The simulation of APEX data: the SENSOR approach

Anko Börner\textsuperscript{a}, Michael Schaepman\textsuperscript{a}, Daniel Schläpfer\textsuperscript{a}, Lorenz Wiest\textsuperscript{b}, Ralf Reulke\textsuperscript{b}

\textsuperscript{a}Remote Sensing Laboratories (RSL), University of Zurich, Switzerland
\textsuperscript{b}German Aerospace Center (DLR), Institute of Space Sensor Technology, Berlin, Germany

ABSTRACT

The consistent simulation of airborne and spaceborne hyperspectral data is an important task and sometimes the only way for the adaptation and optimization of a sensor and its observing conditions, the choice and test of algorithms for data processing, error estimations and the evaluation of the capabilities of the whole sensor system.

The integration of three approaches is suggested for the data simulation of APEX (Airborne Prism Experiment): a) a spectrally consistent approach (e.g. using AVIRIS data), b) a geometrically consistent approach (e.g. using CASI data), and c) an end-to-end simulation of the sensor system. In this paper, the last approach is discussed in detail. Such a technique should be used if there is no simple deterministic relation between input and output parameters.

The simulation environment SENSOR (Software Environment for the Simulation of Optical Remote Sensing Systems) presented here includes a full model of the sensor system, the observed object and the atmosphere. The simulator consists of three parts. The first part describes the geometrical relations between object, sun, and sensor using a ray tracing algorithm. The second part of the simulation environment considers the radiometry. It calculates the at-sensor-radiance using a pre-calculated multidimensional lookup-table for the atmospheric boundary conditions and bi-directional reflectances. Part three consists of an optical and an electronic sensor model for the generation of digital images. Application-specific algorithms for data processing must be considered additionally.

The benefit of using an end-to-end-simulation approach is demonstrated, an example of a simulated APEX data cube is given, and preliminary steps of evaluation of SENSOR are carried out.

Keywords: APEX, hyperspectral, spectroscopy, simulation, optimization

1. INTRODUCTION

The European Space Agency (ESA) has identified the Land Surface Processes and Interactions Mission (LSPIM) (Berger, 1999) as an Earth Explorer Core Mission to increase the understanding of terrestrial ecosystems. Many of ESA's missions are complemented by precursor missions in order to foster the development of methods and algorithms and provide the user community with scientific data prior to the start-off of the mission itself.

In the case of the LSPIM, which is currently in its Phase A (Labandibar, 1999), the Airborne PRISM Experiment (APEX) has been defined as a partial data simulation, validation and calibration source for LSPIM. APEX will be able to simulate the entire reflective part of the spectrum covered by LSPIM as well as adding a to be defined thermal sensor for the emissive part (Schaepman, 1998a).

The directional component of the LSPIM can be simulated using an additional source, namely CHRIS (Lobb, 1999). The directional part of APEX will be achieved by flying pre-calculated flight lines depending on solar angles.
The APEX instrument is defined as an airborne pushbroom imager (Fig. 1) with 1000 pixels across track and a swath width of 2.5 - 5.0 km depending on the flight altitude. 300 spectral bands cover the wavelength range from 400 - 2500 nm. The spectral sampling distance is specified as 5 nm in the VIS/NIR and 10 nm in the SWIR. The APEX specifications were identified based on the Phase A study (Itten, 1997).

Fig. 1: Principle of a pushbroom imaging spectrometer using a prism

The exact system design, the optimization, and evaluation of sensor parameters, observation conditions, and retrieval algorithms have to be performed using different approaches. In general, there are three methods applicable:

i) Physical model

If there is a simple relationship between input and output data, this model should be used. For example, the instantaneous field of view corresponds proportionally to the signal-to-noise ratio and no complex software simulation is required.

ii) Testbed

If the input-output relationship is not simple, it can be useful to use hardware testbeds in order to determine an optimal system configuration, e.g. to evaluate the effect of using a mechanical shutter.

iii) Software simulation

If both methods i) and ii) are not feasible or too expensive, then complex computer models should be used. The main idea is to describe the essential parts of the remote sensing system with physical, mathematical or algorithmic models.

Since a remote sensing system like APEX is a complex device, and since a complete evaluation of the sensor hardware is not possible without the knowledge of its environment (including the observed object) and data processing, the only way to understand the interactions between these parts is the implementation of an complex computer model which describes all the essential processes. The main objectives of this modeling approach are:

- Evaluation of specifications,
- Test of hardware processing facilities,
- Choice and test of retrieval algorithms,
- Accuracy and error estimation of data products, and
- Cost-versus-quality considerations.

For APEX, a software simulator named SENSOR (Software EN vironment for the Simulation of Optical Remote sensing systems) has been developed by a cooperation between the Remote Sensing Laboratories (RSL) and the German Aerospace Cen-
In order to simulate the instrument and to evaluate the impact of changes of the specifications during the construction phase of APEX, the simulator for APEX is capable of monitoring the change of the specifications in terms of uncertainty statements in dedicated data products. The basic concept behind SENSOR is to put all relevant knowledge of hardware of the remote sensing system (e.g. optical distortions, stray light, dark signal, pixel response non-uniformity), the source of radiance, the atmosphere, and the observed object (e.g. slope, aspect) into one model.

In combination with the data processing algorithms, SENSOR represents a powerful tool treating all mentioned parts as elements of a complex, unified system. It is the only way to describe non-trivial relations between any input and output parameter and to draw conclusions concerning single parts of the entire system.

SENSOR pursues a similar philosophy as other end-to-end simulation approaches (e.g. SENSAT, Richter 1990b), but due to its complexity, SENSOR considers effects which are ignored often. It is based on an inheritance of a predecessor proposed by Reulke (1995).

2. SENSOR: AN END-TO-END-SIMULATION TOOL

Fig. 2 depicts a general flow diagram of SENSOR. From left to right all essential input parameters, the processing modules, and the outputs (also input parameters for the next processing step), are given.

**Input**
- Digital elevation model
- Maps (e.g. classes)
- Sun position
- Flight path
- Geometric calibration
- Sky view factor
- Measured or modeled reflectance spectra
- Atmospheric parameters
- Atmospheric LUT's

**SENSOR Modules**
- **Ray tracing**
  - Determination of geometric relations between object, sun, and remote sensing system
- **Radiative transfer**
  - Determination of at-sensor-radiances based on a simulated atmosphere
- **System characterization**
  - Calculation of digital numbers considering effects of the optics and modeling the signal processing chain

**Output**
- Intersection points
- Normal vectors
- Classes of observed surface elements
- Pixel viewing directions
- At-sensor-radiances
- Spectra
- Image cubes

**Fig. 2:** Scheme of the software environment for the simulation of optical remote sensing systems (SENSOR)

SENSOR consists of three main parts: the first one determines the geometric relation between the remote sensing system, the radiant source, and the object, the second one simulates the influence of the atmosphere and calculates the at-sensor-radiances, and the third describes the system hardware consisting of optical and electronic components.
2.1 Ray tracing

The task of the first SENSOR module is to determine the geometric relations between the observed object (in general characterized by a digital elevation model), the radiant source and the remote sensing system applying a ray tracing algorithm.

A flight over a digital elevation model (DEM) is simulated. Information about the position of the sensor (x, y, z), attitude (roll, pitch, yaw), and the geometric calibration data can be provided by data of other sensor systems or can be simulated. Using this information, so-called pixel rays are defined. They describe the viewing direction vector of each detector element at all times during the simulated flight. The determination of the object points observed and their properties - class of the object (e.g. “wood”), sky view factor, normal vector, surface temperature - is the next step. This additional information is given by maps or derived from the DEM directly. Starting from a certain detector element, the pixel ray is tracked until a surface element is hit. The following section gives an idea of this so-called ‘ray tracing’ in a DEM (Fig. 3 illustrates the situation).

Fig. 3: Ray tracing in a digital elevation model

Usually the DEM consists of a set of equidistant distributed samples defining a number of small rectangular patches. In SENSOR, for reasons of simplicity and computing efficiency, these elements are subdivided into two triangles per patch. The size of the DEM elements and the ground sampling distance of the remote sensing system should be approximately the same. The task of ray tracing is to find the triangle, which is seen of a certain detector element. That means to find the intersection point between the pixel ray and the appropriate triangle. In most cases, it is computationally far too expensive to check all triangles whether they were hit by a pixel ray or not, therefore fast algorithms are required. The shadowing effects are considered in the SENSOR ray tracing module as well, but are not discussed in detail in this paper.

After reproducing the pixel ray (by assigning each pixel a position and a viewing direction), the intersection points between the pixel ray and two bounding planes (defined by the maximum and minimum height of the DEM) are calculated. Both intersection points mark a sub-array of the DEM. The intersection point between the pixel ray and the DEM must lie within this sub-array. The next step is to find the triangles which lie on the trace of the pixel ray (Fig. 4). All these candidates have to be tested with respect to an intersection with the current pixel ray. The following equation system solves the problem:

\[ \mathbf{r_{isp}} = \mathbf{r_{0r}} + t_r \cdot \mathbf{r_{dr}} = \mathbf{r_{0s}} + t_{p1} \cdot \mathbf{r_{dp1}} + t_{p2} \cdot \mathbf{r_{dp2}}, \]

where \( \mathbf{r_{isp}} \) is the vector of the intersection point between the pixel ray and the triangle plane, \( \mathbf{r_{0r}} \) the origin of the pixel ray, \( t_r \) the scaling parameter of the pixel ray, \( \mathbf{r_{dr}} \) the pixel viewing direction vector, \( \mathbf{r_{0s}} \) the origin of the triangle plane, \( \mathbf{r_{dp1}} \) and \( \mathbf{r_{dp2}} \) the vectors stretching the triangle plane, \( t_{p1} \) and \( t_{p2} \) the scaling parameters of the triangle plane. It can be checked easily, whether the found intersection point lies within the triangle or not.
The speed of the procedure depends on the geometry of the remote sensing system (e.g. field of view), the flight motions, and the DEM parameters (height differences, resolution, roughness). This procedure is done once per spatial pixel, independent of the number of spectral channels. As an example: Ray tracing algorithms compute about 20 minutes on a Sun Ultra 60 for a simulated APEX scene of 1000 spatial pixels and 1500 lines in a mountainous region (mountain Rigi, Switzerland, height differences of about 1300 m within an area of 10 km $\times$ 10 km).

The output of the ray tracing procedure is:

- Geometry of the object points observed: intersection points with DEM, normal vectors of the observed DEM elements, and
- Properties of the object points observed: classes, sky view factors, temperatures.

The object properties allow to define a link to a spectral library. The library can contain measured or simulated reflectance spectra. An assignment between an object class and a reference spectrum is established using a lookup-table.

### 2.2 Radiative transfer

After the determination of the geometric properties, the at-sensor-radiance for each pixel of each image line in each spectral channel is calculated. The radiative transfer describes the influence of the earth’s atmosphere on the solar irradiance. SENSOR uses data simulated by MODTRAN (Berk, 1989), because of the large wavelength range and the flexibility in varying pixel viewing directions.

Internally, SENSOR applies a sub-channel model. These sub-channels are the smallest units. All calculations are done on this level. All spectral channels of the remote sensing system are built by a number of sub-channels. Width and sampling distance of the sub-channels are set to 1 nm, since MODTRAN’s internal minimum step width (1 cm$^{-1}$) corresponds to a step width of 1 nm at the uppermost wavelength being relevant for APEX (2500 nm or 4000 cm$^{-1}$).

First, the corresponding reflectance value of the observed object is needed for each sub-channel. This information is available after ray tracing using the link between object classes and spectral libraries. The reference spectrum must be given at the same or a better resolution as the sampling distance of the sub-channels. It should be mentioned that spectral features can be modeled correctly only if the sampling theorem is considered. Furthermore, SENSOR is able to include bidirectional reflectance distribution functions (BRDF) provided by spectral libraries.
The second way to retrieve reflectance values is to access a data cube containing reflectance values for each spatial pixel. These data cubes can be obtained by resampling of atmospherically corrected data sets (Schläpfer, 1999) of other, spectrally similar hyperspectral remote sensing systems, such as AVIRIS (Porter, 1987).

The next step is the determination of the at-sensor-radiance of each sub-channel. MODTRAN allows this calculation for a definable channel and a set of environmental parameters. Pre-calculated MODTRAN lookup-tables (LUT) are used in order to increase processing speed. The LUT’s are given for a set of values of all important input parameters. The number of entries per input parameter and their resolution has to be defined by the known APEX specifications and by expected observation conditions.

Therefore, a tool generating these LUT’s is used (Wiest, 1998). Currently, the type of atmosphere, aerosol type, visibility, sensor altitude, ground altitude, sun zenith angle, and wavelength are considered as the varying input parameters.

MODTRAN calculates one integral value (total radiance) for the at-sensor-radiance, belonging to a set of input parameters. Certain input parameters are unknown prior to the simulation, e.g. reflectance, slope, and aspect. All of these parameters have to be one parameter of the MODTRAN LUT, otherwise the pre-calculated total radiance cannot be used as a final value. Instead, the contributing parts of the integral value are taken directly or indirectly from the MODTRAN output files. The essential contributing parts of the at-sensor-radiance are depicted in Fig. 5.

![Contributing parts of total at-sensor-radiance](image)

**Fig. 5: Contributing parts of total at-sensor-radiance (after Wiest, 1998)**

Those parts are:

- $L_{\lambda,g}$ ground emmisivity,
- $L_{\lambda,p,th}$ thermal path radiance,
- $L_{\lambda,p,sc}$ total path scattered solar radiance (including adjacency effects),
- $L_{\lambda,dir}$ direct ground reflected solar radiance,
- $L_{\lambda,diff,sc}$ diffuse ground reflected solar radiance, and
- $L_{\lambda,diff,th}$ diffuse ground reflected thermal radiance.

By introducing additional information, e.g. reflectance, slope, aspect, these parts can be revalued and summed up the total at-sensor-radiance. In comparison to storing the value of the total radiance only, this parametric approach is more flexible. The size of the pre-calculated LUT’s and the computing time for generating them is significantly smaller.

Currently, the MODTRAN LUT used for the APEX simulation consists of seven entries (six radiative contributing parts explained above and the optical depth) for all combinations of the input parameters (Table 1). It has a size of about 7.7 MByte. Its generation takes about half an hour on a Sun Ultra 60. The LUT can be expanded easily, e.g. by adding different aerosol models.
There are two major limitations have to be overcome for the simulation. The first is the interpolation procedure to model appropriate results of intermediate input parameters, which is not trivial in a seven-dimensional data cube. In this work, the nonlinear seven-dimensional hypersphere is expanded into a linear Taylor-series and a multidimensional linear approach is used to solve the equation systems. The second one is the navigation in a multidimensional data cube. The incessantly calculation of the address of an appropriate data set is a very time consuming process. So procedures for a fast data access are necessary.

The result of the module described above is the at-sensor-radiance for each spatial pixel of each image line in each spectral channel. The applied parametric approach allows a flexible use of pre-calculated MODTRAN LUT’s.

### 2.3 System characterization

The module describes the hardware of the remote sensing system considering the aspects of signal and system theory (Jahn, 1995). It is divided into an optical and an electronic part. The aim is the calculation of digital numbers out of the at-sensor-radiance given either by the radiative transfer module of SENSOR or by radiance values provided by other hyperspectral remote sensing systems.

Each spectral APEX channel is characterized by a spectral response function. It describes the sensitivity of the channel regarding to the energy in a certain range of the electromagnetic spectrum. In order to model a continuous response function, SENSOR defines the response function of each channel as a set of discrete samples (sampling distance 1 nm). The first step is the calculation of the number of generated photons arriving the detector for each sub-channel by the equation (Schaepman, 1998a):

\[
n_p = A_e \cdot \omega \cdot t_i \cdot \int_{\lambda_1}^{\lambda_2} \tau_0(\lambda) \cdot \frac{\lambda}{h \cdot c} \cdot L(\lambda) \, d\lambda ,
\]

where \( n_p \) is the number of photons, \( A_e \) the area of the entrance aperture, \( \omega \) the solid angle of the instantaneous field of view, \( t_i \) the integration (exposure) time, \( \tau_0(\lambda) \) the response function of the entire remote sensing system, \( \lambda \) the wavelength, and \( L(\lambda) \) the at-sensor-radiance. The response function of the entire remote sensing system must be given for each pixel on the detector CCD (spatial and spectral direction) and is the result of a calibration process. All other parameters are fixed per simulated flight. First of all, the Poisson distributed photon noise must be considered (yielding to \( n_p' \)).

The number of electrons liberated by the photons is given by

\[
n_e = n_p' \cdot \eta(\lambda) ,
\]
where \( n_e \) is the number of electrons, and \( \eta(\lambda) \) the wavelength dependent quantum efficiency. Most approaches end on this stage, ignoring important noise sources and processing steps. In contrast to these models, SENSOR includes the parameters characterizing the hardware of the remote sensing system. A quasi-convolution in the spectral direction is performed by summing up the generated electrons of the sub-channels into one spectral APEX channel (Fig. 6).

Fig. 6: Sub-channel concept for signal generation within SENSOR

Behind the optics, an electronic data processing chain is modeled. It includes certain noise sources (e.g. pixel response non-uniformity and dark signal), and an analog-digital-converter. The output is one digital number per spatial and spectral pixel, per image line. The transfer of the at-sensor-radiances to digital numbers corresponds to an inverse radiometric calibration. This procedure is performed for each spatial pixel of each line of each spectral channel.

The influence of the point spread function (PSF) of the system and blurring caused by the flight motion during integration time is considered by applying convolution kernels across and along flight track. This is performed for each spectral image, considering neighboring pixels. The result is a simulated APEX data cube.

The cube also includes data of the planned in-flight calibration. Before and after each data acquisition, the system looks into a calibration sphere and a few lines will be scanned with the shutter closed. Black pixels on the outer border of the CCD are used, in order to estimate the dark current of the electronics (Schaepman, 1998b). All these data are introduced in the simulation process as well.

Smile and frown (keystone), caused by the optical components of APEX, are of special interest. If needed, the simulation of these effects with SENSOR may be performed the following way: For each pixel on the CCD (spatial and spectral), separate geometric calibration data are provided (simulated or measured). Then for each pixel the ray tracing procedure has to be car-
ried out separately. After that, the radiation transfer and system hardware modules can be applied the way described above. The simulation of distortion effects using SENSOR is possible, but results in a memory and computational efficiency problem.

Currently, it takes about 20 hours on a Sun Ultra 60 to simulate one entire APEX image cube with a size of 1000 pixels × 1500 image lines × 300 spectral channels. This time would yet increase by applying enhanced convolution algorithms or by using narrower sub-channels.

3. SIMULATED DATA

3.1 Example

Fig. 7 shows an APEX data cube simulated by SENSOR. The image gives an unnatural impression due to a coarse subdivision of the object classes. The first and last lines are the in-flight calibration data (Schaepman, 1998b). The black pixels on the left and right border are used for dark current measurements.

Fig. 7: SENSOR simulated data cube

3.2 Verification

It is an essential prerequisite that the consistency of SENSOR simulated data is verified. The review of the geometric simulation shows no problems and is performed by comparing the original elevation model with all calculated intersection points between pixel rays and DEM.

The following procedure is carried out, in order to estimate the radiometric accuracy of SENSOR simulated data:

- Simulation of a flight over an area (all input parameters are known exactly, e.g. the reflectance spectra of a selected target) - the result is a cube of digital numbers,
- Application of an independent atmosphere correction program (ATCOR, Richter, 1990a) to the simulated data.

The result should be the input reflectance spectra. The comparison of one original, wavelength independent input spectrum (reflectance = 0.3) and a reconstructed output spectrum can be seen in Fig. 8 (only the first 190 APEX channels were considered).
As expected, there are problems in the water vapor absorption bands. Ignoring these parts of the spectrum, the absolute mean deviation between the input and the output spectrum amounts to 0.001 and the standard deviation to 0.084 in this example. In our opinion, these deviations are caused by the parametric approach generating the at-sensor-radiance and differences of the LUT’s generated for SENSOR and ATCOR.

4. CONCLUSIONS

An complex end-to-end simulation tool called SENSOR is introduced. It describes the remote sensing system itself and essential parts of its environment. Using this concept, the interactions between parameters of the complex model, retrieval algorithms, and any output, such as data accuracy and costs, can be evaluated. In principle, it is possible to model all optoelectronic remote sensing systems since merely the part concerning the hardware has to be changed. SENSOR differs from other approaches due to its high complexity. All parts of the complex sensor-environment system are considered. Remarkable features are implemented, e.g. the influence of PSF, different noise sources (photons noise, pixel response non-uniformity, dark signal).

One of the main arguments against a software model is the complete reliance on a selection of known parametrized effects. However, this approach is often the only way to obtain any information on the interaction between the elements mentioned above. Consequently, one important task during the construction of a complex software model such as SENSOR is to find the essential parameters influencing the output values and to include them into the model.

5. OUTLOOK

In the near future, the emphasis will be put on the design and adaptation of APEX hardware components. Additional, the following features will be discussed in order to improve SENSOR’s accuracy and applicability:
• improving the parametric approach for the calculation of MODTRAN LUT's and the at-sensor-radiance, e.g. introducing relative azimuth and sensor zenith,
• modeling of adjacency effects,
• evaluation of more complex convolution algorithms to improve the sub-channel concept,
• including complex optical effects, e.g. straylight models, and
• including complex electronic effects, e.g. temperature dependent noise sources and cross talks.

In general, an optimization or evaluation process using SENSOR works as shown in Fig. 9. A set of system parameters provided by the scientific APEX team and the industrial partners is fed into SENSOR. The simulation results in an artificial image data cube. This data cube will be processed by a number of retrieval algorithms. The output data generated this way can be compared to user driven requirements, e.g. data quality and uncertainties, feasibility considerations, and cost analysis. This results in an evaluation of the input parameters and to an update of the sensor design if necessary.

Fig. 9: General scheme for an optimization or evaluation process using SENSOR

6. REFERENCES


Jahn, H., Reulke, R., 1995: Systemtheoretische Grundlagen optoelektronischer Sensoren; Akademie Verlag, Berlin, pp. 298


