Scanning for Extinct Astrobiological Residues and Current Habitats (SEARCH)

William R. Dieter Robert A. Lodder James E. Lumpp, Jr
Department of Electrical and Computer Engineering
University of Kentucky
Lexington, KY 40506
859-257-1768 859-257-9232 859-257-3895
dieter@uky.edu lodder@uky.edu jel@uky.edu

Abstract—SEARCH is a new method to seek evidence of extinct life and potential habitats. SEARCH combines innovative spectroscopic integrated sensing and processing with a hyperspace data-analysis algorithm. Using UV, visible, and near-IR spectroscopic integrated sensing and processing, SEARCH is designed to explore and quantitatively assess a local region on the surface of a planet or moon as a potential habitat for life, past or present. In the course of collecting geological data, SEARCH spectrometry can investigate planetary processes of relevance to past habitability, including the role of water. In addition to its own investigations, SEARCH can be used at a distance to guide a rover to areas of interest for application of other analysis tools. Several prototype implementations of SEARCH have been developed and initial experimental results are presented along with a design for a full-scale version of SEARCH for Mars exploration.

TABLE OF CONTENTS

1. INTRODUCTION

The Scanning for Extinct Astrobiological Residues and Current Habitats instrument (SEARCH) is a new method to gather geological data and to identify evidence of life, extinct life, and potential habitats. It is suitably small and light to mount on the mast of an autonomous rover such as the Mars Exploration Rovers (MERs) or the Mars Science Laboratory (MSL), and can scan ranges of up to 10 meters. Using a laser-diode array, photodetectors, and on-board processing, SEARCH combines innovative spectroscopic integrated sensing and processing with a hyperspace data-analysis algorithm [1]. Ultraviolet (UV), visible, and near-infrared laser diodes illuminate target points using a precomputed sequence, and a photodetector records the amount of reflected light. For each point illuminated, the resulting reflectance data is processed to separate the contribution of each wavelength of light and classify the substances present.

SEARCH is designed to explore and quantitatively assess a local region on the surface of a planet or moon as a potential habitat for life, past or present. In the course of collecting geological data, SEARCH spectrometry can investigate planetary processes relevant to past habitability, including the role of water. SEARCH is also sensitive to the presence of endolithic biofilms and soil crusts (caused by photosynthetic cyanobacteria capable of living under extreme conditions)[2]. In addition to its own investigations, SEARCH can be used to guide a rover to areas of interest for application of the full suite of analysis systems on board the rover.

SEARCH's laser diode array can obtain profiles of a wide variety of organisms, fossils, or other remnants of once-living organisms. A sufficient number of discrete reflections at different wavelengths from a target provide a unique profile. In addition, amino acids, carbohydrates, polycyclic aromatic hydrocarbons (PAHS), and more complex organic compounds can be identified. Minerals of astrobiological interest, such as amphiboles, silicates, limestone, jarosite, hematite, oxides, feldspars, plagioclase, smectite, halites, apatite, hydroxyapatite, sulfides, sulfates, and others [3], [4], [5] are detectable. Water can be found in liquid or solid phase or as hydrates, enabling a rover to follow the water in its search for life. The data collected by SEARCH are compared to an extensive data bank library of responses from selected terrestrial materials assembled prior to launch. New materials can be added to the library as they are discovered in the rover's environment. By converting data into *tokens,* small but meaningful groups of bytes that identify the substance or substances present, SEARCH can obtain the greatest significance from the data, and greatly reduce the bandwidth required to transmit it back to Earth for comparison with the library.

Several prototype implementations of SEARCH have been developed and are being tested at the University of Kentucky. A full-scale implementation of SEARCH is being designed with 24 lasers in discrete wavelengths

---

[1] 0-7803-8870-4/05/$20.00© 2005 IEEE
[2] IEEEAC paper #1439, Version 3, Updated December 10, 2004

between 300 nm and 2400 nm with 5 rows of each wavelength. This full-scale version is designed to consume less than 4 Watts and weigh less than 600 grams. The rugged laser diodes and detectors allow SEARCH to be packaged in a small, rugged, space qualified package. The following sections describe the theory behind SEARCH and the realization of the system with a series of prototypes. Section 2 outlines the theory behind SEARCH, and Section 3 describes experimental verification of these theories. Section 4 discusses the issues in the implementation of SEARCH instruments. Section 5 describes the SEARCH prototypes currently under development, and Section 6 provides a summary of the SEARCH project.

2. THEORY BEHIND SEARCH

SEARCH combines near-infrared, visible, and ultraviolet spectroscopy with a powerful statistical classification algorithm to detect and identify biomarkers of life, either living or extinct, and identify current habitat indicators or other interesting geological features from a distance. Once such areas have been identified, the rover can approach them and use other instruments for further investigation.

Spectroscopy of Biomarkers and Geology

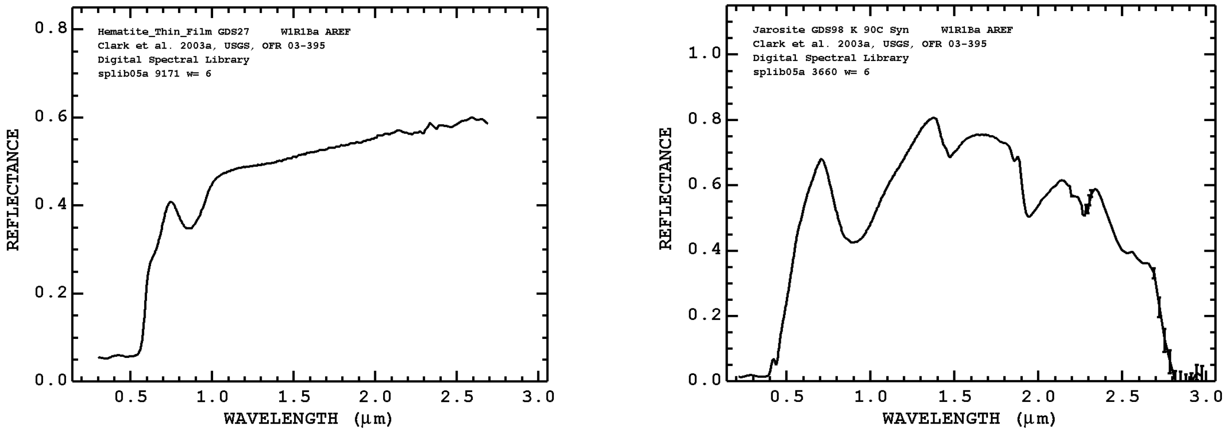
The substances present in biomarkers and geological features determine what spectra SEARCH should use. Water is thought to be a key component to life. The presence of frost or ice is easily detected, along with its temperature (which changes the hydrogen bond strength) [6]. Significant amounts of subsurface ice have been indicated by the Odyssey orbiter, some within several centimeters of the surface, and surface temperatures have been recorded above freezing. However, despite laboratory simulations that show liquid water existing under Martian conditions, various interpretations of the triple point of water and environmental factors affecting it leave the subject in doubt [7], [8].

On Earth, cyanobacteria are known to live in harsh environments where water is intermittently available [9]. These cells form *phototrophic endolithic biofilms,* bacterial films penetrating stone surfaces that are capable of utilizing carbon dioxide in the presence of light as a source of metabolic carbon, on rocky surfaces that derive energy from sunlight. SEARCH is designed to detect and map the locations of such rock outcroppings based on their chemistry and the chemistry remaining from the films. In addition to producing molecules that utilize light as an energy source, cyanobacteria on Earth produce rugged compounds that serve as UV sunscreens [10]. That property can be used with geological context as a chemical marker for the presence of previous biofilms. Conditions supporting carbonate precipitation have been inferred from cyanobacteria because their calcification requires the supersaturation of water with respect to $CaCO_3$ minerals [11]. Repeated calcification of microbial mats and biofilms, usually dominated by cyanobacteria and/or nonphototrophic bacteria, leads to the development of stromatolites, laminated reef-like structures that once were widespread in the marine domain. Three critical steps in cyanobacterial biofilm calcification have been identified: (1) the preliminary supersaturation of the macroenvironment with reference to $CaCO_3$ minerals, (2) the swing of the carbonate equilibrium to exceed the critical supersaturation for $CaCO_3$ mineral formation, and (3) the progression of seed crystal formation.

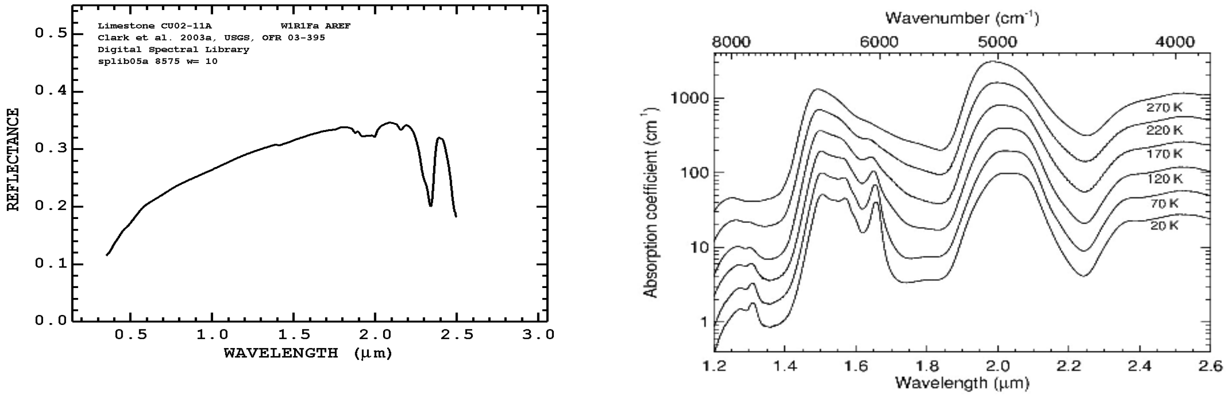
Virtually every organic compound (e.g., polycyclic aromatic hydrocarbons, paraffins, carboxylic acids, and sulfonic acids) has a near-IR spectrum that can be measured, including two classes of terrestrial biomarkers, lipids, and amino acids. Near-infrared spectra consist of overtones and combinations of fundamental mid-infrared bands, giving near-infrared spectra a powerful ability to identify organic compounds while still permitting some penetration of light into samples [12].

SEARCH is also capable of remotely sensing and analyzing of a number of interesting geological and habitat features. Hematite (see Figure 1(a)), a type of iron oxide found in some Martian sediment, has a distinctive spectrum that differentiates it from jarosite (see Figure 1(b)) (found on Earth in association with acidic lakes or hot springs). Hematite nodules probably formed when a stream of groundwater rich in dissolved minerals, including iron, encountered chemical conditions that caused the iron to precipitate. Some of these nodules have a homogeneous internal structure, while others appear to have growth rings or an outer rind. These features suggest that the concretions may have grown in episodes rather than all at once. The size and spacing of hematite nodules within sandstone on Earth depends on many factors, including the salinity, pH, flow path, and flow rate of groundwater through the porous rock. Previous experiments have suggested that the presence of bacteria could speed nodule formation [13].

Other sedimentary rock such as limestone (see Figure 1(c)) also has distinctive UV/visible/near-infrared spectra that can be mapped by SEARCH. The presence of frost or ice is easily detected, along with its temperature (which changes the hydrogen bond strength, see Figure 1(d)). Researchers recently distinguished fossils that formed in ancient flood plains on earth from those that formed in stagnant lakes using a new technique to examine a fossil and identify the setting in which it was initially buried [14]. The method works because ancient soils have a marked effect on the geochemical signature of the bones. Rare earth elements in the lanthanide series appear in only minute concentrations in living creatures. However, when an animal dies, much larger quantities of lanthanide elements in the soil leach into the crystallizing bones, replacing calcium atoms of similar size. These changes, which take up to 30,000 years to occur, reflect a particular soil's geochemistry, permanently



(a) Hematite spectrum[3] (b) Jarosite spectrum



(c) Limestone spectrum (d) Water spectra at different temperatures

Figure 1 - Spectra in the SEARCH wavelength range for hematite, jarosite, limestone, and water have peaks at different wavelengths, allowing them to be differentiated.

capturing a snapshot of soil conditions in the fossilized bones. Researchers have been applying mass spectrometry to analyze the elements of such fossilized specimens, but mass spectrometry is a contact measurement. The outer shell electronic transitions of lanthanide atoms appear in the near-infrared spectral region because the atoms are so large. Because they are electronic transitions, these signals are much stronger than the molecular vibrational overtone and combination bands ordinarily observed in the near infrared, and they might be used as indicators of interesting regions containing fossils.

These substances, and other substances of interest, can be detected by examining their spectral response at strategically chosen wavelengths of light. The full-scale version of SEARCH is expected to use laser diodes to illuminate a target with 24 different wavelengths, selected to match “interesting” substances found in biomarkers and geological features. Table 1 lists each of the wavelengths and which substances have distinctive absorbance or reflectance signals at those wavelengths.

Quantitative determination of the components of soil and stone (e.g., Figure 1(a)-(c), which demonstrate the ability to identify stones that form in the presence of water) can be accomplished down to the noise limit using modern mathematics and parallel pattern-recognition techniques. Furthermore, qualitative analysis, in which the constituents of soils and stones are identified, is also possible (with benefits in analytical speed and accuracy accruing to those using near-IR spectrometry and mathematics to analyze such samples). In contrast to Raman spectrometry, in which only one in 10^9 photons may be Raman scattered, in near-IR spectrometry one in 10 or even more photons are usefully scattered by soils and stones.

To further increase the signal-to-noise ratio, SEARCH uses Walsh-Hadamard encoding sequences of light pulses. In a Walsh-Hadamard sequence multiple laser diodes illuminate the target at the same time, increasing the number of photons received at the photo detector. The Walsh-Hadamard sequence can be demultiplexed to individual

---

[3] USGS Digital Spectral Library splib05a. U.S. Geological Survey, Open File Report 03-395, http://pubs.usgs.gov/of/2003/ofr-03-395/ofr-03-395.html

Table 1 - Wavelengths used by SEARCH

Substances	Wavelength(s)
Residual UV protectant molecules from biofilms	340 nm
Residual phototrophic endolithic biofilms, sulfides	470 nm, 530, nm, 635 nm
Water, -OH, =C=H	1445 nm, 1940 nm,1982 nm
-CH2, -CH3, various -C-H stretches, calcite	1680 nm, 1722 nm, 1734 nm, 1759 nm, 1778 nm, 1818 nm
-NH2, amines, protein amides	2100 nm, 2139 nm, 2180 nm, 2190 nm, 2208 nm, 2230 nm
Lipids, -CH, calcite, limestone	2270 nm, 2310 nm, 2336 nm, 2348 nm
Baseline for reflectance and target angle	1064 nm

wavelength responses with a matrix-vector multiply [15]. Benefits of generating encoding sequences by this method include equivalent numbers of on and off states for each sequence and a constant number of diodes in the on state at each resolution point of a data acquisition period.

BEST Sample Classification

Once the spectrum of a sample has been collected, it must be classified to determine the substance present. The Bootstrap Error-adjusted Single-sample Technique (BEST) [1] is the analytical basis of SEARCH, and the foundation for the geochemical library. Spectra recorded at 24 wavelengths are represented as single points in a 24-dimensional hyperspace. In this scheme, similar samples produce similar spectra that project as "probability orbitals" or "clusters" into similar regions of hyperspace. The BEST metric is a clustering technique for exploring these distributions of spectra in hyperspace.

A sample spectrum is compared to each substance in a geochemical library based on its distance, measured in standard deviation (SD) units, from the known substances. BEST elegantly handles non-uniform standard deviations surrounding each substance, allowing more precise discrimination than other metrics, like Mahalanobis distance [1]. A sample within 3 SD units of a substance is considered to be composed of the matching substance. Any substance more than 3 SD units away from a known substance is considered an unknown substance.

The BEST algorithm can be suitably approximated using multiple linear regression (MLR) to substantially reduce computational requirements. In this implementation, BEST SD units are pre-calculated in a large number of directions from the population means and MLR is used to fit the standard deviation contours as a function of direction. With a sufficient number of terms, the MLR version of the algorithm can predict BEST distances within 5% of true value (see Figure 2 for an example). With the MLR method, the BEST classification algorithm can be performed in situ, allowing a rover to classify large areas, only notifying ground control when an interesting substance is found. An initial library can be computed on Earth based on substances likely to be found in the target environment. When a substance unknown to the BEST library is found, the rover can sample nearby points with similar spectra to create a new library entry for the new substance. Scientists on Earth can determine the type of substance present either by further analyzing raw spectra of the substance provided by SEARCH or by using data from the rover's other instruments.

On-board classification provides a powerful redundancy removal method. Large image regions of spectral data that represent the same substance can be described by a single substance identifier. Moreover, SEARCH can be programmed to scan continuously until any one of a set of substances is found, removing the need to communicate with ground control until an interesting substance is found.

The principal competing technologies for SEARCH are thermal emission spectrometry [16] and Raman spectrometry [17]. The Mini-TES on board the MERs is an infrared spectrometer that can determine the mineralogy of rocks and soils from a distance by detecting their patterns of thermal radiation. It requires passive excitation and provides lower signal-to-noise ratio than should be available using active laser excitation with SEARCH. Raman spectrometry suffers from the generally poor efficiency of Raman scattering, an inelastic process. The Combination Laser Absorption and Raman IR Spectrometer (CLARIS) instrument is a miniature near-IR spectrometer that combines four-instruments-in-one and can perform surface mineralogical studies of Mars. The all-solid-state spectrometer is based on a single tunable diode laser source

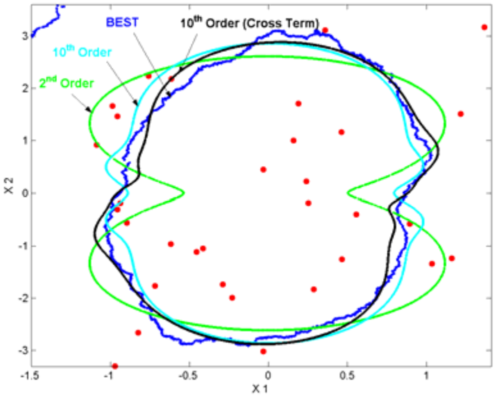
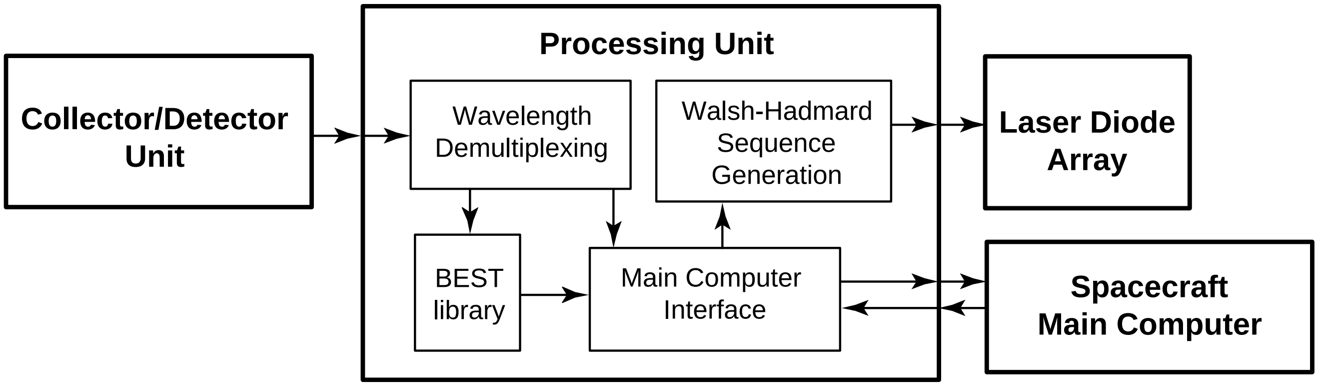
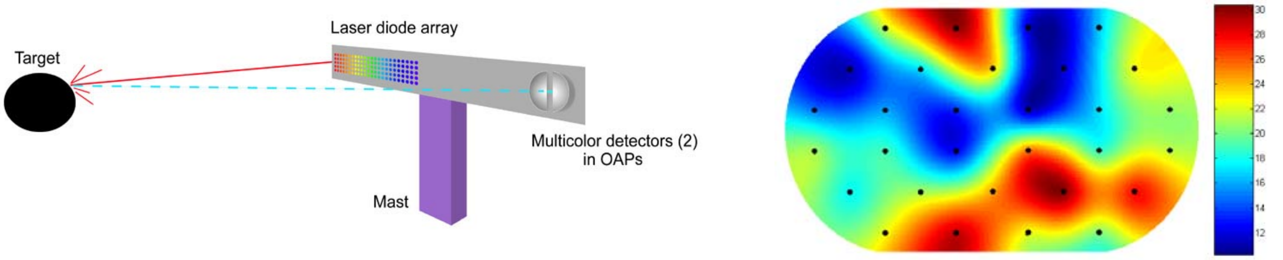


Figure 2 - Modeling BEST SDs. Each line represents a one BEST SD contour. The actual BEST line (dark blue) is estimated by a second order (green), 10th order (light blue), and 10th order with cross terms (black) equation.



(a) Block diagram of SEARCH



(b) Laser diode array and collector/detector unit configuration

(c) Sampling grid over target area

Figure 3 - SEARCH is composed of (a) the Laser Diode Array, the Collector Detector Unit, and the Processing Unit. The SEARCH laser diode array (b) illuminates a target sample with an orthogonal Hadamard pulse sequence. Raster mapping at the black points (c) builds an image indicating what substances are present and their concentrations.

at 1.3 mm (or in combination with a second source at 0.78 mm). CLARIS relies on lenses and InGaAs array detection, and therefore is not as rugged as SEARCH.

3. IMPLEMENTATION OF SEARCH

SEARCH is composed of three principal components: a laser diode array (LDA), a Collector/Detector Unit (CDU), and a processing unit, as shown in the block diagram in Figure 3(a). The LDA for the full configuration contains 5 rows of 24 laser diodes, with each row containing all of the wavelengths listed in Table 1. The CDU consists of a parabolic concentrator focusing light from the LDA on two photodiodes. The processing unit drives the LDA, samples data from the CDU, and classifies the samples using the BEST algorithm, described in Section 2 above. With the LDA and CDU mounted on a mast, as shown in Figure 3(b), SEARCH builds an image by illuminating points in a region with the LDA (see Figure 3(c)) and sensing the results with the CDU.

Laser diodes are positioned within the array so that each laser diode illuminates a different point. When the remote sensing mast on which it is mounted rotates, the spot illuminated by each laser diode in the new position is the same spot previously lit by an adjacent laser diode. Multiple rows of laser diodes are included to reduce the amount of motion needed to complete a scan. With five rows, SEARCH can make five times as many scans as with a single row for the same amount of movement. An added benefit is that should some laser diodes fail, SEARCH can compensate by using laser diodes from a different row to cover for the failed laser diodes. Though scanning speed would be reduced, SEARCH would still be able to collect data using all of its wavelengths.

SEARCH can operate either during day or night. The Walsh-Hadamard sequence provides high enough signal-to-noise ratio to produce usable results in direct sunlight.

SEARCH does not rely on sunlight to illuminate targets, however, and reduced background light at night will reduce the signal-to-noise ratio. Moreover, nighttime operation will allow the rover to map the surrounding area at a time when it and its other sensors are inactive, providing a valuable aid to planning the next day's operations.

SEARCH can scan the surrounding area in one of four modes, depending on operational needs:

(1) **Survey** mode will be used to quickly examine the surroundings whenever the rover moves into a new area. The sensor will scan at low resolution in concentric circles, starting from the ground around the rover and working outward until it reaches 10m.

Table 2 - Time required for each scanning mode

Scanning Mode	Area Covered		Pixels	Scan Time
	Azimuth	Elevation		
Survey	360°	-60° to -12°	52,920	~1 min.
Mapping	360°	-60° to -12°	5,292,000	~1.5 hrs.
Directional Mobile	TBD	TBD	TBD	TBD
Directional Stationary	θ_{az}	θ_{el}	$\theta_{az}\times\theta_{el}\times10^6$	$\theta_{az}\times\theta_{el}\times10^3$ sec.

(2) **Mapping** mode is similar to survey mode, but operates at a higher resolution. It will be used to get a detailed picture of the surrounding area, though it will mostly be used at night due to the amount of time required to perform a detailed mapping.

(3) **Stationary Directional** mode will be used to raster scan rectangular areas in arbitrary directions while the rover is not moving. This mode will be used to scan rock outcroppings, crater walls, or other large, primarily vertical surface features within 10 m of the rover.

(4) The **Mobile Directional** mode will be used to raster scan small rectangular areas while the rover is moving. The sensor may scan in the direction of motion to seek unusual substances in or near the path of the rover, or may scan an area previously scanned to check for changes.

The amount of time required to complete a scan and the amount of data generated depend on the scanning mode, photodetector integration time, and number of samples taken. Measurement at one position is expected to take 1 ms.

A typical sensor mast assembly can rotate a full 360° in azimuth, but has a limited elevation range [20]. From approximately 2 m off the ground, SEARCH has a resolution of 1 cm at 10 m, assuming a 1 mrad pointing accuracy. When scanning the full 360° in the azimuth with 1 mrad steps, the instrument will sample approximately 6300 pixels per circular scan. 840 circular scans are required to cover an area around the rover from –60° below level to 10 m from the base of the mast on level ground. The amount of data generated depends on the scanning mode. Table 2 summarizes the area that can be covered in each mode, the number of pixels sampled, and the amount of time required to complete the scan. Though Survey Mode and Mapping Mode cover the same area, Survey Mode completes much more quickly because it only samples one tenth the pixels in both the azimuth and elevation.

The area to be covered in the directional modes will not be known until a scan is started. The amount of raw data required for each image is 16-bits per wavelength per pixel. Thus, the above full azimuth 10 m scan would require 233 MB (at 2 bytes/pixel/wavelength) to store raw wavelength data in an uncompressed format. Substances found in the BEST library can be represented with a 16-bit number. Using the 16-bit substance value for each recognized substance would reduce the above data requirement to 10.6 MB assuming all pixels are found in the BEST database. Both raw and substance data will likely compress very well assuming many adjacent pixels will contain the same substance as their neighbors.

Redundancy removal will play an important role in reducing the storage and transmission requirements for both raw, unprocessed wavelength data and classified sample data. Classifying samples with the BEST algorithm can be thought of as reducing the size of the 24-column 16-bit raw wavelength data with a 24-to-1 compression ratio to get the single 16-bit substance identifier. Classified samples are expected to be the primary data transferred from SEARCH to Earth, though raw wavelength data will be stored temporarily and transferred upon request. Additional data redundancy removal will be applied to raw and classified samples. When communication bandwidth between Earth and a rover is at a premium, SEARCH can continue to operate, transmitting data only when specific substances are found, reducing the communication requirement to close to zero when no interesting substances are present.

4. SEARCH EVOLUTION

SEARCH began as an experiment in pharmaceutical cleaning validation. Several experiments have been performed subsequently to test the approach. Based on this earlier work, we are developing our current generation prototype to test SEARCH in more realistic conditions as an integrated instrument on an interplanetary rover.

Pharmaceutical Validation

A remote sensing experiment was conducted to test the hypothesis that digital images of light scatter could verify the amount of contaminant present on a surface [18]. This pharmaceutical remote-sensing problem is similar to the remote sensing of chemistry on Mars. Light from a single wavelength laser was transmitted through commercially available Teflon printed microscope slides and the scattered light was imaged with a digital camera. The slides were treated with 200-microliter aliquots of various concentrations of bovine serum albumin (BSA) one day prior to scanning, and the solutions were allowed to evaporate. Quantification based on the means of the 30 scattering images collected from 30 locations on the slides resulted in an r^2 of 0.99 (shown in Figure 4). For this graph, the standard error of prediction (SEP) was 70 ng cm^{-2} (SEE= 50 ng cm^{-2}, with a 0.35% relative SEP over the range of concentrations. Cross-validation was performed with the f test at p= 0.05).

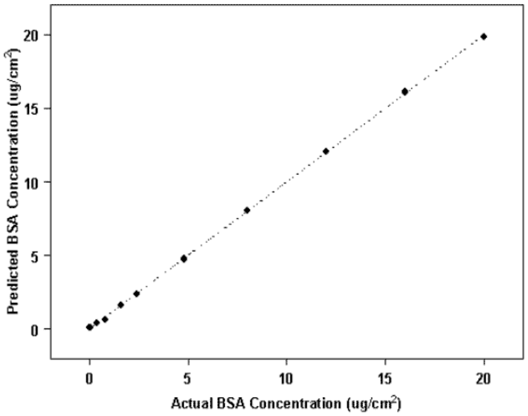


Figure 4 - Predicted vs. actual BSA concentration for a single laser system. Cross validation from mean images, r2= 0.99. ICI imaging of protein distribution.

These results suggest that laser light scattering is able to quantify proteins at levels more than adequate for pharmaceutical cleaning validation. The ability to detect such molecules via remote sensing will be a powerful addition to a Mars rover. An experiment in cleaning validation using raster-scanned pseudorandom orthogonally modulated laser diodes similar to the SEARCH instrument has also been conducted. The utility of laser scattering can be greatly increased by imaging more than a single point on the target surface at a time. In this way, the scanning speed can be improved markedly. The raster scanning approach can be employed in an integrated computational imaging (ICI) system with multiple lasers operated in parallel to image a surface using only a single detector. Together, the laser array and single detector comprise an ISP (integrated sensing and processing) system for ICI in cleaning validation and Mars exploration.

This ICI approach was demonstrated by scanning a slide with three wells containing 0, 10, and 20 micrograms/cm^2 BSA. A set of three orthogonal binary pulse sequences (see Figure 5) was applied to the lasers at each raster position on the slide surface. The covariance between the signal observed at the detector and each of the pulse sequences was proportional to the light scattering from each well (see Figure 6). In turn, the light scattering from each well was proportional to the BSA concentration on the surface (see Figure 7). BSA solution was uniformly distributed over identical surface areas for each sample and allowed to dry before scanning.

By integrating the sensing and processing functions, a simple, rugged, inexpensive imager can be constructed with excellent detection limits (<100 ng protein / cm^2). Scanning multiple points in each well permits mapping of protein distribution, similar to the mapping functions SEARCH will perform on Mars.

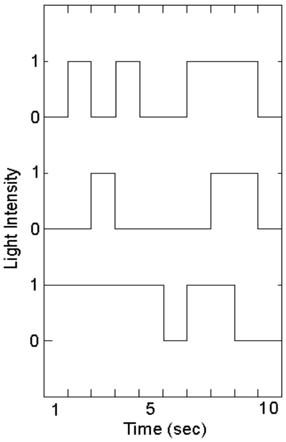


Figure 5 - Pulse sequences applied to each of the three lasers illuminating a target.

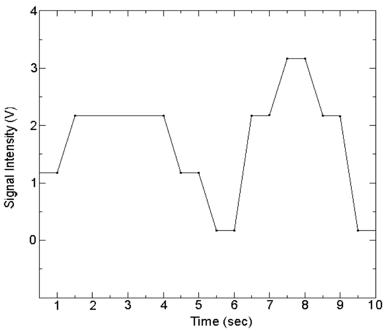


Figure 6 - Sample waveform (in volts, 2 Hz sampling) read from the detector as the three lasers illuminate different concentrations in the three wells

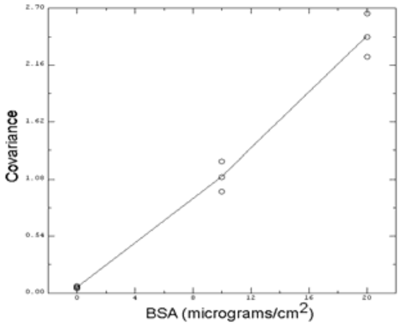


Figure 7 - BSA concentration is directly proportional to the covariance between the detector signal and the waveform applied to each of the three laser diodes.

Experiments with Biofilms

The SEARCH approach was first applied to imaging endolithic biofilms in a robotic instrument, assembled from commercial components. The instrument, based on liquid crystal tunable filters, has been successfully tested. The science hypothesis has been further confirmed in the lab with a number of heat-sterilized samples of zoetic materials, including remnant biofilms, organic, and inorganic compounds.

A simple SEARCH prototype sensor with a 2x2 light-emitting diode (LED) array and a silicon phototransistor detector was constructed in order to analyze a printed test target. The array was designed with each of the four LEDs modulated at different audio frequencies. Two LEDs had emission peaks at 530 nm and two LEDS had emission peaks at 630 nm. The prototype interfaced directly with a notebook computer which digitized the samples with a sound card, and all data were processed in Matlab v6.5. Fourier transforms were calculated on the received signals rendering the signal from each of the four LEDs. In this manner, the spatially resolved absorption spectrum [log(1/R)] of a particular surface could be determined directly from a comparison of peak heights in the FFT of the signal from photodetector.

The same process used in mapping a test pattern was applied to mapping a biofilm species (*Gloeocapsa*) on limestone. The experiment tested the hypothesis that the imager correctly differentiated regions containing blue-green algae from regions scraped clean with a rock abrasion tool. A grid pattern was scraped into a limestone rock as shown in Figure 8. The 4-diode sensor was passed over the top of the surface in 5 sweeps, and a 3-d mesh plot was constructed from the Fourier transformed data (shown in Figure 9). The intensity plot corresponds well to the grid pattern on the rock, indicating that the sensor could easily map biofilms deposited on the limestone surface.

In another experiment, biofilms were killed by baking to demonstrate the ability to detect them by their residual chemical signatures. In Figure 10, a biofilm is scanned in the near-infrared region before baking at 160 C for three hours (black line, top). The spectrum of clean limestone is subtracted from each of these three spectra. The principal chemical observed in the spectrum is water (1450 and 1940 nm) retained in the cells of the biofilm. After one baking, the spectrum of the rock appears in blue. A small negative residual spectrum of water appears because even clean limestone (the blank) contains a small amount of water. An additional baking produced the spectrum shown with the red line. Using the 470 nm, 510 nm, and 650 nm diode wavelengths like those in SEARCH produced the background-corrected spectra in Figure 11 after one and two bakings. These spectra confirm that nonliving biofilms still produce detectable spectrometric signatures.

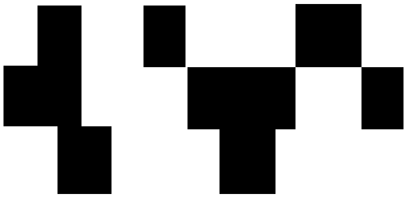


Figure 8 - Grid pattern abraded from the surface of limestone.

5. SEARCH PROTOTYPES

We are currently developing a family of prototype implementations of SEARCH to move the design from the laboratory 2x2 LED testbed to the final full 5x24 laser configuration suitable for use as an integrated instrument on a space vehicle.

The current SEARCH prototype being evaluated at the University of Kentucky is shown in Figure 12. It includes a 5x5 LED array, a photodetector circuit, and embedded microcontroller collectively referred to as the *scan unit*. A laptop PC running Matlab sends configuration and scan commands to the scan unit and is referred to as the *control PC*. The control PC is responsible for processing the scan data from the scan unit and classifying material in each scan using BEST. The scan unit in the current prototype is aimed manually at surfaces to be scanned and can perform scans either at the request of the control PC or at a regular pre-defined rate.

The LEDs in the 5x5 array were chosen in the visible range at commonly available frequencies for ease of testing. This also allowed classification of a diverse set of materials while keeping development cost low. Each LED is driven with an individual transistor switch, and can be switched on or off at rates up to 100 KHz. The scan unit can be set to perform a scan (a sequence of LED patterns) as a pre-defined set, or individual sequences can be requested by the control PC. The LEDs are housed and aligned by an aluminum frame to collimate their light, which would otherwise spread much more than light from more expensive solid-state lasers.

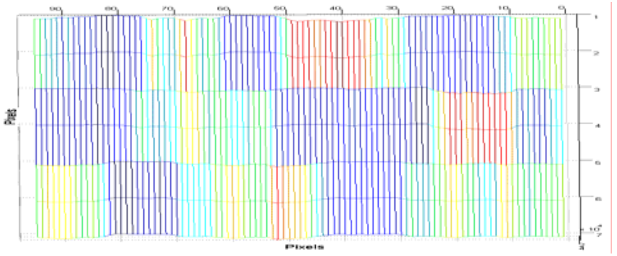


Figure 9 - 3-d mesh plot of 2-dimensional position vs. intensity. The cleaned areas on the stone appear as light blue, yellow or red, while the film-covered areas appear in dark blue.

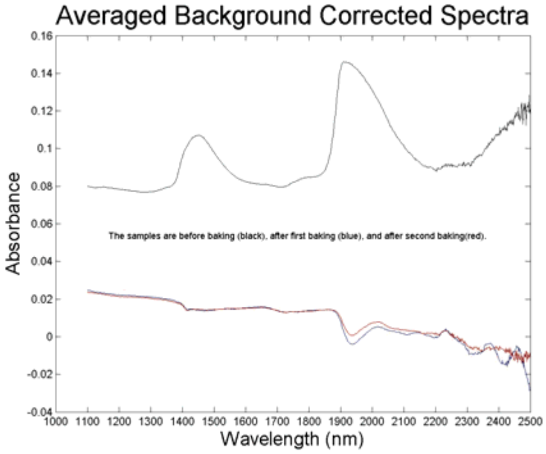


Figure 10 – Near-IR spectra of biofilm on stone before backing (black), after one baking at 160 C (blue), and after two bakings at 160 C (red).

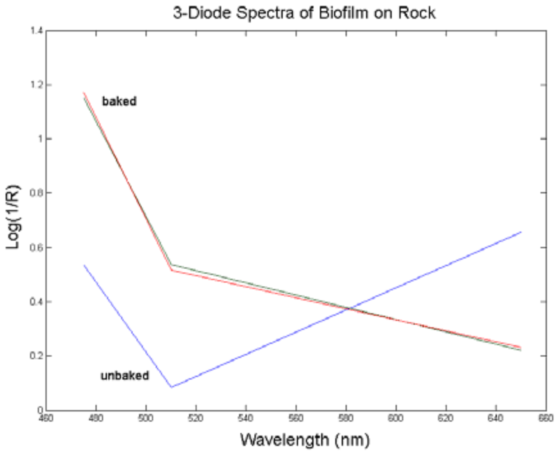


Figure 11 - Visible spectra of biofilm on stone before baking (blue), after one baking at 160 C (green), and after two bakings at 160 C (red).

The photodetector circuit uses a single detector with peak response in the frequency range covered by the LEDs in the array. The current prototype does not include a reflector for the detector as the response from the large spot size of the bright LEDs gives strong responses at up to 5 meters, which is more than sufficient for the current testing.

An 8051-based microcontroller is used to control the scan unit. The controller has a serial interface to the control PC and general-purpose I/O pins connect to the LED array. The controller uses a 100,000 sample per second analog-to-digital converter to sample the output of the photodetector. The scan unit uses very little power, the vast majority of which is consumed by the LEDs when they are on, and the unit can run off internal batteries or an external supply.

The control PC uses a serial connection and Matlab serial interface routines to control the scan unit. A Matlab program reads wavelength data from the scan unit, performs an inverse Hadamard transformation on the data to get the raw wavelength data, and then uses BEST to classify the sample.

As we gain experience with the current prototype we are continuing the development of future generations of the SEARCH system culminating in the final 5x24 laser array with the frequencies listed in Table 1 and packaged for use on a spacecraft.

Evolution of the SEARCH prototype

The current SEARCH prototype will provide the basis for evolution to the full-scale SEARCH instrument. The current 5x5 scan will be mounted on a pan and tilt unit that will allow the scan unit to raster over an area. This will allow each LED to target a specific point in space so that multiple frequencies can be aligned on a specific point in 3D space. Next, the LEDs will be replaced with solid-state lasers. This will allow scans at longer ranges and also at a higher resolution with a finer spot size. The PC control unit will then be replaced with a DSP processor, and finally, the same DSP will be used to control the laser array and the photodetectors. Once this integrated instrument is verified and validated it can be packaged and space qualified for use on future rovers.

Ultimate SEARCH implementation

The ultimate version of SEARCH is envisioned to be a self-contained unit using all the wavelengths described in Table 1. Mounted on the mast of a rover, SEARCH could guide the missions to regions containing interesting substances. The range the rover can cover will be greatly increased because it will be able to evaluate a region without having to make physical contact.

The 5x24 array of laser diodes will be pulsed in an orthogonal (Hadamard) sequence with approximately half of the laser diodes illuminated at each step to increase the amount of light incident on the photodiode, improving the signal-to-noise ratio (SNR) of the photodiode measurement. Each step of the orthogonal sequence is sampled with two redundant photodiodes, and the result is demultiplexed to determine the contribution of each wavelength. The entire sequence is planned to take less than 1 ms. If one or more laser diodes fail, they are removed from the sequence and the scanning order is updated to compensate for the lost laser diodes. The proper function of the lasers and the detectors can easily be verified by a simple test sequence using a calibration target in view of the SEARCH Instrument.

The spatial resolution of the SEARCH laser optics is designed to match the movement of the rover mast (1.0 cm^2 at a distance of 10.0 m from the target). The shiny rock surfaces frequently seen in Viking lander images were postulated to be desert varnish [19] because of their resemblance to this biofilm. Viking's image resolution was about 0.5 cm^2, thus the SEARCH resolution is adequate to detect such features. The 1.0 cm^2 resolution will also be adequate to detect the other targets sought as described above. SEARCH is to be mounted on the rover mast, so as the Rover approaches the targets the resolution will improve in inverse proportion to the laser path length. Thus, at 1.0 m from a target, the resolution will be 1.0 mm^2.

As described in Section 0, SEARCH will perform either stationary or mobile scans to map the surrounding area and search for interesting substances. Sampling each of the photodiodes with its own 16-bit A/D channel, at rates of over 100K samples/sec, SEARCH will process the sample data, producing a classified substance image at the same rate at which new wavelength data for each pixel are sampled. Each pixel in the classified substance image takes 16-bits of storage (8-bit substance identifier + 8-bit distance in SD space).

A stationary scan produces a raw wavelength image, a classified substance image, and a 256 bit summary indicating which substances were detected in the scan. The summary can be reviewed by a human operator or some higher level navigation system on a rover and used for navigational purposes or to correlate data from other instruments on the rover.

If the SEARCH system contains at least 254 MB of storage, then two compressed raw wavelength images and two classified substance images at full resolution with a 360° azimuth range and elevation over a range of 70° can be stored. The corresponding uncompressed substance image requires 10.6 MB of storage. Demultiplexing and classifying one target location using the MLR method requires approximately 150,000 floating-point operations (FLOPs). Thus the processing unit must sustain 185 MFLOPS to process pixels at a rate of 1 target location per millisecond.

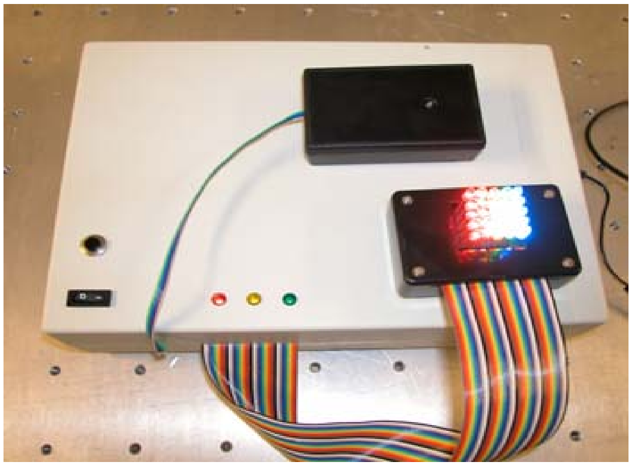


Figure 12 - The SEARCH 5x5 Prototype in its portable casing, designed to enable field testing.

During mobile scan mode, SEARCH can be given a set of substances considered interesting, and a range over which to scan. SEARCH can then scan continuously until requested to stop or until an interesting substance is detected which would trigger a state update notification to the rover's main computer system.

The BEST library, in the SEARCH Instrument, can be configured to contain a set of pre-defined substances. New library entries can be generated off-line (on Earth) based on wavelength data from unclassifiable pixels and data from other instrument in the rover. When operating autonomously, SEARCH could be programmed to group unclassifiable substances with similar spectra, and use the BEST method to create new library entries improving the compression for these new substances. Later, the spectra for these new substances can be transmitted (back to Earth) to assist in further evaluation and identification of the substances.

6. CONCLUSION

SEARCH is a new technique for noncontact classification of geological formations and biofilms. It uses a laser diode array to illuminate a target with a wide range of wavelengths. Reflected light is measured and processed to determine the target's spectral reflectance, and to identify what substances are present at the target. SEARCH can be implemented with robust, well-understood technologies, like commercially available laser diodes, photodetectors, and microprocessors.

SEARCH promises several advantages over currently used alternatives. Unlike thermal emission spectrometry, SEARCH does not rely on ambient light, allowing nighttime scanning when a rover is less utilized, and allowing it to be used to help plan the next day's mission operations. Furthermore, SEARCH has the potential to achieve much higher signal-to-noise ratios than either thermal emission spectrometry or Raman spectrometry, and it does not require lenses, electromechanical focusing mechanisms, or other components that need fine adjustment, reducing the possibility of misalignment due to vibration in transit.

Preliminary experiments support the theory behind SEARCH. Experiments have used SEARCH technology to classify proteins in the lab, and measure biofilms, both in the lab and in outdoor environments. Both living bioflims and residues of dead biofilms have been detected on stone surfaces.

Several generations of the SEARCH prototypes have been developed at the University of Kentucky with the goal of eventually producing a space-qualified instrument. Though the current implementations are not specifically designed to withstand wide temperature ranges, high radiation, or a vacuum, there are no basic principles to prevent such an implementation. For example, currently available space grade DSP processors have enough processing power to handle the BEST classification and drive a 24x5 laser diode array. Redundant laser diodes and sensors will provide the ability to tolerate single-point failures.

The current generation of SEARCH uses a 5x5 visible-light LED array and a complementary photodetector, under the control of a microcontroller linked to a laptop. Results of experiments both inside and outside the lab with this portable prototype are being used to guide the next generation of SEARCH development. The next generation of SEARCH devices is expected to replace the external laptop with a digital signal processor (DSP) chip for classifying substances. It is also planned to use laser diodes to achieve better scanning distance and spatial discrimination than possible with LEDs.

7. ACKNOWLEDGEMENTS

The authors thank Dr. Gil Levin for his advice and consultation on the application of the SEARCH instrument to remote sensing on Mars, and for the acronym. They also thank Clay Harris, David Jackson, and Aaron Urbas for assistance with the instrument and data collection.

REFERENCES

[1] L.A. Cassis and R.A. Lodder, “Near-IR imaging of atheromas in living arterial tissue,” *Analytical Chemistry*, **65,** 9, 1247-56, 1993.

[2] F. Garcia-Pichel, S. L. Johnson, D. Youngkin, and J. Belnap, Small-scale vertical distribution of bacterial biomass and diversity in biological soil crusts from arid lands in the Colorado Plateau, *Microbial Ecol* 46:312-321, 2003.

[3] Harald Furnes, Neil R. Banerjee, Karlis Muehlenbachs, Hubert Staudigel, Maarten de Wit, Early Life Recorded in Archean Pillow Lavas, *Science*, Vol. 304, Issue 5670, 578-581, 23 April 2004.

[4] Gernot Arp, Andreas Reimer, Joachim Reitner, Photosynthesis-Induced Biofilm Calcification and Calcium Concentrations in Phanerozoic Oceans, *Science* Vol. 292, 1701-1704, 1 June 2001.

[5] D.C. Catling, On Earth, as it is on Mars? *Nature* 429:707-708, 17 June 2004.

[6] W. M. Grundy and B. Schmitt, The Temperature-Dependent Near-Infrared Absorption Spectrum of Hexagonal H_2O Ice, *J. Geophys. Res.* 103, 25,809-25,822, 1998.

[7] Stephen M. Clifford and Timothy J. Parker, The Evolution of the Martian Hydrosphere: Implications for the Fate of a Primordial Ocean and the Current State of the Northern Plains. *Icarus* 154, 40-79, 2001.

[8] Opportunity rover finds strong evidence Meridiani Planum was wet, NASA/Jet Propulsion Laboratory press release, March 2, 2004, Available at `http://marsrovers.jpl.nasa.gov/newsroom/pressreleases/20040302a.html`.

[9] Amanda Lowell, Kah-Siew Ho, and Robert A. Lodder, Hyperspectral Imaging of Endolithic Biofilms Using a Robotic Probe, *Contact in Context*:v1i1/planetprobe, 1((1)), 1-10, 2002. `http://www.contactincontext.org/`

[10] F. Garcia-Pichel and R. W. Castenholz, Occurrence of UV-Absorbing, Mycosporine-Like Compounds among Cyanobacterial Isolates and an Estimate of Their Screening Capacity, *Appl. Environ. Microbiol.*, Vol. 59, No. 1, 163-169, Jan 1993.

[11] Gernot Arp, Andreas Reimer, Joachim Reitner, Photosynthesis-Induced Biofilm Calcification and Calcium Concentrations in Phanerozoic Oceans. *Science* Vol. 292, 1701-1704, 1 June 2001.

[12] R.J. Dempsey, D.G. Davis, R.G. Buice, Jr., R.A. Lodder, Biological and Medical Applications of Near-Infrared Spectrometry, *Appl. Spectrosc.*, 50(2), 18A-34A, 1996.

[13] M. A. Chan, *et al.*, A possible terrestrial analogue for hematite concretions on Mars. *Nature* 429(June 17):731-734, 2004.

[14] C. A. Metzger, D. O. Terry, and D. E. Grandstaff, Effect of paleosol formation on rare earth element signatures in fossil bone. *Geology* 32(June):497-500, 2004.

[15] H. E. B. Da Silva, C. Pasquini, Dual-Beam Near-Infrared Hadamard Spectrophotometer *Appl. Spectrosc. 55* (6), 715-721, 2001.

[16] Miniature Thermal Emission Spectrometer (Mini-TES). `http://marsrovers.jpl.nasa.gov/mission/spacecraft_instru_minites.html`

[17] C. R. Webster, R. L. Herman, R. D. May, W. S. Woodward, R. W. Zurek, Y. L. Yung, G. R. Rossman,

Combination Laser Absorption and Raman IR Spectrometer (CLARIS) for in situ Atmospheric, Particulate, & Mineralogical Studies of Mars , `http://laserweb.jpl.nasa.gov/planetaryinstruments/claris.html`

[18] Aaron A. Urbas and Robert A. Lodder, In situ spectroscopic cleaning validation, *NIR News*, 14(2), 8-10, 2003

[19] H. J. Moore, personal communication to G.V. Levin, 1979

[20] R. Heninger, Mars Science Laboratory Mission 2009 Landed Science Payload Proposal Information Package, NASA Jet Propulsion Laboratory, D-27202, MSL-225-0502, April 2004. `http://centauri.larc.nasa.gov/msl/PIP-FINAL-14Apr04.pdf`

BIOGRAPHY

William R. Dieter is an Assistant Professor of Computer and Electrical Engineering at University of Kentucky. His current research interests include power-aware software for embedded systems, real-time systems, and high-performance computing. Dieter received a Ph. D. in Electrical Engineering from University of Kentucky in 2001 and BS in Computer and Electrical Engineering from Purdue University in 1993. He is a member of IEEE, ACM, IEEE Computer Society, and USENIX.

Robert A. Lodder received his B.S. degree cum laude in Natural Science from Xavier University (Cincinnati, Ohio) in 1981. After deciding to pursue a career in chemistry he worked under Professor Richard T. O'Neill at Xavier, and received his M.S. in Chemistry in 1983. He received his Ph. D. in Analytical Chemistry in 1988 after working with Professor Gary M. Hieftje at Indiana University, and is currently Professor of Pharmaceutical Sciences at the College of Pharmacy, University of Kentucky Medical Center. Dr. Lodder holds joint appointments in the Department of Electrical and Computer Engineering, and the Division of Analytical Chemistry of the Department of Chemistry at Kentucky. He serves as Editor of the astroanalytical chemistry and astrobiology journal Contact in Context, and as a member of the U.S. Food and Drug Administration Advisory Committee on Pharmaceutical Science, Process Analytical Technologies subcommittee. Dr. Lodder is a first-prize winner in the 1990 international IBM Supercomputing Competition, as well as a winner of a National Science Foundation Young Investigator Award, the American Society of Agricultural Engineers Paper Award, a Buchi NIR Award, the Tomas Hirschfeld Award in Near-IR Spectroscopy and a Research and Development 100 Award.

James E. Lumpp, Jr. James E. Lumpp, Jr. is an Associate Professor in the Department of Electrical and Computer Engineering at the University of Kentucky. He received the BSEE and MSEE degrees from the School of Electrical Engineering at Purdue University in 1988 and 1989 respectively, and the Ph. D. from the Department of Electrical and Computer Engineering at the University of Iowa in 1993. He joined the faculty at the University of Kentucky in 1993. He is a Member of IEEE, ACM, AIAA, and SWE and the Center for Advanced Networking and the Advanced Science and Technology Commercialization Center at the University of Kentucky. He has worked as a visiting scientist or intern at AT&T Bell Labs, Rockwell International, Northrop Defense Systems Division, and Oak Ridge National Laboratory. His research interests include distributed embedded systems, safety critical systems, and high-performance distributed computing.