left(\mathbf{D}^T \mathbf{D} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) \mathbf{x}_l \\ & + \left(\frac{1}{\sigma_n^2} \mathbf{D}^T \mathbf{z} + \frac{m_g}{\sigma_g^2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)^T \mathbf{x}_l \\ & + \frac{1}{\sigma_w^2} (w - m_w)^2 + \frac{1}{\sigma_\alpha^2} (\alpha - m_\alpha)^2 \end{aligned} \quad (77)$$

as the new objective function. $\mathbf{x}_l = [g \ \varepsilon]^T$ is the vector of linear parameters and

$$\mathbf{D} = [\mathbf{T}_w \mathbf{L}^\downarrow + \mathbf{L}_{w,\alpha}^\uparrow \mid \mathbf{T}_w (\mathbf{b}_t + \mathbf{L}^\downarrow)]. \quad (78)$$

$\mathbf{D}$ is thus dependent on the non-linear parameters w , t , and α . We note that the first two terms in (77) can be identified as a quadratic programming (QP) problem and that the last two terms are constant with respect to g and ε . Thus, we solve for the non-linear parameters (α, w, t) in an iterative search procedure where we evaluate (77) in each step. In each iteration we solve for $\mathbf{x}_l$, which is a quadratic problem on the form

$$\min_{\mathbf{x}} \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{f}^T \mathbf{x} \text{ such that } \mathbf{A} \mathbf{x} \leq \mathbf{b}. \quad (79)$$

The matrix $\mathbf{H}$ and the vector $\mathbf{f}$ are identified with the corresponding parts of (77) so that

$$\mathbf{H} = \frac{1}{\sigma_g^2} \left(\mathbf{D}^T \mathbf{D} + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right) \quad (80)$$

$$\mathbf{f} = \frac{1}{\sigma_n^2} \mathbf{D}^T \mathbf{z} + \frac{m_g}{\sigma_g^2} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (81)$$

and the QP problem is solved using an active set method [23]. Similarly as for the linear problems, there are different cases to consider, as described below.

A. Graybody assumption, unknown gain

The objective function is

$$\frac{1}{2} \mathbf{x}_l^T \mathbf{H} \mathbf{x}_l + \mathbf{f}^T \mathbf{x}_l + c(\mathbf{x}_{nl}) \quad (82)$$

where $\mathbf{H}$ and $\mathbf{f}$ are given by (80), (81) and (78), and the function c is

$$c = \frac{1}{\sigma_w^2} (w - m_w)^2 + \frac{1}{\sigma_\alpha^2} (\alpha - m_\alpha)^2, \quad (83)$$

as given by (77).

B. Near-blackbody assumption

The near-blackbody assumption is simply a special case of the graybody assumption, where we set $\mathbf{L}^\dagger = \mathbf{0}$ and a suitable lower bound ε_{lb} on the emissivity, for example 0.98. The matrix $\mathbf{D}$ becomes

$$\mathbf{D} = [\mathbf{L}_{w,\alpha}^\dagger \mid \mathbf{T}_w \mathbf{b}_t]. \quad (84)$$

$\mathbf{H}$ and $\mathbf{f}$ are computed by inserting the (84) in (80) and (81), and c stays the same as above (83).

C. Blackbody assumption

Assuming the object is a blackbody, the emissivity ε disappears (so that $\mathbf{x}_l = k$), and the equations boil down to:

$$\mathbf{H} = \frac{1}{\sigma_g^2} (\mathbf{L}^{\dagger T} \mathbf{L}^\dagger + 1) \quad (85)$$

$$\mathbf{f} = \frac{1}{\sigma_g^2} ((\mathbf{T}_w \mathbf{b}_t)^T \mathbf{L}^\dagger + m_g) + \frac{1}{\sigma_n^2} \mathbf{z}^T \mathbf{L}^\dagger \quad (86)$$

$$c = \frac{1}{\sigma_n^2} \mathbf{z}^T \mathbf{T}_w \mathbf{b}_t + (\mathbf{T}_w \mathbf{b}_t)^T (\mathbf{T}_w \mathbf{b}_t) + \frac{1}{\sigma_w^2} (w - m_w)^2 + \frac{1}{\sigma_\alpha^2} (\alpha - m_\alpha)^2 \quad (87)$$

REFERENCES

- [1] P. W. T. Yuen and G. Bishop, "Enhancements of target detection using atmospheric correction preprocessing techniques in hyperspectral remote sensing," in *Proceedings of SPIE (Military Remote Sensing)*, vol. 5613, pp. 111–118, 2004.
- [2] A. Gillspie, S. Rokugawa, T. Matsunaga, J. S. Cothorn, S. Hook, A. B. Kahle, "A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 4, 1998.
- [3] M. D. Abel, J. M. Zennera, G. A. Petrick, A. T. Buswell, M. L. Pilati, W. R. Czyzewski, L. P. Alessandro, and S. K. Weaver, "A new approach to atmospheric correction of hyperspectral data," in *Proceedings of SPIE (Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery VIII)*, vol. 4725, (Orlando, FL), pp. 72–82, 2002.
- [4] K. L. Hirsch, L. Balick, C. Borel, and P. McLachlan, "A comparison of four methods for determining precipitable water vapor content from multi-spectral data," in *Proceedings of SPIE (Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery VII)*, vol. 4381, (Orlando, FL), pp. 417–428, 2001.
- [5] D. Gu, A. Kahle, and F. D. Palluconi, "Autonomous atmospheric compensation (AAC) of high-resolution hyperspectral thermal infrared remote-sensing imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 6, pp. 2557–2570, 1999.
- [6] B. R. Johnson and S. J. Young, *In-scene Atmospheric Compensation: Application to SEBASS Data at the ARM Site.*, Aerospace Report No. ATR-99(8407)-1 Parts I and II, 1998.
- [7] E. Hernandez-Bacquero, *Characterization of the Earth's surface and atmosphere from multispectral and hyperspectral thermal imagery*, PhD dissertation, Rochester Institute of Technology, 2000.
- [8] C. C. Borel, "ARTEMISS – an algorithm to retrieve temperature and emissivity from hyper-spectral thermal image data," in *28th Annual GOMACTech Conference, Hyperspectral Imaging Session*, (Tampa, FL), 2003.
- [9] M. R. Smith, A. R. Gillespie, H. Mizzon, L. K. Balick, J. C. Jimi, $\frac{1}{2}$ nez-Mui, $\frac{1}{2}$ oz, and J. A. Sobrino, "In-scene atmospheric correction of hyperspectral thermal infrared images with nadir, horizontal, and oblique view angles," *Int. Journal of Remote Sensing*, vol. 34, no. 9–10, 2013.
- [10] Y.-G. Qian, N. Wang, C. Gao, Y.-Y. Jia, L. Ma, H. Wu, Z.-L. Li, and L. Tang, "Preliminary evaluation of linear spectral emissivity constraint temperature and emissivity separation method for contrast samples from hyperspectral thermal infrared data," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 465–468, 2013.

- [11] J. Cui, B. Yan, X. Dong, S. Zhang, J. Zhang, F. Tian, and R. Wang, "Temperature and emissivity separation and mineral mapping based on airborne TASI hyperspectral thermal infrared data," in *International Journal of Applied Earth Observation and Geoinformation*, vol. 40, pp. 19–28, 2015.
- [12] M.-A. Gagnon, P. Tremblay, S. Savary, V. Farley, P. Lagueux, and M. Chamberland, "Airborne thermal hyperspectral imaging of urban and rural areas," in *Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 1369–1372, 2014.
- [13] X. Zhong, J. Labed, G. Zhou, K. Shao, and Z.-L. Li, "A Multi-Channel Method for Retrieving Surface Temperature for High-Emissivity Surfaces from Hyperspectral Thermal Infrared Images," *Sensors*, vol. 15, no. 6, pp. 13406–13423, 2015.
- [14] M. Fox, J. Gruninger, J. Lee, A. J. Ratkowski, and M. L. Hoke, "Atmospheric parameterization for model-based thermal infrared atmospheric correction of spectral imagery," in *Proceedings of SPIE (Imaging Spectrometry VIII)*, vol. 4816, pp. 93–103, 2002.
- [15] K. Chandra and G. Healey, "Using coupled subspace models for reflectance/illumination separation," in *Proceedings of SPIE (Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X)*, vol. 5425, (Orlando, FL), pp. 538–548, 2005.
- [16] MODTRAN®, <http://www.modtran.org>.
- [17] I. Renhorn, "Modelling of imaging spectral sensors," Scientific report FOI-R--2118--SE, Swedish Defence Research Agency (FOI), 2006.
- [18] M. Alonso and E. J. Finn, *Physics*. Addison-Wesley, 1992.
- [19] L. Guanter, R. Richter, and J. Moreno, "Spectral calibration of hyperspectral imagery using atmospheric absorption features," *Applied Optics*, vol. 45, no. 10, pp. 2360–2370, 2006.
- [20] J. Nocedal and S. J. Wright, *Numerical Optimization*, Springer, 2006.
- [21] MATLAB®, <http://www.mathworks.com>.
- [22] R. D. Kaiser, D. L. Vittoe, and A. K. Andrews, "Detecting low-emissivity objects in LWIR hyperspectral data and the corresponding impact on atmospheric compensation," *Proceedings of SPIE (Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery IX)*, vol. 5093, paper no. 705, 2003. doi: 10.1117/12.497122
- [23] P. E. Gill, W. Murray, and M. Wright, *Practical Optimization*. London, UK: Academic Press, 1981.

award-winning.

Jörgen Ahlberg received his M.Sc. degree in Computer Science and Engineering in 1996 and his Ph.D. in Electrical Engineering in 2002, both from Linköping University, Sweden. He then held positions as scientist and research leader at FOI, the Swedish Defence Research Agency for nine years. He is currently an Adjunct Senior Lecturer at Linköping University and runs R&D projects at Visage Technologies, Termisk Systemteknik, and Glana Sensors, companies specializing in visual, thermal and hyperspectral computer vision respectively. Research interests, in the general area of image analysis and vision, includes tracking and analysis of facial images as well as automatic detection, recognition, and tracking in thermal and hyperspectral systems. He has published more than 40 scientific papers, of which four