Reusable Pattern Recognizers for Solar Imagery
Continuing Proposal to 632-07

Task Leader

Michael Turmon, JPL, turmon@jpl.nasa.gov, 818 393 5370, PI

Product Description

This task will develop image analysis technologies which automatically recognize scientifically relevant features in multispectral images, with emphasis on solar imagery. System operation begins by using scientists' labelings automatically to learn a statistical model for the features of interest. As new images are acquired, this model is coupled to general-purpose image-analysis software which integrates the resulting spatial cues to define local, pixel-level assessments of activity. These pixel-level cues are then progressively linked into a coherent, object-level scene description of the phenomenon of concern. The result is a directly-interpretable *summary of scientifically relevant information* in the image. Our use of open standards for structured data (e.g., object models and datasets in XML with suitable schemata) ensures portability across applications and future extensibility.

	Current	***FY 1999***	***FY 2000***	***FY 2001***
Capability	Hand-label and annotate images	Automated data fusion & labeling	Link recognized objects in time	Link into phys. relevant groups
Metric	45 min/image (trained expert)	3 min/image (software tech.)	+ 5 min/image	+ 3 min/image
Comment	neither repeatable nor objective	objective, repeatable	no existing automatic system for this task	no existing automatic system

This is a **continuing**, **push** (60%) / **pull** task.

Benefits

This technology will increase the science return and cost-effectiveness of space missions. Development of portable models for data allows exchange and re-use of pattern recognizers across missions, enhancing model quality and amortizing costs over a larger mission base. Higher-level scene descriptions, coupled with scientists' ranking of the importance of features, allow principled prioritization for image acquisition and downlink as well as enabling intelligent data compression based on identified features.

For more on the applicability of this work, see the Customer Relevance section.

Technical Approach

Many problems of spacecraft autonomy and high-level data analysis can be couched in terms of finding structures in images. The work described here is aimed at a trainable, hierarchical approach to this problem, in which scene descriptions are inferred from pixel-by-pixel data upwards toward more useful object-level representations. Placement of the inference problem in a statistical framework allows determination of model parameters from prior information and scientist-provided data. To ground these ideas, we

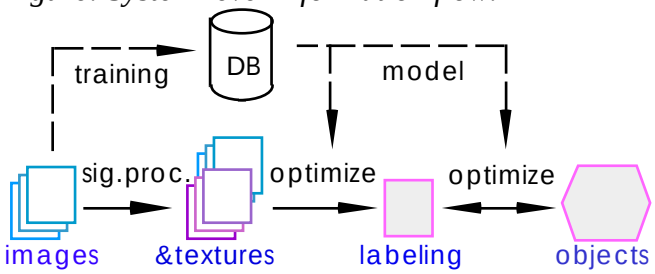
have used a family of problems in the area of solar physics as a testbed. The central tasks here are to find, track, and characterize the shapes of active regions (e.g., sunspots) in full-disk solar images. To do this, we have used two levels of representational abstraction: a pixel-level labeling of the scene, and, at a higher level, a 'membership function', defined by just a few parameters, giving the subjective degree-of-belief that a given site is part of an active region.

Tackling such problems provides an arena for developing a new toolbox of statistical models for multispectral scientific images, including optimization technology for fitting models to data and database technology for efficient maintenance of training data and the models themselves. The first question is addressed by developing numerical methods for training flexible models by large-scale optimization, efficiently sampling from the model probability distributions, and applying pattern recognition methods to large-scale spatio-temporal data sets. The latter issue is attacked by developing model-databases to store and access trainable statistical models of observed data, the training data, and training experiments combining models and data. Such databases allow appropriate models to be selected and validated from a given general family.

Tech-I: Introduction and Foundation

The research program outlined below finds structures in images by defining representations for the information to be extracted from the data, building statistical models for the representations and their relationship to the data, and learning the detailed characteristics of the model from a database of labeled imagery. Once the representation and the precise form of the model are known, optimization and pattern recognition methods are used to find the best representation for a new data set.

Figure: System-level information flow.



The figure shows this process schematically. The principal data path starts with an image dataset, possibly consisting of images in several modalities and resolutions. A preliminary image feature extraction phase augments this dataset with derived images such as rotation-invariant texture or principal component vector images. Then an optimization phase adapts a multilevel representation to the derived image features. This representation can be stored in a model database for later recall, evaluation, and comparison with other training sessions. The whole process is regulated by expert-supplied labelled images, which are also stored in the database. These labelings are used in the beginning to determine the model parameters which link the observations to the derived representation. The statistical framework in which the model is defined makes this training phase possible.

The final feature recognition system can be used in ground-based data analysis, in which case the final, optimal representations are to be stored in a database. Alternatively, a trained version of the recognizer — just the bottom portion of the figure — can be in operation on the spacecraft itself. The information gained is used to direct exploration, guide data acquisition, and maximize utility of data uplink.

One quantitative formulation of this processing mechanism is as follows. Image data $\mathbf{x}$, indexed by location $s = (s_1, s_2)$ and modality m are observed. From this data, we wish to determine an image decomposition $\mathbf{y}$ for all sites $s \in N$, say, for some region of interest. We furthermore wish to determine a higher-level representation $\mathbf{h}$ which has characteristics that will influence the image decomposition (e.g., Turmon and Mukhtar 1997). Formally, we assume there is a Markov relationship between the levels of the model, so that the joint probability of a data-set $\mathbf{x}$ and its interpretation $(\mathbf{y}, \mathbf{h})$ is

$$P(\mathbf{x}, \mathbf{y}, \mathbf{h}) = P(\mathbf{x} \mid \mathbf{y}) \;\; P(\mathbf{y} \mid \mathbf{h}) \;\; P(\mathbf{h}) \;\; .$$

We wish to infer the interpretation given the data; accordingly, we may maximize the posterior distribution $P(\mathbf{y}, \mathbf{h} \mid \mathbf{x}) = P(\mathbf{x}, \mathbf{y}, \mathbf{h}) / P(\mathbf{x})$, or, equivalently, its logarithm:

$$(\mathbf{y}, \mathbf{h}) \leftarrow \max_{(\mathbf{y}, \mathbf{h})} \log P(\mathbf{x} \mid \mathbf{y}) + \log P(\mathbf{y} \mid \mathbf{h}) + \log P(\mathbf{h})$$

where we have dropped the term $P(\mathbf{x})$, which is constant with respect to $\mathbf{y}$ and $\mathbf{h}$ and does not affect the maximization. The interpretation of the above is simple: maximize fidelity to the data (the first term) subject to a constraint on the reasonableness of the interpretation $(\mathbf{y}, \mathbf{h})$. Typically, this reasonableness condition takes the form of the spatial smoothness of $\mathbf{y}$ as a function of the location variable s, perhaps also coupled with an "Occam's razor" constraint on the complexity of the interpretation $\mathbf{h}$.

It is in the maximization of the above expression that the optimization sub-goal of this task comes to the fore. It is not unusual for the first-level interpretation $\mathbf{y}$ to consist of thousands or millions of free parameters, which must be estimated from the millions of per-pixel observations. The next level of abstraction might consist of only hundreds of free parameters, but these typically have a complex relation to the original data. Often, simulated annealing (Geman and Geman, 1984) or Metropolis-Hastings methods (Besag et al, 1995) are used to perform this maximization.

As mentioned above, the probability distributions defined here have free parameters to allow them to adapt to different data sources. In a problem with two object classes, for example, $P(\mathbf{x}|\mathbf{y})$ can be determined by standard methods if examples of $\mathbf{x}$ from both classes are provided by domain experts. Within this statistical context, a principled approach to model selection is possible. We note that this probabilistic model is the way we *naturally account for multispectral inputs:* in this context, the per-pixel observable $\mathbf{x}$ becomes a vector, but this does not change the underlying probabilistic machinery.

The resulting object abstraction becomes the tool for linking observations across time. This not only reduces uncertainty in active region locations, it is fundamental to studies of longer-term sunspot evolution. To do this, the membership function is indexed in time and in space. There is a deterministic component of motion, as well as a stochastic

component to allow for variations among the population of objects. The membership function is then adjusted to account for past position as well as the current data.

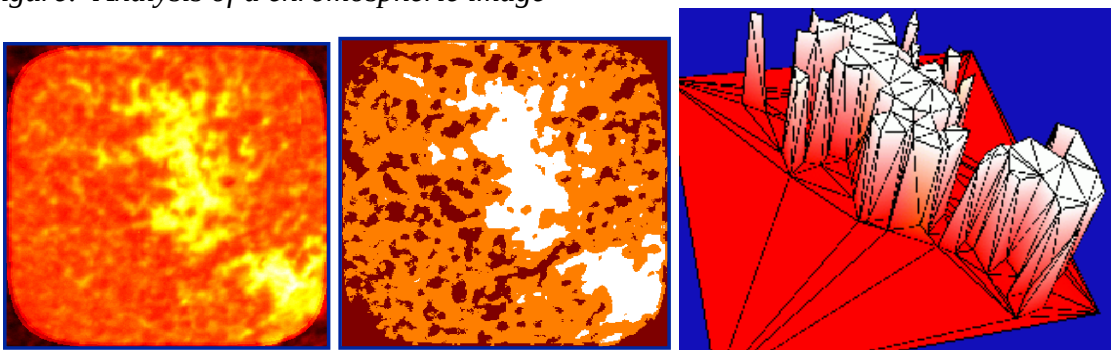
Tech-II: Application to Solar Feature Identification

This segment of the proposed research will identify and characterize the various surface structures of the Sun which contribute to observed changes in solar total and spectral irradiances. Various space-borne observations of total solar irradiance over the last two decades have established conclusively that total solar irradiance is not constant; rather, it changes on time scales from minutes to the 11-year solar activity cycle (e.g. Willson and Hudson, 1988) and may be responsible for slow climatic changes such as those that produced the Little Ice Ages (Hansen et al., 1993). To understand the physical causes of irradiance changes, it is necessary to study the spatial and temporal characteristics of structures in the various layers of the solar atmosphere. For this purpose, during FY99, we have used the high-precision images of the Sun taken by the Michelson Doppler Imager (MDI) aboard the SoHO (Solar and Helioseismological Observatory) satellite (Scherrer et al., 1995).

This research requires processing of thousands of images, for which we have established a state-of-the-art statistical pattern recognition system, building on related methods developed for chromospheric intensity maps (Turmon and Pap, 1997). In contrast to expert labeling on the one hand, and elementary pixel-level image processing on the other, we propose to use the hierarchical image segmentation technique, as outlined above, which will provide more objective and accurate measurements of the area, position, intensity, and contrast of sunspots, faculae, and the magnetic network. The proposed system will greatly enhance the objectivity, repeatability, and conciseness of current state-of-the-art methods, which generally are based on hand-labeling or heuristic region-growing (e.g., Worden 1996).

As indicated, this problem fits naturally into the framework of section I. The observed data (in this application, a magnetogram and photogram from the MDI CCD array) is $\mathbf{x}$, and the object type at each pixel is $\mathbf{y}$. In this analysis, we anticipate using four classes: quiet sun, faculae, active regions (sunspots), and unidentified outliers. A high-level description of the sunspots can be determined and becomes $\mathbf{h}$, which will be used for spot tracking, age determination, and shape evolution. A sample of the use of this setup for a comparable dataset (ground-based chromospheric images, see Turmon and Mukhtar, 1998) is shown below. The left panel is a detail of a chromospheric image; its labeling is in the center panel, and the higher-level description of the active region is at right: it uses only about 200 real numbers to achieve a coarse-grained representation. The statistical approach advocated above allows scientist input in developing the models for $P(\mathbf{x} \mid \mathbf{y})$. A graphical user interface has been integrated into the standard SAOtng image browser, which this team now co-develops, to help scientists perform this model-selection step automatically.

Figure: Analysis of a chromospheric image



Tech-III: Reusable Frameworks for Feature Recognition

While recognizing the importance of applications to motivate and guide research, it is essential that the software system we develop be open, flexible, and reusable in other application areas both in solar physics applications in the SEC enterprise and in other enterprises. A major thrust of this proposal is to ensure that the developed software base is reusable in these applications. This is done via mechanisms of software and data abstraction.

Several kinds of data, having a diverse structure, must be maintained in the pattern recognition environment advocated here: image sets, labeled images, calibration parameters, and various statistical models. The key to obtaining reusability is to adopt an expressive language for these data; we have used the emerging XML (extensible markup language) standard. This is a tuple-based textual language which is capable of encoding structures as general as graphs, and having extensions for external data (XLink) and mathematical formulas (MathML). It is reconfigurable for various applications via a sort of schema (DTD, or document type definition). Since XML parsers exist for many languages (C, Java, and Python, among others), it is easy to make an XML encoding interoperable with other formats, and to build an interface to allow persistent objects via interchange with a database.

In FY99, for example, we have developed a flexible schema for encoding generative statistical models in XML using the well-known Bayes net formalism (e.g., Pearl 1988). The resulting statistical models are used to compute the probabilities needed in the optimization strategies of section I; they may also produce samples from a generative model to check its validity . The figure shows a mixture model for sunspot data and its XML representation as shown in a browser/editor. These model files may be exchanged and compared, allowing scientists to build on the calibration and modeling efforts of the community. This has already found application in the CNES Picard mission, which can use parts of the active region models developed for SoHO/MDI images to find and screen out active region effects from its delicate measurement of solar diameter.

The other major capability of the system we advocate is abstraction of the algorithms used. To do this, we may leverage the portable language for data and algorithmic parameters referred to above. A typical algorithm is given input in this language of its parameters plus an image-set and an annotation of image-features, and produces an output in the same form. This homogeneity allows an object-oriented calling mechanism driven by an extensible table of recognized image-types and operations. A generic image operation consists of a series of such calls, and is invoked by sending the manager the operation name and the image-set to operate on. The execution manager takes care of parameter query, operation sequencing, and metadata maintenance. We have begun to prototype this mechanism in the high-level object-oriented scripting language Python, which has been used recently to integrate large-scale scientific codes. The basic image operations, implemented in a lower-level language, are combined by the object-oriented execution manager. This approach allows easy extensibility via derived classes.

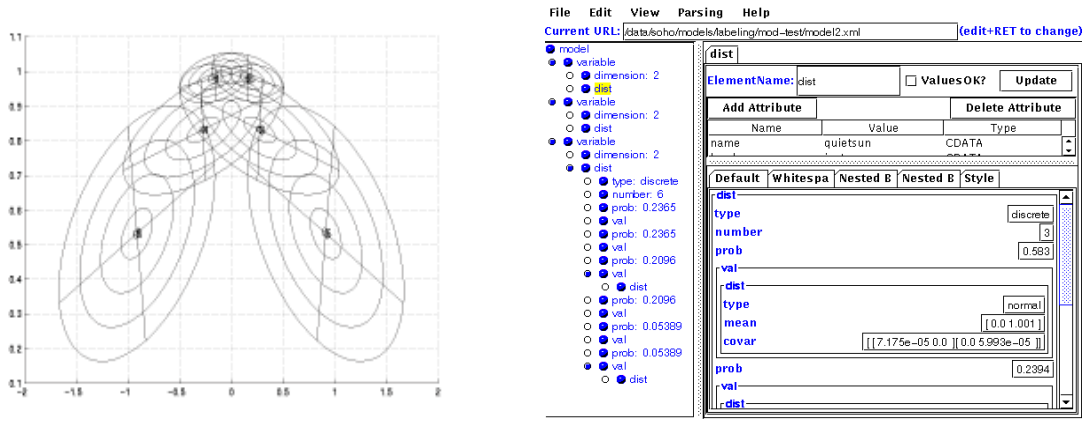


Figure: Model for sunspots in SoHO/MDI observables, and XML rendering

Tech-IV: Conclusion

A broadly applicable framework has been presented for the problem of automatically and objectively finding and classifying objects present in scientific imagery. The essential ingredient is a testable statistical model relating the data to the objects of interest. To support the development of such models, we have introduced database technology for storing scientist-labeled data and complete model specifications, allowing practicing scientists to develop a toolkit of valid statistical models for the data they work with. Novel optimization technology then supports the fitting of object representations to the observed data. Furthermore, having specified a model, this fitting process can execute automatically, allowing recognition to be done in situ which gives an objective basis for optimally use of limited observation time and downlink capability.

Status and Milestones

Activity in FY99 proceeded on several levels. Programmatically, support for software development and scientific research on the connections between active region occurrence and solar irradiance on SoHO was secured through a three-year award from NASA SEC. The PI and his collaborator Judit Pap held a two-day meeting with the SoHO imaging team at Stanford to discuss calibration issues and to further the goal of integrating our labeling system into the MDI data pipeline. A follow-up meeting to discuss related modeling issues will be held in Los Angeles in July, attended by the PI and solar physicists from Stanford, UCLA, Big Bear Observatory, and CSUN San Fernando

Observatory. The PI was made a co-investigator on the CNES Picard satellite, due for launch in 2003. The active region location techniques will be used to identify and screen out sunspots from a delicate radius determination. The PI's software package is also being installed at Big Bear Solar Observatory for labeling their images. On the general-purpose algorithm development side, the multilevel modeling technique developed by the PI won a Best Paper award at the 1998 Computational Statistics meeting. The work on portable models for scientific data will be presented at the 1999 Conference on Highly Structured Stochastic Systems. Scientific results on the relations between activity and irradiance were presented at the SoHO-6 meeting, the AGU, and the Birmingham (U.K.) irradiance meeting. A paper describing the results is being prepared for the Astrophysical Journal.

FY1999, Q4 — Product will fully support multimode image inputs, including ability to define reusable models for multimodal inupts.

FY2000, Q4 — Include temporal contextual information in region identification. Define and use portable datatypes for training and region-identification sessions.

FY2001, Q4 — Product will allow diverse object descriptions, e.g. multi-object clusters such as sunspot groups.

Customer Relevance

See also the Benefits section.

Co-Funding

In FY1999, co-funding agreement with SEC for adaptation of this active-region identification software for the SoHO/MDI dataset was initiated at a level of about $50K/year over three years. In the mid-term, the PI of the CNES PICARD satellite has selected this system for active region identification. These agreements indicate the strong interest of the solar physics community in the proposed methods. As the product generality and TRL grow, co-funding agreements with STEREO mission (e.g., Dr. Jeff Kuhn) will be defined.

Technology Endorsements

Judit Pap, Senior Research Scientist, UCLA Dept. of Physics and Astronomy
Philip Scherrer, SoHO/MDI PI and Professor, Stanford Dept. of Physics
Richard Bogart, MDI Data Scientist, Stanford Ctr. for Space Science and Astrophysics
Claus Frölich, SoHO/Virgo PI and Director, World Radiation Center, Davos, Switzerland
Luc Damé, Picard PI, Centre Nationale de Recherche Scientifique, Paris
Jeff Kuhn, APT Instrument Developer and Astronomer, U. Hawaii Institute for Astronomy
William Marquette, Director, Big Bear Solar Observatory

Vita

Dr. Michael Turmon is a senior member of the technical staff in the Machine Learning Systems Group at the Jet Propulsion Laboratory. He received Bachelor's degrees in Computer Science and in Electrical Engineering from Washington University in St.Louis. He obtained his Ph.D. in Electrical Engineering from Cornell University in 1995. He was a National Science Foundation Graduate Fellow, and was invited to apply for an AT&T Graduate Fellowship. Michael's areas of research are in the theory of generalization in neural networks, and applications of model-based and nonparametric statistics to image

understanding. Michael has been theory co-chair on the program committee of the annual NIPS (Neural Information Processing Systems) conference, chaired a session at the IMS (Institute for Mathematical Statistics) meeting, and is co-investigator on the NASA/ESA SoHO spacecraft. He has reviewed for the learning theory sessions of NIPS, for the Knowledge Discovery in Databases conference, and the Computer Vision and Pattern Recognition conference. He has been an invited participant in NIPS and AMS/SIAM/IMS workshops on learning and generalization. Michael has refereed for the IEEE Transactions on Signal Processing, the Journal of the American Statistical Society, and the Machine Learning Journal. He is a member of the IEEE (Information Theory Society) and the IMS.

Technical References

S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Patt. Anal. and Mach. Intell.*, **6**:721-741, 1984.

U. Grenander and M. I. Miller, "Representation of knowledge in complex systems," *J. Roy. Stat. Soc.* (Ser. B), p. 549-582, 1994.

P. J. Green, "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination," *Biometrika*, 82(4):711-732, 1995.

J. Hansen, A. Lacis, R. Ruedy, M. Sato, and H. Wilson, *Natl. Geogr. Explor.*, 1993, **9**, 142.

J. Pap, M. Turmon, and S. Mukhtar, "Automatically finding solar active regions using SoHO/MDI photograms and magnetograms," *Proc. SoHO 6/GONG '98 Workshop on Structure and Dynamics of the Sun*, 1998, pp. 979-984.

J. Pap, M. Turmon, S. Mukhtar, R. Bogart, R. Ulrich, C. Frolich and Ch. Wehrli, "Automated recognition and characterization of solar active regions based on the SoHO/MDI images," In *Proc. 31st ESLAB Symposium*, 1997.

J. Pearl, *Probabilistic reasoning in intelligent systems*, Morgan-Kaufman, 1988.

P. Scherrer et al., "The Solar Oscillations Investigation – Michelson Doppler Imager," *Solar Physics*, 1995, pp. 129-188.

M. Turmon and S. Mukhtar, "Recognizing chromospheric objects via Markov Chain Monte Carlo," *Proc. IEEE Int'l Conf. Image Process.*, 1997, pp. 320-323.

M. Turmon and S. Mukhtar, "Representing solar active regions with triangulations," *Proc. Compstat-1998*, pp. 473-478.

M. Turmon, E. Mjolsness, V. Gluzman, and L. Ramsey, "A language for probabilistic modeling of scientific data," In *Proc. Second Conf. Highly Structured Stochastic Systems*, 1999.

R. C. Willson and H. S. Hudson, 1988, *Nature*, **332**, p. 810.

J. Worden, *A Three Component Proxy Model for the Solar Far Ultraviolet Irradiance*. PhD thesis, Univ. Colorado, Boulder, CO, 1996.

From: Judit Pap <pap@astro.UCLA.edu>
To: mjt@aig.jpl.nasa.gov
Date: Fri, 11 Jun 1999 19:09:11 -0700 (PDT)

To: Autonomy and Information Management Team (632-07)
Subject: Pattern Recognition in Scientific Imagery Task
Reference: Letter of Support

The purpose of this letter is to express my support for Dr. Michael Turmon's work on feature recognition in solar images. We have collaborated for over three years on various aspects of this problem. During that time, Michael has proven to be an intelligent, hard-working researcher who has supervised the construction of a large piece of software for this scientific problem.

Algorithms to automatically recover solar features from high-resolution images represent an important scientific tool in relating how features on the Sun influence the radiation incident on Earth. Knowing more about this relationship is one way to understand how energy is stored and transmitted by the Sun.

An important part of this effort in the Sun-Earth Connection theme is understanding the links between SOHO irradiance data as measured by the VIRGO experiment, and the spatially-resolved active region behavior as captured by MDI. The software Michael developed is a promising method of making this link; in fact, due to the volume of data involved, such automated methods are the only feasible way for our community to proceed in this task. For instance, recent analysis of a fourteen-month trial period involved processing over 2000, two-megabyte images. This work has led to the active interest of the SoHO/MDI PI, Phil Scherrer, the MDI data scientist, Richard Bogart, and the SOHO/Virgo PI, Claus Frolich.

Furthermore, the methods developed are founded on widely-accepted statistical methodology, which should contribute to the acceptance of these results in the solar physics community. Besides the work indicated above, obvious avenues of approach exist in using these methods on similar solar datasets taken from other instruments on SOHO, and for example, the telescopes planned for launch on the STEREO project and various planned NASA experiments after STEREO.

Additionally, the research led by Dr. Turmon at JPL will lead to a better understanding of the physical origin of the observed irradiance changes. It must be underscored that the space-borne irradiance measurements over the last two decades have established conclusively that the Earth's climate, radiative environment, and upper atmospheric chemistry are influenced by the varying solar energy flux. Therefore, understanding the underlying mechanisms of irradiance variability goes well beyond the interest of our scientific peers.

Dr. Judit Pap
Research Astronomer
UCLA Department of Physics and Astronomy

```
From: Jeff Kuhn <kuhn@pelea.ifa.hawaii.edu>
To: Michael.J.Turmon@jpl.nasa.gov
Subject: Pattern Recognition in Scientific Imagery
Date: Wed, 16 Jun 1999 10:36:51 -1000 (HST)
Cc: Paul.E.Stolorz@jpl.nasa.gov, kuhn@ifa.hawaii.edu
```

Dear Mike,

I want to underscore the importance of the pattern recognition work you are doing.

There are several points that deserve wider recognition:

1) While it is now clear that the sun is a major contributor to climate change, it is not at all evident what the physics of this interaction is. Our research here is at an exploratory stage and we need clear analytic tools to make sense of the photospheric signatures of global
solar change.

2) The well-founded analytic techniques you are using and developing are unique and not reproduced in any other work. I think this is exactly the correct approach needed for making a quantitative advance on questions like how the solar convection zone gates the entropy flux from the deep solar interior. These are important questions that should find interest well beyond the solar physics (or statistical analysis) communities.

I hope that the feature/region-identification techniques will continue to be developed and applied to these important problems.

Sincerely,
Jeff Kuhn
Professor, Institute for Astronomy, and
University of Hawaii

From: Rick Bogart <rick@rick.stanford.edu>
To: Michael Turmon <turmon@aig.jpl.nasa.gov>
Subject: Letter of Support
Date: Mon, 21 Jun 1999 11:45:08 -0700
Cc: Paul.Stolorz@jpl.nasa.gov, Pscherrer@rick.stanford.edu

To: Autonomy and Information Management Team (632-07)
Subject: Pattern Recognition in Scientific Imagery Task
Reference: Letter of Support

I am writing to you at the request of Dr. Michael Turmon of the Machine Learning Group at JPL to express my interest in and support for the work on feature recognition in solar images being conducted as part of the Statistical Pattern Recognition in Scientific Imagery task under his leadership.

Algorithms to automatically identify and categorize solar features from high-resolution images are an important scientific tool in several critical areas of solar research: understanding the contributions of the various features of solar magnetic activity to the overall photospheric flux budget, with its implications for understanding the influence of variations in the solar radiation incident on Earth for climate change; tracking the surface motions to better understand the detailed dynamics of the upper boundary of the convection zone; and analyzing the growth, decay, and morphology of magnetic features in order to better elucidate their structure and relation to the overall magnetic cycle and to understand how energy is stored and redistributed within the sun's convective envelope.

The Michelson Doppler Imager (MDI) on SoHO has amassed over the last 3 years a huge database of continuum imagery useful for these purposes whose scientific potential has barely been tapped. Part of this dataset is the sets of about 20 images per day (photograms and magnetograms) that have been used by Dr. Turmon's group to derive several daily feature maps extending over the mission lifetime to date. The automated methods developed for that purpose are the only feasible way to proceed with the task of identifying region types. The maps of feature classes on the solar disk produced by this software are of great interest to the team of solar physicists who are studying the irradiance aspect of the Sun-Earth connection. They are also of potential interest to the teams studying the nature of the features themselves, and the underlying techniques promise to be of great use in the analysis of the vastly greater amount of MDI data accumulated each minute for months at a time to explore the dynamics of the convection zone. Efforts are underway to agree on a taxonomy of physically relevant solar structures so that such maps can be produced as one of the standard data products of the Solar Oscillations Investigation running MDI, with braod applicability in several areas of research.

Dr. Rick Bogart
SOI/MDI Data Scientist

To: Michael Turmon <turmon@aig.jpl.nasa.gov>
From: Luc Damé <luc.dame@aerov.jussieu.fr>
Subject: Your participation in PICARD is essential
Date: Mon, 14 Jun 1999 22:16:09 +0200

To: Autonomy and Information Management Team (632-07)
Subject: Pattern Recognition in Scientific Imagery Task
Reference: Letter of Support

The purpose of this letter is to express my interest in the work on feature recognition in solar images being done as part of the Statistical Pattern Recognition in Solar Imagery task in the Machine Learning Group at JPL under the leadership of Dr. Michael Turmon.

Algorithms to automatically recover solar features from high-resolution images are an important scientific tool in many investigations of the Sun's effects on the Earth. Extending far beyond the obvious relations of these features to irradiance changes, identifying the features is a crucial enabling technology for many solar science missions.

The CNES microsatellite PICARD, which is designed to measure small oscillations in the solar limb, is such a mission. In order to precisely determine the location of the solar limb, the active regions must be identified and removed. The well-calibrated techniques developed by Dr. Turmon are a key step in this process because of the extreme sensitivity of the limb measurement. Furthermore, since these techniques are founded on a mathematical model, they are inherently repeatable. Because of the historical import of all climate-related data, the repeatability is a highly desirable feature: future researchers will be able to compare their data to the PICARD limb measurements by using these objectively-defined mathematical procedures rather than a specially-built set of computer programs.

Finally, I note that the explicit modeling of the activity patterns allows new missions to build on the experience of past missions. For example, the modeling procedures to be used on the PICARD data were originally developed for SoHO/MDI. We can anticipate that the PICARD models will in turn be useful to other solar missions.

Dr. Luc Dame
PI of PICARD

Luc Dame
Service d'Aeronomie du CNRS
BP 3, 91371 VERRIERES-LE-BUISSON CEDEX, FRANCE
Tel: +33 1 64 47 43 28
Fax: +33 1 69 20 29 99
e-mail: luc.dame@aerov.jussieu.fr
ldame@solar.stanford.edu