
Copyright 2012 Society of Photo-Optical Instrumentation Engineers. One print or electronic copy may be made for personal use only. Systematic reproduction and distribution, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

http://dx.doi.org/DOI# TBD

See article below.
High-speed atmospheric correction for spectral image processing

Timothy Perkins\textsuperscript{a}, Steven Adler-Golden\textsuperscript{*}, Patrice Cappelaere\textsuperscript{b}, and Daniel Mandl\textsuperscript{b}

\textsuperscript{a}Spectral Sciences, Inc., 4 Fourth Avenue, Burlington MA 01803-3304
\textsuperscript{b}Vightel Corp., 8303 Winter Wind Ct., Ellicott City, MD 21043-6059
\textsuperscript{c}NASA Goddard Space Flight Center, 8800 Greenbelt Rd., Greenbelt MD 20771

ABSTRACT

Land and ocean data product generation from visible-through-shortwave-infrared multispectral and hyperspectral imagery requires atmospheric correction or compensation, that is, the removal of atmospheric absorption and scattering effects that contaminate the measured spectra. We have recently developed a prototype software system for automated, low-latency, high-accuracy atmospheric correction based on a C++-language version of the Spectral Sciences, Inc. FLAASH\textsuperscript{TM} code. In this system, pre-calculated look-up tables replace on-the-fly MODTRAN\textsuperscript{®} radiative transfer calculations, while the portable C++ code enables parallel processing on multicore/multiprocessor computer systems. The initial software has been installed on the Sensor Web at NASA Goddard Space Flight Center, where it is currently atmospherically correcting new data from the EO-1 Hyperion and ALI sensors. Computation time is around 10 s per data cube per processor. Further development will be conducted to implement the new atmospheric correction software on board the upcoming HyspIRI mission’s Intelligent Payload Module, where it would generate data products in near-real time for Direct Broadcast to the ground. The rapid turn-around of data products made possible by this software would benefit a broad range of applications in areas of emergency response, environmental monitoring and national defense.

Keywords: atmosphere, correction, compensation, FLAASH, hyperspectral, multispectral, algorithms, remote sensing

1. INTRODUCTION

Remotely sensed spectral imagery of the Earth's surface can be used to fullest advantage when the influence of the atmosphere has been removed and the measurements are reduced to units of reflectance. Elimination of molecular and particulate scattering and absorption from the data is desired for many applications, such as when comparisons are to be made with data taken in the laboratory or under different atmospheric or viewing conditions. This process, which transforms the data from spectral radiance to spectral reflectance, is known as atmospheric correction, compensation, or removal. First-principles atmospheric correction of visible-near-infrared-shortwave infrared (VNIR-SWIR) hyperspectral imagery (HSI) typically consists of two steps. The first is the retrieval of atmospheric parameters, including an aerosol description (most importantly, the visibility or optical depth) and the column water vapor amount. The second step is the solution of the radiative transfer (RT) equation for the retrieved aerosol and water vapor and transformation from radiance to reflectance.

A number of atmospheric correction algorithms have been developed over the years for both hyperspectral and multispectral imagery [1-8]. Among these is the FLAASH\textsuperscript{TM} (Fast Line-of-sight Atmospheric Analysis of Spectral Hypercubes) code [8-15], developed by Spectral Sciences, Inc. working in close collaboration with, and with support from, the US Air Force, NASA, and other US Government agencies. A commercial version of the code is sold as an add-on to the well-known ENVI (Environment for Visualizing Images) software package.

This paper reports development of new prototype FLAASH-based software suitable for automated, low-latency atmospheric correction on both ground-based and on-board processing systems. The software is based on a C++-language version of FLAASH, denote FLAASH-C, that includes new pre-calculated RT look-up tables, which replace on-the-fly calculations using MODTRAN\textsuperscript{®} [16,17] (the MODTRAN trademark is being used with express permission of the owner, the United States of America, as represented by the United States Air Force). The portable C++ code enables
parallel processing on multicore/multiprocessor computer systems. Initial software has been installed on the Sensor Web at NASA Goddard Space Flight Center, where it is currently atmospherically correcting new data from the EO-1 Hyperion and ALI sensors. Computation time is around 10 s per data cube for a single processor.

Further development is anticipated to implement the FLAASH software on board the Intelligent Payload Module for NASA’s planned HyspIRI mission, where it would generate data products in near-real time for Direct Broadcast to the ground. The rapid turn-around of data products made possible by on-board atmospheric correction would benefit a broad range of applications in areas of emergency response, environmental monitoring and national defense.

2. FLAASH ALGORITHM OVERVIEW

2.1 Radiative transfer equation and solution

FLAASH solves for the pixel surface reflectance $\rho$ using a standard at-sensor radiance equation which may be written as [1,6,8,18]

$$L^* = \frac{a \rho}{1 - \rho eS} + \frac{b \rho e}{1 - \rho eS} + L^*_a$$

(1)

Here $\rho_e$ is a spatially averaged surface reflectance, $S$ is the spherical albedo of the atmosphere from the ground, $L^*_a$ is the radiance backscattered by the atmosphere, and $a$ and $b$ are coefficients that solely depend on atmospheric and geometric conditions. The second term corresponds to the radiance from the surface that is diffusely transmitted into the sensor, giving rise to the “adjacency effect”. These quantities are all implicitly wavelength dependent and therefore bandpass-dependent, and are computed by integrating over the sensor spectral response functions.

The values of $a$, $b$, $S$ and $L^*_a$ in Equation 1 are determined from MODTRAN outputs of total and direct-from-the-ground spectral radiances computed at three different surface reflectance values, such as 0, 0.5 and 1. The latest versions of FLAASH use MODTRAN5 [16] with defaults of 5 cm$^{-1}$ spectral resolution for hyperspectral images and 15 cm$^{-1}$ resolution for multispectral images. The model atmosphere is based on user selections of one of the six standard MODTRAN atmosphere types (Tropical, Mid-latitude Summer, Mid-latitude Winter, Sub-artic Summer, Sub-artic Winter and U.S. Standard) and one of the four standard aerosol types (rural, urban, maritime and tropospheric). The viewing and solar angles of the measurement and nominal values for the surface elevation, aerosol type and visible range (visibility) for the scene are initially specified, as well as the sensor wavelength band centers and instrument functions.

The variable atmospheric quantities that have the greatest impact on VNIR-SWIR wavelengths are the column water vapor and the aerosol amount. Strong water absorption bands are located around 0.82, 0.94, 1.13, 1.4 and 1.9 µm. The centers of the stronger bands are usually too opaque for retrieving useful surface information. However, elsewhere the atmospheric correction processing can compensate for this absorption using a water amount derived from hyperspectral measurements at one of the near-IR bands. Here the MODTRAN calculations are iterated over a series of varying column water vapor amounts, and the water vapor is retrieved from the near-IR bands on a pixel-by-pixel basis [11]. In the next step an automated visibility estimate for the scene is retrieved using suitably “dark” pixels, which provide a stable estimate of the upwelling radiance scattered by aerosols [13]. A refined water vapor retrieval with the updated visibility is then performed to establish the atmospheric description used to calculate the surface reflectance.

The averaging implied in $\rho_e$ is a convolution with a spatial point spread function (PSF) [13]. Strictly speaking, different PSFs apply when $\rho_e$ appears in the numerator and denominator of Equation 1. However, since $\rho_e S$ is generally very small, we approximate the denominator PSF with the numerator PSF, which describes the upward diffuse transmittance. $\rho_e$ is estimated from an approximate form of Equation 1,

$$L_e^* = \frac{(a+b) \rho_e}{1 - \rho_e S} + L^*_a$$

(2)
Here \( L_* \) is the radiance image convolved with the PSF. The convolution is performed with a fast Fourier transform method using a reduced resolution image for further time savings. Then Equation 1 is then solved for \( \rho \). Finally, an optional post-processing step called spectral polishing may be used to suppress atmospheric residuals in the reflectance spectra [9,11].

The most time-consuming part of the MODTRAN atmospheric simulation is the calculation of multiple scattering, which is needed to accurately represent heavy aerosol loadings and Rayleigh scattering. The best balance of speed and accuracy is provided by MODTRAN’s “DISORT scaling” method [14]. This method works by first running fast two-stream scattering calculations over the full wavelength range, then spectrally rescaling the results by interpolating from a handful of more time-consuming, but more accurate, discrete ordinates calculations performed at atmospheric window wavelengths. Further runtime speed improvements have been achieved by replacing these on-the-fly calculations with large, pre-calculated lookup tables, as described below.

### 2.2 Wavelength and spectral smile recalibration

FLAASH provides the option of recalibrating hyperspectral sensor wavelengths using sharp molecular absorption features in the atmosphere, such as water vapor, oxygen and carbon dioxide bands. The feature matching is based on a transformed spectrum called the Normalized Optical Depth Derivative (NODD) [19] which minimizes sensitivity to both the surface reflectance and the molecular column density. The implementation in FLAASH uses an average spectrum of 10 to 20 fairly bright pixels chosen at random. The recalibration technique essentially consists of matching MODTRAN-calculated spectral shapes with the average spectrum by shifting the channel center wavelengths until the best NODD fit is obtained. Since most hyperspectral sensors contain more than one spectrometer, a wavelength shift is determined separately for each spectrometer and then applied to all of its channels.

While hyperspectral sensors are designed to provide nominally uniform spectral and radiometric calibration across the image, in practice the data may exhibit significant wavelength and slit function dependences on the image column. Contributors may include uncorrected spectral “smile” and misalignment between the detector array and the spectrometer. We refer to all of these dependences as smile effects. FLAASH can accept column-dependent calibration information as optional inputs and adjust the atmospheric compensation accordingly. FLAASH-C develops and applies separate atmospheric corrections to selected individual columns and interpolates the correction in between those columns. The smile compensation may be combined with the wavelength recalibration option; here, a separate recalibration is performed on each selected column. Smile compensation is now included in the default FLAASH-C atmospheric correction for NASA’s EO-1 Hyperion sensor [14].

### 2.3 Ground-leaving radiance output

Many hyperspectral sensors cover the atmospheric window region between ~2.0-2.4 \( \mu \text{m} \), where thermal emission from hot materials can be observed. To characterize the emission in terms of surface temperature and emissivity it is necessary to remove atmospheric and solar dependences from the measured spectrum. Accordingly, a new output, atmospherically corrected ground-leaving spectral radiance, has been added to FLAASH-C. This is the spectral radiance with atmospheric absorption and scattering removed—i.e., the product of the FLAASH-retrieved reflectance and the solar function. This quantity contains both solar reflected and emitted components. Figure 1 shows an example of ground leaving radiance for a Hyperion image of a forest fire. The solar reflected component can be estimated using cooler pixels in the scene and removed, leaving the emitted component, from which a map of elevated surface temperature can be derived. Besides the emission from the surface, the estimated thermal signature depicted here also exhibits molecular emission at wavelengths around 2000 nm from the hot CO\(_2\) released by the fire, and elevated radiance in the visible wavelengths due to scattering by the optically thick smoke plume. The results of such analysis should facilitate identification and characterization of forest fires as well as geothermal phenomena such as lava flows and hot springs.
2.4 Calculation speedup using Look-Up Tables (LUTs)

For hyperspectral imagery the two most time-consuming parts of the FLAASH atmospheric correction are the MODTRAN computations and the solution of the radiative transfer (RT) equation 1 for each spectral channel of each pixel. A timesaving strategy that we developed for the latter is based on spatially resampling the pixels into “superpixels” that are assigned common water vapor amounts and \( \rho_e \) values [10]. With 4x4 superpixels the procedure speeds the solution by around a factor of four. With this speedup, the MODTRAN calculations then consume typically half or more of the remaining computation time in FLAASH. Achieving significant further time savings therefore requires either speeding up or eliminating the MODTRAN calculations. An obvious approach is to construct pre-calculated look-up tables (LUTs) of the RT quantities, which may be interpolated during FLAASH execution to match the conditions of the observation. The challenge is to build a general set of tables for all sensors and viewing conditions, given the need to cover not only a wide range of atmospheric parameters and surface elevations but also a wide range of sensor altitudes for both airborne and spaceborne platforms and off-nadir as well as nadir views.

In an initial study of LUT feasibility we addressed the simpler problem of constructing LUTs for the JPL AVIRIS sensor viewing nadir from 20 km AGL. This allowed us to develop appropriate parameter grid spacings and evaluate interpolation errors. In our latest work we have extended the LUT approach to satellite platforms with nadir and moderately off-nadir views. The view-independent atmospheric variables in the LUT are the solar zenith angle (sza), visibility (in km) or aerosol optical depth (AOD), surface elevation, and column water vapor amount. The viewing geometry is expressed in terms of two angles, an elevation angle relative to nadir and an azimuth angle relative to the solar azimuth. Since the viewing geometries of the intended missions are likely to be no more than ~20 degrees off-nadir, two off-nadir geometries were selected to span this 20 degree space, with each geometry sampled at four solar azimuth angles. A finer gridding of the other four parameters mentioned above was selected, producing a total of approximately 435,000 simulations. For some variables, such as the solar zenith angle and aerosol optical depth, the grid parameters were sampled on a nonlinear scale to produce a more uniform separation between the model calculations. The column water vapor amount represents the largest of the grid dimensions, with 20 linearly-spaced values to account for the diversity of climates and seasons. The base model atmosphere was taken as the MODTRAN Tropical model to enable coverage of a very wide range of water amounts without saturation. Since VNIR-SWIR spectra are rather insensitive to atmospheric temperature and pressure, the Tropical model turns out to be a satisfactory choice for most, if not all, atmospheric conditions. The aerosol model was the MODTRAN rural aerosol, which works well over most land scenes. The calculations, performed over the 350 nm-2500 nm wavelength range at 5 cm\(^{-1}\) resolution, required roughly one week on a single computer, and produced around 16 GB of data containing the FLAASH RT parameters.
The RT database was then reduced to a manageable size for on-board processing using a principal component analysis (PCA) transformation. In this representation the data are expressed as linear combinations of orthogonal basis vectors, and, by restricting the combination to a subset of the most significant basis vectors, constitutes a reduced-dimensional, compressed database. The standard method for PCA using singular value decomposition (SVD) analysis derives the basis vectors as eigenvectors of the data covariance matrix. However, the calculation of the covariance matrix and subsequent diagonalization requires far too much computer memory to efficiently apply to the LUT data due to its very high dimensionality, equaling the number of wavelength values. Therefore an alternative method based on the Nonlinear Iterative Partial Least Squares (NIPALS) algorithm [20] was adopted. This iterative approach requires much less computer memory and permits the successive calculation of basis vectors until an acceptable level of accuracy has been obtained. By retaining 24 to 32 eigenvectors, sufficient for sub-percent-level residual for the entire LUT, the database was reduced to ~100 MB, for a compression ratio of 160:1. Interpolation from the LUT is performed linearly on all dimensions except for the solar azimuth angle, where the interpolation is linear in angle cosine to provide better accuracy.

An illustration of the new LUT-based FLAASH-C atmospheric correction with data from the NASA EO-1 Hyperion sensor is given in Figure 2. Agreement with the corresponding MODTRAN-based correction is excellent (generally at the sub-percent level); the LUT-based and MODTRAN-based results overlay almost perfectly at nearly all wavelengths, the exceptions being at the edges of the water absorption bands.

Figure 2. FLAASH-C reflectance retrievals from Hyperion data using the original MODTRAN-based method (blue) and the LUT (“GLUT”) method (red).

A timing test was performed running FLAASH-C on a single PC processor with both Hyperion hyperspectral data and Advanced Land Imager (ALI) multispectral data. Operation with the LUT saved nearly half the computation time with ALI data and 2/3 of the computation time with Hyperion data. The computational saving is smaller with the ALI data because water vapor retrievals, and the associated MODTRAN calculations, are not performed with those data. Total FLAASH-C execution time was 14 s with a typical ALI image (~2000x3500 pixels) and 12 s with a typical Hyperion image (~256x3500 pixels).

The remaining envisioned speedup for FLAASH, which will dramatically shorten both the atmospheric retrieval and correction times, will be operation with parallel processors. FLAASH-C is currently parallelized with MPI (Message Passive Interface) routines. We anticipate providing multi-threading parallelization for FLAASH-C in the near future. With either method the PSF spatial convolution step is parallelized by band, while the other processing steps are parallelized spatially, i.e., by image line or pixel.
3. SOFTWARE DEPLOYMENTS

The newest capabilities of the FLAASH-C code, described in the previous sections, are now operational on NASA’s Goddard Space Flight Center Elastic Cloud and SensorWeb systems. The code is currently processing new images from both ALI and Hyperion automatically as they are downlinked from the EO-1 satellite. The code has also been integrated into a Web Coverage Processing Service (WCPS), allowing NASA researchers the ability to process archival EO-1 data upon request through a web-based interface. This latest capability allows a user to selectively reprocess a specific dataset with different runtime options as necessary, which is then made available for download to the user through the web service.

For upcoming HyspIRI planning missions, three hardware systems are currently being prototyped by Goddard Space Flight Center and the Jet Propulsion Laboratory to support on-board spectral image processing. The SpaceCube is a prototype FPGA-based system based on VIRTEX-5. Goddard is also developing a multicore system for the HyspIRI Intelligent Payload Module (IPM) based on the DOD-provided MAESTRO chip, using a Tilera64 design but with 49 cores. The IPM is a hybrid system that also contains radiation-hardened FPGAs for low-level data processing. When embedded on these parallel-processing architectures, FLAASH-C is anticipated to provide a rapid and fully automated atmospheric correction capability for on-board and direct-broadcast data processing.

4. SUMMARY AND CONCLUSIONS

The FLAASH atmospheric correction code has been developed by Spectral Sciences, Inc. and the US Government over the past fifteen years to provide accurate and fast first-principles atmospheric correction of VNIR-SWIR hyperspectral and multispectral imagery. This paper describes the latest version of the algorithm, as implemented in both IDL and C languages, including the recent development of MODTRAN look-up tables (LUTs) for improved speed and portability. FLAASH-C with LUTs for satellite sensors is currently operational on NASA’s Sensor Web, where it provides automated processing of imagery from the EO-1 Hyperion and ALI sensors. FLAASH-IDL is in widespread use throughout the remote sensing community as an ENVI add-on. The latest upgrades to that code, which offer more accurate visibility retrieval [13], are being provided with ENVI 5.0.

Validation of FLAASH has been ongoing. Our studies include comparisons of retrieved reflectances with “ground truth” spectra [9], comparisons of visibility retrievals using different FLAASH methods [13], and comparisons of derived aerosol atmospheric properties with results from field radiometers [15]. Comparisons of FLAASH with other atmospheric correction codes have also been reported [9,21,22]. Computation times are now as low as ~10 s per image per computer processor, and the latest visibility retrieval algorithms broaden the range of scenes for which automated aerosol retrieval may be performed. With these advances, accurate, real-time, on-board atmospheric correction of hyperspectral and multispectral imagery is within reach.

ACKNOWLEDGEMENTS

We are grateful to NASA for support of the FLAASH LUT development and Sensor Web code integration under an SBIR program, Contract No. NNX11CF11P.

REFERENCES


