Innovative Information Technology for Space Weather Research

February 3, 2003

Proposal Summary

Solar activity is closely related to the near earth environment – summarized descriptively as space weather. Changes in space weather have adverse effect on many aspects of life and systems on earth and in space, such as safety of spacecrafts and astronauts, interruption of communication and navigation systems, damage to power grids and trans-continental pipelines due to grounding electric currents caused by geomagnetic storms, and hazardous conditions to polar routed flights. Real-time, high-quality data and data processing would be a key element to forecast space weather promptly and accurately. We propose to use the innovative computation and information technologies (IT) for real time space weather monitoring and forecasting.

(1) We will use the artificial intelligence for pattern recognition, such as segmentation, Support Vector Machines (SVMs), and neural network to detect and characterize three important solar activities in real-time: filament eruptions, flares, and emerging flux regions (EFRs). Such detection methods will be initially applied to data obtained by our 5-station full-disk Global Hα Network which monitors solar activity continuously with a cadence of 1 minute and a pixel resolution of 1 arcsec. Full disk magnetograms and other ground-based and space data will be added as real time access to these data will be improved. Combining the real time detection with the recent statistical study on the relationship among filament eruptions, flares, EFRs, coronal mass ejections (CMEs), and geomagnetic storms, we will establish real time report of solar events and automatic forecasting of earth directed CMEs and subsequent geomagnetic storms.

(2) We combine state-of-art parallel computing techniques with phase diversity speckle imaging techniques, to yield near real-time diffraction limited images with a cadence of approximately 10 sec. For the 65cm telescope at Big Bear Solar Observatory (BBSO), the diffraction limit is about $0.15''$ at 5000Å. The parallel computer cluster will include 64 high speed Pentium processors. We will utilize the multiplicity of parallel paradigms to optimize the calculation of phase diverse speckle imaging to improve calculation speed. With such data, we will monitor flare producing active regions continuously and carry out targeted studies of the evolution and flows in flare producing active regions. Observations will be made in the photospheric continuum, as well as in the G-band (around 4300Å). High resolution magetograms will be obtained by speckle inversion. Our primary scientific objectives is to understand energy storage and release processes in extremely small scales. This includes: (1) understanding the complicated structure and flows in flux emergence, and their role in triggering flares; (2)detecting and characterizing the possible motion of magnetic footpoints in regions involved in different scales of energy release. The real time monitoring will be integrated into our space weather forecasting system.

The real-time phase diversity speckle imaging will be complement and supplement to the Adaptive Optics (AO) systems currently being developed at BBSO and National Solar Observatory (NSO), under NSF support. The hardware and data processing software of the parallel cluster will have direct application to the 4-meter Advanced Technology Solar Telescope (ATST) and the planned 1.6-m New Solar Telescope (NST) at BBSO.

(3) We will develop Web based software tools to post our processed data, events and forecasting in real time, and to be integrated with current solar activity and space weather prediction Web pages at BBSO. This will also be a part of Virtual Solar Observatory (VSO) being developed by the solar physics community. We will also develop a high speed server and advanced user interface and data visualization, so the real-time data and forecasting can be available to the community efficiently.

Finally, let us address two criteria required for an NSF proposal:

(1) Intellectual merit: Inplementing IT in space weather research will be a truly multi-discipline project involving faculty members in solar physics, imaging technology and parallel computing technology from 4 institutions in 3 countries. Research team members are well known in their respective research fields. Space weather, artificial intelligence and parallel computing are three areas standing in the frontier of science and technology.

(2) Broad Impact: the proposed work will effectively integrate research and education. Three Ph.D. students will recieve training and finish their Ph.D. theses in this project. We also anticipate

that 5 to 8 undergraduate students to be involved in the project. Several new courses will be developed during the tenure of the project, and annual workshop will be held involving students and researchers. The technology to be developed, will not only be useful for solar physics, but also will have impacts on medical and military applications. All our developed codes will be freely available to the community. Of course, as we stated earlier, space weather forecasting is having huge impact on life and systems on earth and space under this era of modern technology.

A. Background

The sun is the source of space weather – the disturbances whose radiative, field and particle energy directly impacts earth. Solar activity changes the radiative and particle output of the sun, producing corresponding changes in the near-earth space environment, as well as at the earth's surface. The most dramatic events on the sun, insofar as space weather effects are concerned, are solar flares and coronal mass ejections (CMEs).

A flare is defined as a sudden, rapid, and intense variation in brightness. A solar flare occurs when magnetic energy that has built up in the solar atmosphere is suddenly released. Radiation is emitted across virtually the entire electromagnetic spectrum, from radio waves at the long wavelength end, through optical emission to X-rays and gamma rays at the short wavelength end. The amount of energy released is the equivalent of millions of 100-megaton hydrogen bombs exploding at the same time! Flares are often associated with CMEs.

A CME is due to a large scale rearrangement of the solar magnetic field which can induce a geoeffective solar wind structure. Each CME may carry away a mass of up to 10^{13} kg and release up to 10^{25} J of energy from coronal magnetic fields (Harrison, 1994, 1995). Following their discovery, CMEs were soon found to be correlated with the occurrence of geomagnetic storms. CMEs are easily observed at the solar limb, where they are seen against a dark background. However, the earth-directed, disk CMEs are much harder to detect – although they are the ones that have the most important geomagnetic effects. Therefore, detecting the early manifestation of CMEs on the solar disk is essential to understanding earth-directed CMEs themselves. Establishing the correlation among the occurrence of CMEs and some other solar surface phenomena, such as filament eruptions and flares, is a long-standing problem in solar physics. After firm relationships are established, the correlated solar surface phenomena can provide an early warning of the occurrence of earth-directed CMEs and geo-effectivness of CMEs. A key challenge to move this research area forward is to have real-time, automatic and accurate detection and characterization of solar events, and to use the results for real time space weather forecasts.

Space weather forecast faces another challenge: obtaining and utilizing real-time diffraction limited images. An important goal in ground-based astronomy is to improve the angular resolution that can be achieved. The angular resolution is nearly always limited by phase aberrations introduced by atmospheric turbulence. For the special case of ground-based solar astronomy, the spatial resolution is typically limited to about $0.5''$ for short-exposure images ($\leq$20 ms) and to about $1.0''$ for long-exposure images (about 1 s). $1''$ is about 730 km on the sun. Many basic processes on the Sun, however, take place at scales below $1.0''$. The photon mean free path in the lower photosphere corresponds to about $0.1''$ at disk center. Magnetic structures may occur on even smaller scales. For example, magnetic elements outside of sunspots have typical diameters smaller than $0.2''$ (Keller, 1995). Despite their small size, these small structures are believed to play an important role in large-scale phenomena such as the solar magnetic dynamo or triggering of solar flares and CMEs. To understand a variety of solar phenomena, it is, therefore, indispensable to reach a spatial resolution well below the seeing limit and possibly approaching the diffraction limit of existing and future large solar telescopes. A number of sophisticated techniques have been conceived to combat the deleterious effects of atmospheric turbulence in astronomical imaging in general. Among these are speckle imaging, phase-diversity and phase-diverse speckle imaging. Phase-diversity speckle imaging techniques are particularly attractive for solar astronomy because (1) they require relatively simple and inexpensive instrumentation, (2) they perform well with relatively few images in high-signal regimes, (3) they lead to a joint estimation of the object and the wavefront, and (4) they obviate the need for complicated calibration.

Speckle imaging and phase diverse speckle imaging apply complex numerical algorithms for data analysis. Usually, only a few data sets are analyzed per year, and only at the end of this process does one discover whether something really interesting has been captured. This is due to the long computation time, so raw images have to be saved to be processed off-line, and a hard disk is filled up very quickly. The time-lag between the observation and the data analysis is far too long for any sort of rapid response, such as would be needed for space weather warnings and flare forecasting. Furthermore, the time lag renders the whole scientific enterprise less efficient than it needs to be considering today's computer technologies. The power of parallel computing has just started to be exploited in solar physics. Supercomputers have been used for numerical calculations in astrophysics, however they cannot be used for real-time data processing because the observational data cannot be transferred to the supercomputer centers in real-time.

Parallel processing of solar data will literally provide a new window through which we can observe the Sun in exquisite detail and study the evolution of granulation and sunspots in the active regions leading to major flares. The underlying data processing algorithms are understood, but the complexity is such that only parallel computing enables us to visualize and interpret large data sets effectively. This is an excellent chance for scientists and engineers to do cutting-edge, interdisciplinary work. The development of multithreaded numerical algorithms for high spatial resolution imaging becomes even more important in the context of proposed, next generation, 4 meter-class solar telescopes (Advanced Technology Solar Telescope). All the work proposed here will be beneficial in exploiting the capabilities of such an advanced telescope. Scalable numerical algorithms and computer architectures ensure that we will be in step with the demands for higher spatial, spectral, and temporal resolution and larger fields-of-view observations. In addition, Big Bear Solar Observatory (BBSO) is planning an intermediate size telescope: 1.6-meter New Solar Telescope, which will be operational in 2005/2006, a number of years before ATST. Therefore, we will utilize the proposed techniques and tools in the near future.

B. Automatic Detection and Real-Time Monitoring of Solar Activity

B.1 Introduction and Current Status

In the last few years, we have gained significant experience in the detection of sudden disappearance of solar filaments. This was the center part of the Ph.D. thesis of J. Gao (2002). He combined thresholding and region growing methods which were based on segmentation principle (Beveridge, 1989) in the detection. A sample result of the detection is shown in Figure 1, where the raw data are displayed in the top row and detected filaments are in the bottom row. The positions and sizes of all the filaments are generated automatically. A large filament labeled “A” in Figure 1, disappeared on May 21, 1999, and was detected by the program. In this figure, one can see, by detailed inspection, that the disappearance/appearance of even small filaments are detected, demonstrating the sensitivity of the program. This program is running daily at BBSO, providing real time detection of filaments and statistical properties of filaments in various sizes.

B.2 Objectives

We propose to advance our studies in the following two areas:

(1) We will develop new detection methods to detect filament eruptions, flares and New Flux Emergences (EFRs), and compare the effectiveness of different methods.

These three kinds of features have different morphologies in solar chromosphere. Filaments are elongated dark features, flares are rapid brightenings and EFRs have dark fibril structure in the center of the growing region (arch filaments) and bright patches at footpoints. These three activities have different time scales: impulsive flares occur in on time scales of minutes; filament eruptions may be associated with flares, and the process lasts between minutes and hours; an EFR may take hours to a day to be fully developed. We will develop pattern recognition methods to detect all three kinds of dynamic phenomenon. Examples of existing solar feature include statistical pattern recognition method developed by Turmon et al. (2002) to label faculae and sunspots; neural network detection of flares by Fernandez Borda et al. (2002); and Machine Learning systems reviewed by Mjolsness and Decoste (2001). However, we found that detection of dynamic solar activities is still far from being mature.

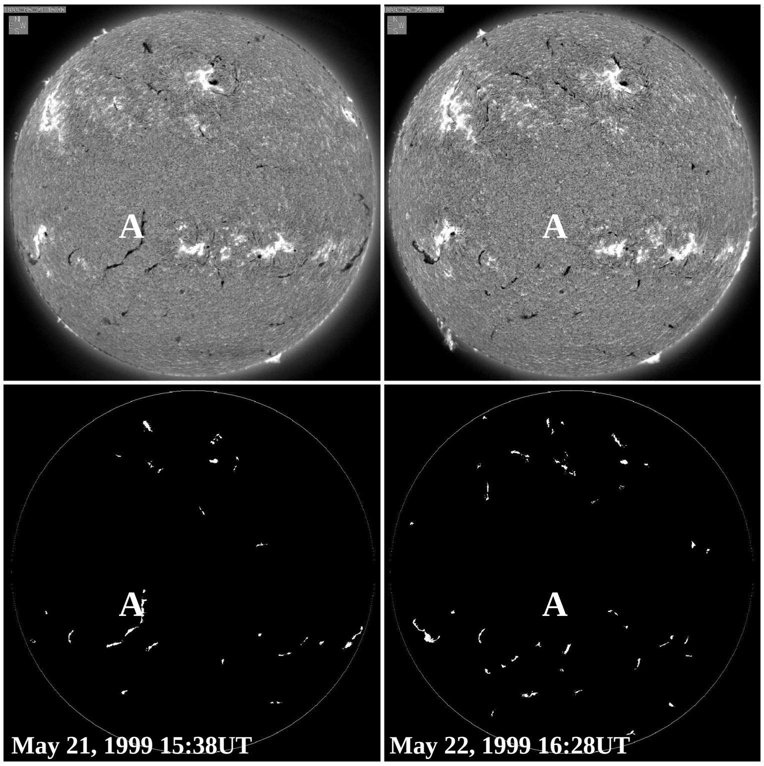

Figure 1: Images that demonstrate our detection of a filament disappearance: The top row shows two Hα images obtained on May 21 and 22, 1999, respectively; the bottom row shows filaments detected by our program. A big filament marked by "A" clearly disappeared, and was just as clearly detected by our program.

(2) We will characterize the detected solar features. There are many properties of solar feature worthy monitoring and we can not include t all. We will concentrate on the following features which are most relevant to the space weather research:

(a) We will record the orientation of filaments, in-conjunction with the magnetic polarity detection from full disk magnetograms to establish the orientation of magnetic clouds associated with the filament eruptions and CMEs (Yurchyshyn et al. 2001, 2003). We have reported on the relationship between the projection speed of CMEs, measured at $20R_\odot$ from SOHO/LASCO images, and the hourly averaged magnitude of the southwardly directed magnetic field, B_z, in interplanetary ejecta, as measured by the Advanced Composition Explorer (ACE) magnetometer. We found that intensity of the B_z is an exponential function of the CME's speeds. We also present data which support earlier conclusions on correlation of B_z and the Dst index. We have developed a method to predict the intensity of geomagnetic storms, which is based on the above mentioned correlations and the fact that the orientation and chirality of the erupted solar filaments correspond to the orientation and chirality of interplanetary ejecta. Such a method will be used in real-time, if we can detect the filament eruption in real time. This will give a 24 to 48 hours advanced warning of arrival of geomagnetic storms. Based on about 20 cases studied so far (not in real time), our prediction is remarkably accurate. The objective of the proposed study is to have automatic real-time predictions of earth-directed CMEs which have potential to produce geo-magnetic storms.

(b) We will automatically calculate the eruption speed of erupting filaments by tracking the front of filaments as a function of time. This will aid us to calculate the speed of CMEs.

(c) We will find the morphology of solar flares: in particular, if they are two-ribbon or compact flares. From one of our on-going studies, statistically, two ribbon flares may be more closely associated with CMEs.

(d) We will characterize the growth and separation speed of EFRs. EFRs are closely associated with the triggering of flare and CMEs (e.g. Nitta and Hudson, 2001; Feynman and Martin, 1995). In addition, the magnetic orientation of EFRs will be detected in real-time, as the regions that do not follow the Hale-Nicholson law would be more likely to produce solar flares.

B.3 Utilizing State-of-Art Pattern Recognition Methods

In this section, we discuss several methods that we intend to use for pattern recognition based on solar observations. The scientific goal is to develop new automatic procedures to detect disappearing filaments, flares and EFRs. The tasks of segmentation and classification are challenging since it is very difficult to take into account the complexity involved in the human classification process. The initial data to be used will be from the full disk Global Hα Network. The five stations in the network are at BBSO, Kanzelhöhe Solar Observatory (KSO) in Austria, Catania Astrophysical Observatory (CAO) in Italy, Yunnan Astronomical Observatory (YNAO) and Huairou Solar Observing Station (HSOS) in China. Nominally, each station obtains Hα image every minute with 1 arcsec pixel resolution. For the detection of EFRs, complementary real-time full disk magnetograms are very useful. BBSO can obtain such data set from Big Bear station of Global Oscillation Network Group (GONG). Of course, our techniques are applicable to other ground-based and space data sets, when they are available in real-time.

Features Extraction Using Image Segmentation Techniques

We have conducted the preliminary experiments of filament detection using local thresholding and morphological operators. The method consists of an advanced local thresholding preprocessing stage, separating dark features from a solar disk image, followed by a series of directional morphological filtering which extracts different shapes. Although, the method has been favorably compared against the existing technique (Gao et al., 2002), the step of local thresholding presents somewhat heuristic and consumes more time.

With additional work on filament registration, we propose the following steps. The initial preprocessing involves brightness normalization and limb darkening removal. We approximate the limb darkening by a mathematical function for adjustment. We also perform the equal-area normalization. Using the disk center and radius, the disk is transformed to achieve uniform area of solar features throughout the disk. After the preprocessing, we perform a global thresholding and employ morphological closing operations with multi-directional linear structuring elements to extract elongated shapes in the image. After logical union of directional filtering results, remaining noises are removed from the final outcome using morphological dilation and erosion with a circular structuring element.

We will extensively test the performance of the new procedures on serial solar images. The experiences in the segmentation of filaments will be transformed into the detection of other solar features, such as flares and EFRs.

Features Classification Using Neural Network Approaches

A neural network is an information-processing system that has certain performance characteristics in common with biological neural networks. Such a technique has been used to detect solar flares (Fernandez Borda et al., 2002). Neural networks have been developed as generalizations of mathematical models of human cognition or neural biology. A neural network is characterized by (1) its pattern of connections between the neurons, (2) its method of determining the weights on the connections, and (3) its activation function. The architecture consists of an input layer, a number of hidden layers, and an output layer. The processing ability of the network is stored in the inter-unit connection weights, obtained by a process of learning from a set of training patterns.

We have conducted research on an Adaptive-Resonance-Theory (ART) based neural architecture for pattern recognition and image enhancement without prior knowledge (Shih and Chang, 1992). In order to build the training set for the neural network, the solar features are selected from the BBSO database. We are currently using the following image features: (1) mean brightness of the image frame, (2) standard deviation of the brightness, (3) the brightness of the pixel of maximum brightness derivative between consecutive images, (4) radial position of that pixel, (5) variation of mean brightness between consecutive images, and (6) contrast between the point with maximum brightness derivative and the minimum of 5 by 5 neighborhood. The multilayer back-propagation neural network (Rumelhart et al., 1986) will be adopted for the solar features classification. The approach in our preliminary study looks promising. We will explore more image features statistically or structurely to classify different types of solar features.

RBF and SVM Approches

The focus of solar phenomena detection is the development of efficient feature-based classifiers. Two advanced techniques other than back-propagation neural networks are proposed: RBF classifiers and SVM classifiers. Radial Basis Function (RBF) network is a two-layer, hybrid feed forward learning network. It is a fully connected network and generally used as a classification tool. It was introduced by Broomhead and Lowe (1988). In a RBF model, the first layer from input nodes to hidden neurons is an unsupervised layer and the second layer from hidden neurons to output nodes is a supervised layer. Each hidden neuron has a radially symmetric basis function as its activation function. The characteristic feature of the basis function is that its response decreases (or increases) monotonically with distance from the center of the basis function. The centers, the distance scale, and the figure of the basis function are the parameters of the model. Typically, Gaussian-like and direct or inverse multiquadric-like radial basis functions are mostly used as activation functions. The purpose of the hidden neurons is to cluster the input data and reduce the dimensionality.

The transformation between the input nodes and hidden neurons is non-linear. It means that the first layer connecting the input nodes and the hidden neurons has unit weight and does not have to be trained. However, the transformation of the second layer is linear. The objective of the RBF network is by training data to minimize the sum of square errors and find the optimal weights between hidden neurons and output nodes. These optimal weights can classify effectively the test

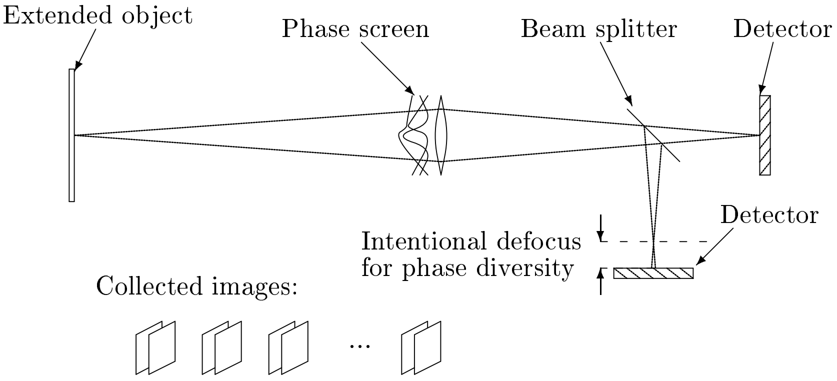

Figure 2: Typical data collection scheme, illustrating the data collection model. Only the straight through path is used for speckle interferometry. Short exposure images are collected for each of multiple atmospheric realization. For phase diversity methods, the conventional short-exposure images are supplemented with additional, simultaneous images that are intentionally defocused.

data into correct corresponding classes. There are some different training procedures like two-stage training, back-propagation training, and support-vector training that can be applied to RBF network (Schwenker et al., 2000). The applications of RBF network include the classification of 3-D visual objects, high resolution, handwritten digits (Schwenker et al., 2000), and face recognition (Er et al., 2002; Huang and Wechsler, 1999).

The basic idea of Support Vector Machine (SVM) is to separate two input pattern vectors into two classes by a hyperplane in the high dimensional space. What this means is that we always increment the dimension by one into if we fail to separate the input pattern vectors in dimension d. We do not stop the process until the dimension such that the desired property is satisfied. Hence, this dimension is called VC dimension (Vapnik, 1995).

SVM has recently been introduced as a new technique for solving a variety of learning and function estimation problems. Some applications using SVM include pharmaceutical data analysis, dynamic reconstruction of a chaotic system, and time series prediction.

C. Real-Time Image Reconstruction with Parallel Computer Cluster

C.1 Introduction – Two Different Methods

Seeing is the primary obstacle to obtaining high resolution time-series solar observations from the ground. Turbulence in the atmosphere along the line of sight randomly distorts the wavefronts. The result is blurring and geometrical distortions in the collected images.

Adaptive optics (AO) can fix some of the problem, but the high bandwidth correction needed allows only partial correction. Also AO is not so effective for large fields of view because of anisoplanatism. Large fields of view are required for space weather applications. The solution is post processing of image data in subfields smaller than the anisoplanatic angle, which for many applications is on the order of $\lesssim 5''$. The problem is to separate the object and the aberrations, two unknown quantities, in the images in the presence of noise, a third unknown. The exposure time for each frame must be short enough that the evolving atmosphere can be regarded as frozen during the exposure. The data collection model is shown in Fig. 2. Two different approaches have been developed and used for larger solar data sets: Speckle Interferometry and Phase Diversity.

Speckle imaging is a relatively mature technique for obtaining fine-resolution images in the

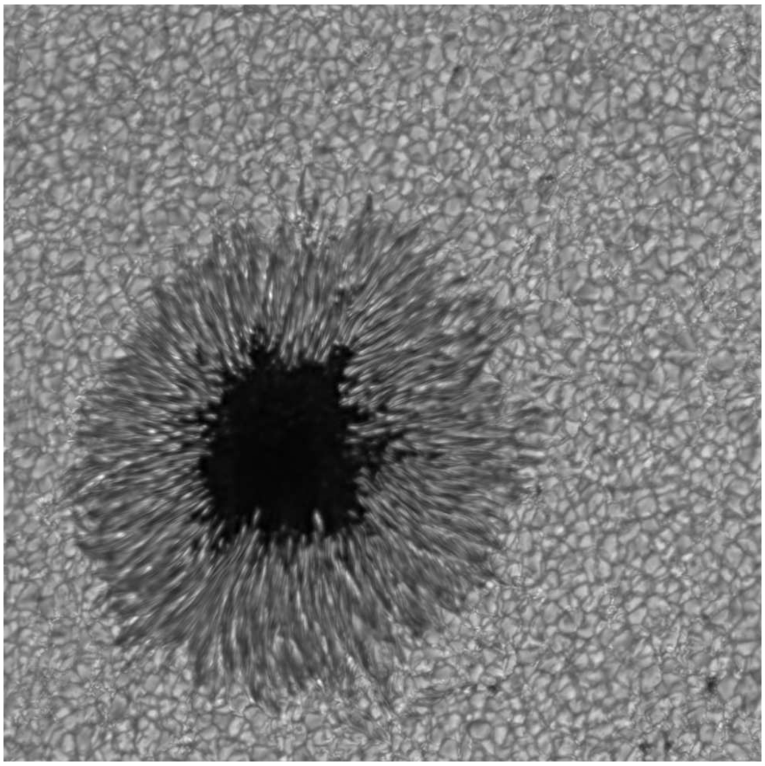

Figure 3: Speckle reconstruction of active region NOAA 8673 observed with the 1k × 1k pixel, 12-bit SMD CCD camera on 27 August 1999. The FOV is $76'' \times 76''$, the exposure time is 4 ms, and the Fried-parameter amounts to 12.6 cm.

presence of atmospheric turbulence. This technique requires the collection of many short-exposure images of a static object. Clever processing of this short-exposure time series affords the restoration of fine-resolution information that would be irretrievable if a single, long-exposure image were collected (Dainty, 1984). Speckle imaging requires the collection of 100 or more images so that the ensemble average over the class of all possible realizations may be approximated by an arithmetic average over a finite number of realizations. Speckle-imaging methods have been successfully adapted to the solar-imaging problem (Keller and von der Lühe, 1992; de Boer and Kneer, 1992; von der Lühe, 1993, 1994). The parallel operation for speckle imaging has been implemented at BBSO (Denker et al., 2001). This will not be one of the primary topics in the proposal. However, our experience in speckle imaging with parallel cluster will aid the more important component of the proposal: the phase diversity speckle imaging. Figure 3 shows an example of speckle reconstruction using 1K by 1K camera at BBSO.

Phase Diversity (PD) methods, see Section C.2, utilize a model of the image formation process to constrain the estimation of the aberrations and the common object in two image channels with a known difference in the pupil phase, usually a focus shift. An image collected in this second optical channel will contain the effects of the unknown phase aberrations but will also be influenced by the intentional defocus, which adds a known quadratic phase.

It is somewhat remarkable that estimates for the object and the unknown aberrations can be made from these two images, given the known quadratic phase diversity. However, in stressing regimes (e.g., poor seeing or weak signal levels) a single pair of phase-diversity images may not contain enough information to estimate jointly and with high fidelity the object and wavefront. Even under favorable conditions for which phase diversity is able to produce good wavefront estimates, we have observed that object information at isolated spatial frequencies may be irretrievably lost, resulting in object estimates with significant artifacts. These cases motivate a third fine-resolution imaging technique referred to as Phase-Diverse Speckle (PDS – see Paxman et al., 1992b), which requires the simultaneous collection of one conventional short-exposure image and at least one short-exposure image with phase diversity, for each of multiple atmospheric realizations. The problem with missing information at isolated frequencies can be fixed by using PD separately on each image pair in a PDS data set but then using the full set to make a single restoration of the common object (Löfdahl and Scharmer, 1994), a technique referred to as Partitioned PDS (Paxman et al., 1996). For the problem with poor seeing and weak signal, however, the solution is to jointly process the whole set (Paxman et al., 1992b; Paxman and Seldin, 1994); this method goes by the name Joint PDS (Paxman et al., 1996).

C.2 Phase Diverse Speckle Imaging

First proposed by Gonsalves and Chidlaw (1979); Gonsalves (1982), PD was not developed for solar imaging until in the nineties, when PDS (Paxman et al., 1992b) was usually incorporated. Typically, PD methods use two light channels, one in-focus and one out-of-focus image; the aberrations are then estimated and corrected by combining the information in images from the two channels. The out-of-focus channel can be conveniently arranged with a simple beam splitter and a second detector array that is translated along the optical axis as in Fig. 2 or with a combined beam splitter and mirror that puts both images on the same detector, e.g. as shown in Fig. 4.

PD methods, like Speckle methods, require short-exposure images ($\lesssim$ 10 ms) to 'freeze' the wavefront aberrations, which makes it possible to separate the object information and the information on atmospheric turbulence. Also the preprocessing is similar with the usual corrections by means of the average dark and flat field image. The image displacements within the sequence are estimated along with the other aberrations. The image formation model utilized by PD methods is only valid within the isoplanatic patch so large field of view PD restorations are typically mosaics of smaller restored subfields.

The first uses of PD to retrieve fine-resolution solar images were reported by Löfdahl and Scharmer (1993, 1994); Seldin and Paxman (1994). Since then it's been used for large scale data processing by Löfdahl et al. (1998, 2001), for investigations of G-band bright point dynamics

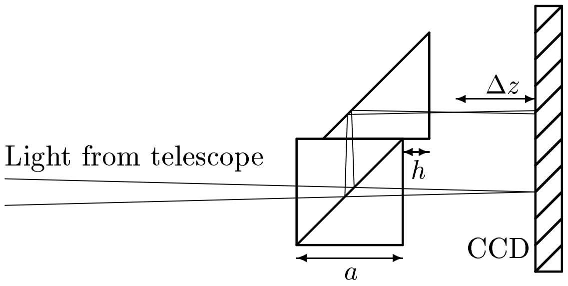

Figure 4: A simple PD detector can be made by gluing a cube beam splitter and a prism together in this configuration. Varying the size of the cube, a, and a displacement, h, between cube and prism allows some freedom for setting the amount of diversity displacement, Δz, and matching the separation of the two images to the detector size.

(Berger et al., 1998b,a; van Ballegooijen et al., 1998). Other projects include reconstruction of white-light, line-wing, and magnetogram data from a low-order AO system for demonstrating new results in the evolution of small-scale magnetic field, as well as sunspot photometry (Tritschler and Schmidt, 2002; Keller et al., 2000).

Very recently, Löfdahl (2002) developed an algorithm for Multi-Frame Blind Deconvolution (MFBD) together with a set of linear equality constraints that describe the data collection scheme. PD and JPDS are cases that can be treated with this approach, as demonstrated by Löfdahl and Scharmer (2002, see Fig. 5), who JPDS processed a long sequence of images from the new Swedish 1-m Solar Telescope for studying the fine structure of sunspot penumbrae (Scharmer et al., 2002), see images and a movie at `http://www.solarphysics.kva.se/NatureNov2002/`. The new formulation also accommodates joint processing of other kinds of data sets, e.g. several simultaneous PDS data sets in different wavelengths, see Fig. 6. The JPDS codes currently used at NSO which are developed by Paxman in the Environmental Research Institute of Michigan (known as ERIM codes) are proprietary and the source code is not available. Our proposed new method will have the greater flexibility and will be free available to science and engineering community. Dr. Löfdahl will be the main contributor to this effort.

Here follows a description of the approach of Löfdahl (2002). Unless auxiliary information is used, the Blind Deconvolution problem for a single frame is not well posed because the object and PSF information in a data frame cannot be separated. There are different ways of bringing auxiliary information to bear on the problem. MFBD uses several frames which helps somewhat, because the solutions are constrained by a requirement that the object be the same, but is often not enough to get useful results without further constraints. One class of MFBD methods constrain the solutions by requiring that the PSFs correspond to wavefronts over a certain pupil geometry, expanded in a finite basis. This is an effective approach but there is still a problem of uniqueness in that different phases can give the same PSF. Phase Diversity and the more general PDS methods are special cases of this class of MFBD, where the observations are usually arranged so that in-focus data is collected together with intentionally defocused data, where information on the object is sacrificed for more information on the aberrations. The known differences and similarities between the phases are used to get better estimates.

We first present a technique for jointly estimating the common object and the aberrations in a series of images that differ only in the aberrations. With no extra information beyond the image formation model, including aberrations from the phase in the generalized pupil transmission function, it is a Maximum-Likelihood (ML) Multi-Frame Blind Deconvolution (MFBD) (Schulz, 1993) method. Constraining the PSFs to be physical by requiring that they come from an underlying parameterization of the phase over the pupil is a powerful technique. Although such methods do work (Schulz et al., 1997; Van Kampen and Paxman, 1998), methods using more information work better (Tyler et al., 1998). Data sets used in Phase-Diverse Speckle (PDS) interferometry (Paxman

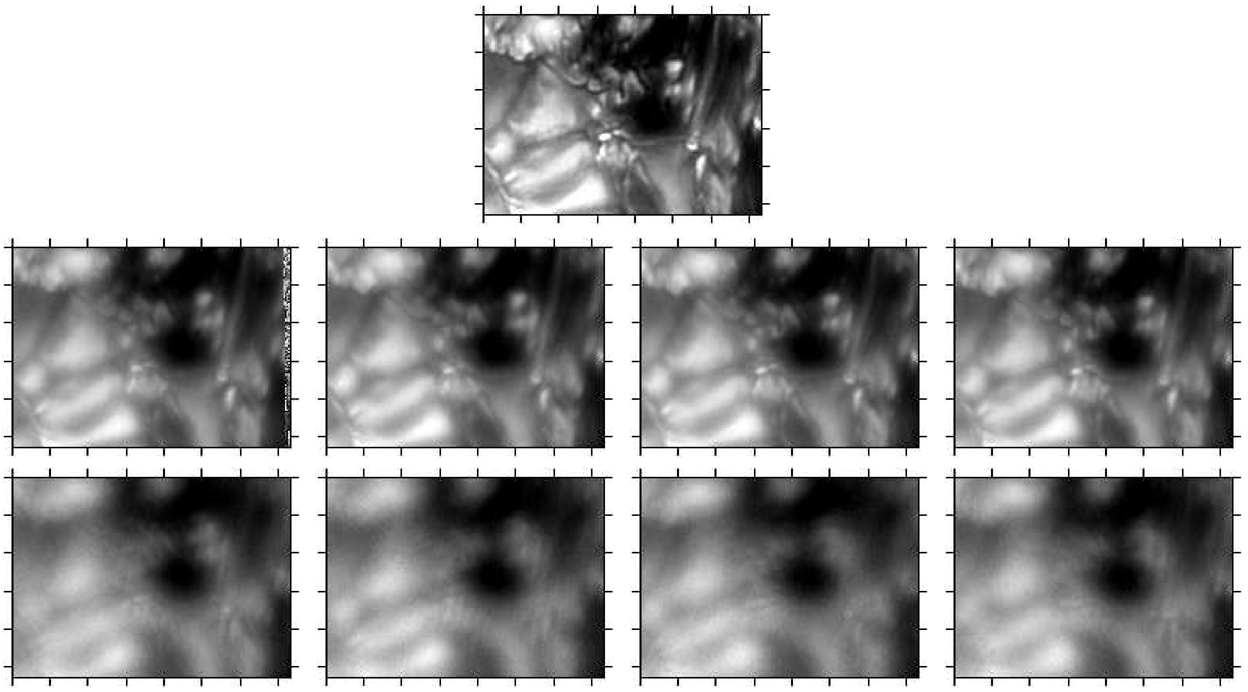

Figure 5: JPDS restoration of G-band 430.5 nm images from the new Swedish 1-m Solar Telescope. Top row: restored object; mid row: conventional in-focus images; bottom row: corresponding diversity out-of-focus images. Tick marks are $1''$ apart.

et al., 1992b; Löfdahl and Scharmer, 1994; Paxman et al., 1996) has such extra information in the form of two (or more) imaging channels with a known difference in phase (at least to type), with Phase Diversity (PD) (Gonsalves, 1982; Paxman et al., 1992a; Löfdahl and Scharmer, 1994) as the special case of only one such pair (or set). These are all different data collection schemes lending themselves to similar joint inversion techniques. The purpose of the formulation presented here is to recognize the similarities and outlining a method for treating them all with a single algorithm.

The intensity distribution in the focal plane of a telescope $i(\vec{x})$ is given by the convolution of the intensity distribution of the object $i_0(\vec{x})$ and the point spread function (PSF) $h(\vec{x})$ which describes the atmospheric turbulence along the line-of-sight

$$i(\vec{x}) = i_0(\vec{x}) * h(\vec{x}) \quad . \tag{1}$$

We use an isoplanatic image formation model with Gaussian white noise, which means that we assume that the optical system can be characterized by a generalized pupil function, which can be written, for an image frame with number $j \in \{1, \dots, J\}$, as

$$P_j = A_j \exp\{i\phi_j\}, \tag{2}$$

where ϕ_j is the phase and A_j is a binary function that specifies the geometrical extent of the corresponding pupil. A data frame d_j can then be expressed as the convolution of an object, f, and a point spread function (PSF), $s_j = |\,\mathfrak{F}^{-1}\{P_j\}|^2$. In the Fourier domain we get

$$D_j(u) = F(u) \cdot S_j(u) + N_j(u), \tag{3}$$

where u is the 2-D spatial frequency coordinate, S_j is the OTF, N_j is an additive noise term with Gaussian statistics, appropriate for low contrast objects like the solar photosphere, see also Eq. (1). For brevity, we will drop the u coordinate for the remainder of this section.

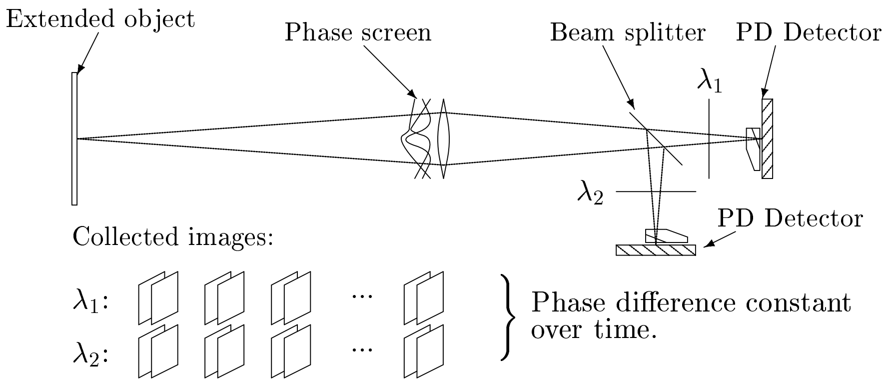

Figure 6: Extended data collection scheme. Here, the two detectors are equipped with PD beam splitters similar to the one in Fig. 4, so each camera collects PDS data sets, simultaneously in two different wavelengths, λ_1 and λ_2, respectively. See also Fig. 2.

We parameterize the unknown pupil phases by expanding them in a suitable basis, $\{\psi_m\}$, allowing for a part of the phase, θ_j, to be excepted from the expansion,

$$\phi_j = \theta_j + \sum_m^M \alpha_{jm}\psi_m; \qquad \forall j. \tag{4}$$

The Gaussian white noise assumption allows us to use the inverse Wiener filter estimate of the object,

$$F = \frac{1}{Q}\sum_j S_j^* D_j, \tag{5}$$

where

$$Q = \gamma_{\mathrm{obj}} + \sum_j^J |S_j|^2, \tag{6}$$

to derive a metric in a form that does not explicitly involve the object (Gonsalves, 1982) and that has been shown to correspond to a ML estimate of the phases (Paxman et al., 1992a). With two regularization parameters (Vogel et al., 1998), this metric can be written as

$$L(\boldsymbol{\alpha}) = \sum_u \left[\sum_j^J |D_j|^2 - \frac{|\sum_j^J D_j^* S_j|^2}{Q}\right] + \frac{\gamma_{\mathrm{wf}}}{2}\sum_m^M \frac{1}{\lambda_m}\sum_j^J |\alpha_{jm}|^2. \tag{7}$$

When minimizing L, the γ_{obj} term in Q has the effect of establishing stability with respect to perturbations in the object. Its use in Eq. (5) suggests setting γ_{obj} to something proportional to the inverse of the signal to noise ratios of the image data frames. The other regularization parameter, γ_{wf}, stabilizes the wavefront estimates and can be set by examining the relation between L and the RMS of the wavefront (Engl et al., 1996). For simplicity, this term is given under the assumption that the ψ_m are atmospheric Karhunen–Loève functions (Roddier, 1990), where λ_m is the expected variance of mode m. In the general case, the wavefront regularization term is a matrix operation involving the covariances (Vogel et al., 1998).

Note that, although presented in a PD setting, this metric has nothing to do with PD per se, it is just a MFBD metric. The PD part enters in the parameterization in earlier PD methods and, equivalently, in the LECs in this presentation.

Efficient minimization of L requires the gradients and, for some methods, the Hessian with respect to the aberration parameters.

We lexicographically arrange all the α_{jm} in a single column vector, $\boldsymbol{\alpha}$, with $N = JM$ elements, so that we can also refer to them as α_n, with a single index $n = (j-1)M + m$,

$$\boldsymbol{\alpha} = [\alpha_{11} \quad \alpha_{12} \quad \ldots \quad \alpha_{JM}]^{\mathrm{T}} = [\alpha_1 \quad \ldots \quad \alpha_N]^{\mathrm{T}}. \tag{8}$$

We can then write the normal equations of the MFBD problem as a matrix equation,

$$\mathbf{A}^{\mathrm{MFBD}} \cdot \delta\boldsymbol{\alpha} - \mathbf{b}^{\mathrm{MFBD}} \simeq 0, \tag{9}$$

where the elements of the right hand side vector can be written as

$$b_{jm}^{\mathrm{MFBD}} = \left\langle -2\sqrt{J}\,\mathrm{Im}\left[P_j^*\,\mathfrak{F}\{p_j\,\mathrm{Re}[\mathfrak{F}^{-1}\{F^* D_j - |F|^2 S_j\}]\}\right], \psi_m \right\rangle + \gamma_{\mathrm{wf}}\frac{\alpha_{jm}}{\lambda_m}. \tag{10}$$

The constrained minimization problem in $\boldsymbol{\alpha}$ can be transformed into an unconstrained minimization problem in a reduced set of variables, $\boldsymbol{\beta}$, an N' element vector of parameters, $\beta_{n'}$. The normal equations for the transformed problem are obtained by left-multiplying Eq. (9) with the null space matrix of the constraints pertinent to the data collection scheme used, $\mathbf{Q}_2^{\mathrm{T}}$, and substituting $\mathbf{Q}_2\delta\boldsymbol{\beta}$ for $\delta\boldsymbol{\alpha}$,

$$\mathbf{Q}_2^{\mathrm{T}}\mathbf{A}^{\mathrm{MFBD}}\mathbf{Q}_2 \cdot \delta\boldsymbol{\beta} - \mathbf{Q}_2^{\mathrm{T}} \cdot \mathbf{b}^{\mathrm{MFBD}} \simeq 0. \tag{11}$$

Once we have a solution for $\boldsymbol{\beta}$, we easily get the solution for $\boldsymbol{\alpha}$ by back substitution. With the appropriate null matrix, this is equivalent to the traditional JPDS algorithm for Gaussian additive noise.

Phase diversity imaging is particularly suited to data obtained behind an AO system since it does not require an amplitude calibration.

C.3 Parallel Computing Implementation of Phase-Diverse Speckle Imaging

Our research objective is to create, in the terminology of `http://www.darpa.mil/baa/baa02-09.htm`, a High Productivity Computing System (HPCS) for space weather research that will use state-of-art parallel computing methodologies and tools to optimize phase diverse speckle codes, and to obtain large size diffraction limited data in real time.

As we mentioned earlier, BBSO has developed a 32-node parallel system to carry-out real time speckle reconstruction (Denker et al., 2001) for the 1024 by 1024 image sequences. The code is written in C++. It provides one processed image each minute. The number of raw data used is about 100 (selected out of 500). For the current phase-diverse speckle imaging system in the National Solar Observatory (NSO), it uses 10 pairs of best images out of several hundreds. The current computation time is around 1 hour. Our goal is to provide real time data in the cadence of 10 seconds.

Given the high volume of data and processing times, that involve image preprocessing, FFT operations, iterative methods that themselves involve sparse matrix computations and FFT operations, parallelism provides probably the only choice to reduce running and response times to achieve real-time or close to real-time performance. In general, the use of parallelism can improve performance in two ways:

1. better *scalability* of the massive number of images is possible, since larger images or increasing number of images can be handled within the same response time (response time remains the same, large problem sizes can be solved within the same amount of time),

2. a significant reduction in *response time* for a given problem size can be achieved (problem size remains the same, response time decreases).

In this research effort we will mainly concentrate on achieving objective 2; objective 1, however, may be dealt with if increasing the number of images processed could substantially improve performance or the reliability of the predictions.

Achieving high performance in any multiprocessor system is a challenging task. This is more so for the cluster of PC workstations that will be used for computational support in this project. Several issues need to be resolved related to optimizing processor utilization (eg. workload balancing among the processors), minimizing interprocessor communication and synchronization time, and achieving optimal data distribution and decomposition. We have been able to deal with all these issues successively in the past (Gerbessiotis and Valiant, 1994; Gerbessiotis and Siniolakis, 1999, 2002b,a; Gerbessiotis et al., 2002).

The implementation of the HPCS system for space weather prediction will be carried out on top of the Message Passing Interface architecture (MPI) (Pacheco, 1997; Snir et al., 1998; Gropp et al., 1998, 1999b,a); the LAM-MPI (Burns et al., 1994) implementation is readily available and preinstalled on various Linux distributions. We have had considerable experience dealing with such issues in a PC cluster environment (Gerbessiotis, 2002a,c,b). In our implementation efforts we will use C/C++ and avoid a less efficient mixing of C and the ANA programming language that was used before. This would make our system simpler to use and decrease response times to satisfy the criterion of real-time performance.

A variety of parallel paradigms will be explored and utilized in this research effort. Because of the sheer volume of image data *task parallelism* Kumar et al. (2001) involving the image sequences will be used. For real time speckle reconstruction for example, this would lead to the distribution of 100 or more images evenly among the cluster processors to initiate individual and independent processing. Such an approach may not require any kind of complex and nontrivial data distribution other than the obvious one of distributing the images. For phase-diverse speckle imaging the number of isoplanatic subfields of approximately 100×100 pixels may reach up to 100 in number, with 6-20 image frames per subfield. The image frames of a subfield after processing would generate a restored image of the subfield and the combination of the restored images of the 100 subfields would create a restored image of the entire field of view (FOV). This natural data distribution of the entire FOV into subfields will be utilized for the task of allocating jobs (subfield processing) to processors and the subsequent combination of the restored subfields. Keeping track of the busy processors from the idle or unavailable processors and the assignment of processors to various tasks will involve the use of a static load balancing scheme in the proposed HPCS system that will ensure that processor utilization be maintained high; such a method will also warrant some level of fault tolerance and increased availability.

The nature of the problems that need to be solved and the use of task parallelism, does not necessitate at least initially the use of dynamic load balancing approaches to replace the static approaches that will be employed. High availability will be maintained by keeping track of the state of the nodes of the cluster; this information will be collected and used by the load-balancer responsible for task allocation.

In the recent past we have performed work on parallel numerical computing applications and we designed, analyzed and also implemented architecture independent parallel algorithms for various decomposition methods (Cholesky, Gauss-Jordan elimination, etc.) and computational problems in finance (eg. option price valuations on binomial and trinomial trees and with the Blacks-Scholes formula using the explicit finite difference method). We showed theoretically the optimality of our designs and verified these conclusions by carrying out experimental studies that confirmed the theoretical predictions on a variety of parallel platforms that included the Cray T3D, IBM SP2, SGI Power Challenge, and cluster of SMP PCs. In all such problems we had to deal with issues related to data distribution and work-load balancing in achieving optimal performance. In addition the architecture independent design of our algorithms allowed us to reliably predict their performance before we ran them on a particular parallel platform. These theoretical prediction tools (Valiant, 1990; Gerbessiotis and Valiant, 1994) will be employed in this research effort, as they will allow us to decide early on which approach will lead to significant improvements in performance.

We have also worked on problems of combinatorial computing such as sorting, searching and selection and their applications to parallel heaps and parallel priority queues. In such a setting we explored whether for certain problems it makes more sense to design a purely parallel and new (but also, complex) data structure or it makes more sense to use a global structure consisting of local disjoint copies of the corresponding sequential structure on every processor and use parallel methods to search and retrieve items from the data structure and maintain the global structure synchronized. Such an experience will also be useful in this effort.

The end-result of this research effort will be the use of our experiences in parallel algorithm design and implementation in building a High productivity computing system for space weather research. Our proposed cluster will be stationed at BBSO in California. In our initial effort we will use the PC Cluster Laboratory that has been established at the CS Department at NJIT with partial support from NSF under MRI grant NSF/9977508. As part of an on-going effort under this project, design of parallel algorithms and development of parallel code for a variety of problems has been completed. The developed code so far includes

1. primitive operations related to broadcasting and parallel prefix, memory efficient (but synchronization inefficient) and memory inefficient (but synchronization efficient) versions of dense matrix multiplication,
2. binomial-tree based stock option price valuations,
3. trinomial-tree based parallel option price valuations,
4. parallel option valuations by the Black-Scholes formula using the explicit finite difference method,
5. dense matrix computations (eg. matrix inversion, Gauss-Jordan elimination and Cholesky factorization) and functions that allow the distribution to processors and collection from processors of matrices under different distribution patterns, and
6. development of benchmark programs that establish the communication and synchronization capabilities of the communication network of a parallel platform.

According to Moore's law of exponential growth in computing power, we can expect an increase of computing power by a factor of 16 over the next 6 years which will allow us to operate 1k × 1k or even 2k × 2k pixel CCD cameras for phase-diverse speckle imaging. This will be right in time for a next generation of 4 meter-class telescopes which will provide a spatial resolution of $0\overset{\prime\prime}{.}05$ at 500 nm, i.e., a field of view (FOV) of about $100'' \times 100''$ for a 2k × 2k pixel CCD camera which is sufficient to cover large sunspots and (substantial parts of) active regions. We are certain that large format, high frame rate CCD cameras for phase-diverse speckle imaging in combination with parallel computers will become standard post-focus instruments for advanced solar telescopes.

D. Distribution and Application of Data and Results for Space Weather Research

D.1 Scientific Applications of Reconstructed Data

The system discussed above will be a powerful tool for many areas in high resolution solar physics research. As the scientific objective is for real space weather, we will concentrate on a targeted study: diffraction limited atmospheric structure of flaring active regions.

There are two very important aspects of evolution of active regions which may be closely associated with solar flares: (1) new flux emergence and (2) unusual flow motions. These two

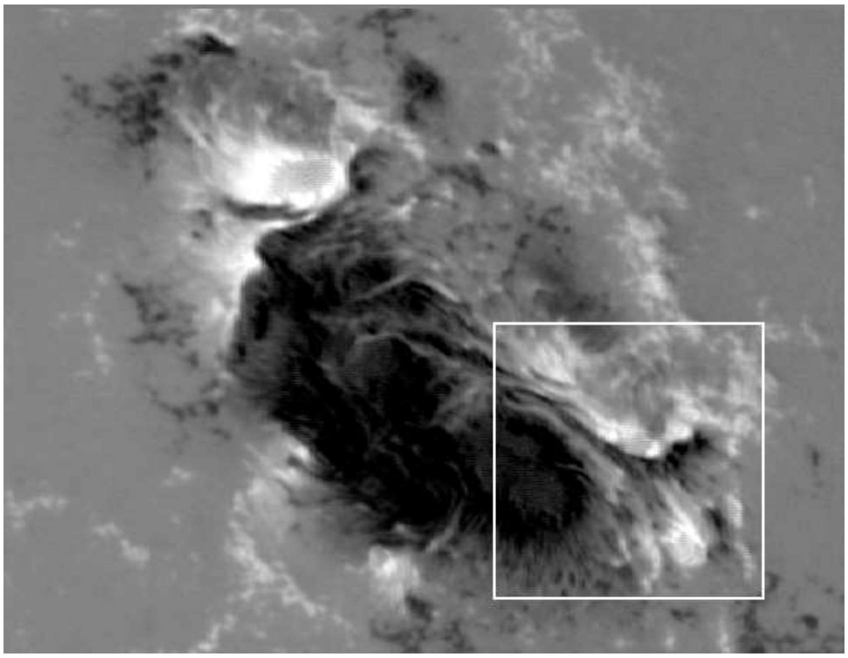

Figure 7: A line-of-sight magnetogram of June 10, 1991 active region NOAA 6659 obtained at Big Bear Solar Observatory. It is a δ sunspot group. The field of view is 5 by 4 arcmin. White and black gray scales represent field strength of positive and negative magnetic polarities, respectively. The dominant negative spots are surrounded by opposite polarity satellite spots. On the lower-right corner (marked by a white box), magnetic channel structure is visible due to new flux emergence.

kinds of phenomenon may also be associated (Tanaka, 1991; Zirin and Wang, 1993). A few well-known active regions showed the correlation between flux emergence, flows and major flares. Two of the best examples are March 1989 NOAA 5395 (Wang et al., 1991) and June 1990 NOAA 6659 active regions (Zirin and Wang, 1993) of NOAA 6659. In Figure 7, we demonstrate the "magnetic channel" structure as discussed by Zirin and Wang (1993). The channel was created by the new flux emergence and appears as complicated flows along the channels. The area with strong flows, new flux emergence, and complicated channel structure was also obviously associated with several major flares in this active region. The flux emergence destabilizes the existing magnetic configuration and cause magnetic reconnection. Flow motions may contribute to the twist of flux ropes by moving footpoints around. Such flow and channel structure is not well understood – it requires long sequences of high resolution observations. Studies at BBSO have demonstrated the value of diffraction limited observations in the flaring region (Denker and Wang, 1998).

If the speckle and phase diversity observing sequences are processed off-line (as we are doing now), the chance of studying the long term evolution of an active region and catching flares during the observations is slim. Because we store all the raw images in hard disk, it is usually filled up in half hour, and it will take many hours to several days to process the data. Our proposed system will only save constructed image data – therefore, the disk will be big enough for several day's continuous observation. The real time data product with the diffraction-limited quality is essential in this study. When this parallel cluster is built and operational, we will follow all flare productive regions, with an image scale of $0.''1$. The active regions flows will be derived by local correlation tracking and feature tracking. We expect to answer the following questions:

1. Do flaring active regions always have the complicated flows as discussed above?

2. Are these flows always associated with flux emergence?

3. Are flows unusually strong before major flares? Monitoring of flows will provide a prediction

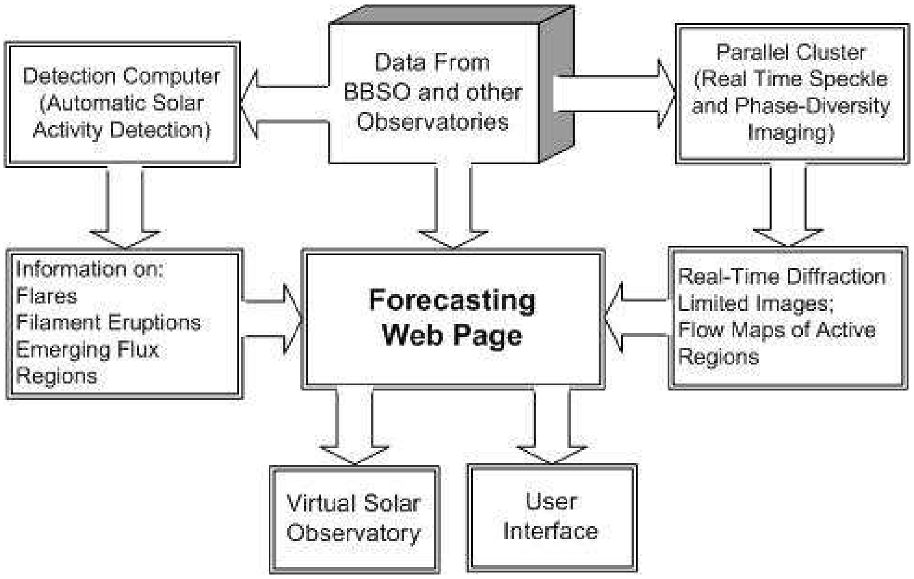

Figure 8: A block diagram showing the interconnection of different research components.

tool for solar flares and space weather.

We anticipate that we will answer the above questions better, when we combine our observations with high resolution flare observations of satellites such as RHESSI and TRACE, in understanding the detailed energy build-up and release processes.

D.2 Real Time Web Access of Data and Space Weather Forecasting

Our ultimate goal is to provide high quality data and space weather forecasting in real time. BBSO now has established a web server, and our automatic solar activity detection and real-time diffraction solar images will be important additions to the system.

In our current on-line system, every hour, images from our Global Hα Network are transferred to the web-based BBSO Active Region Monitor (ARM, `www.bbso.njit.edu/arm`), which also includes the most recent full-disk EUV, soft X-ray, continuum and magnetogram data from the Solar and Heliospheric Observatory (SoHO) and YOHKOH (until December 2001). ARM also includes a variety of active region properties from the National Oceanic and Atmospheric Administration's (NOAA) Space Environment Center (SEC), such as up-to-date active region positions and flare identifications. Furthermore, we have developed a Flare Prediction System (FPS) which estimates the probability for each active region to produce C-, M-, or X-class flares based on nearly eight years of NOAA data from cycle 22. This, in addition to BBSO's daily solar activity reports, has proven to be a useful resource for activity forecasting (Gallagher et al., 2002).

The new Web pages will include real time report of all the filaments (and their disappearances), flares and EFRs. The size, location and two ribbon separation speed of flares, and information of disappearing filaments will be posted in real time. The real-time diffraction limited active region images will be also posted on the web, along with flows derived in real time by local correlation tracking and feature tracking.

Figure 8 shows the structure of new contents of space weather forecasting based on results from our proposed research, and the link among different components.

Virtual Solar Observatory (VSO) is a community effort, which attempts to distribute ground-based and space data to users uniformly and efficiently (Hill, 2001). We will actively collaborate with NSO group to include our real-time data and forecasting as a part of VSO. The VSO effort is lead by NSO. One of Co-PIs in our proposal, Dr. Keller is a team member of VSO. In addition, we include a high-end server in our budget and propose to develop a state-of-art user interface for quick and easy access of the data by the space weather community.

In order to maintain our objective of building a High Productivity Computing System for space weather research, we shall also develop a web-based interface to the parallel computing module previously described. The results of the data processing will become available through this interface. Open-source software tools will be used for the development of this interface and the access of the raw and the processed data. A Linux-based Apache web-server will be utilized. Web-based forms in HTML will be processed by Perl scripts and if needed C programs to generate and return the required responses by accessing a MySQL database server/engine.

The storage subsystem stores data according to different data type, logical data organization and transaction patterns. These data are managed by local Data Base Management Systems (DBMSs) specialized to their respective medium. Moreover, since data retrivals exhibit a temporal correlation such that a newly generated images are retrieved more frequently and older ones less frequently. It is cost-effective to have a hierarchical storage architecture, where older images are stored in a lower-cost but slower storage system (e.g., optical systems) and more recent images in a more expensive but faster storage system (e.g., magnetic disks).

E. Education and Outreach

Education and outreach are extremely important for the future of research in science and technology. Our proposed effort includes a strong component in the integrated research and education.

BBSO highly prizes its special role in the education of the next generation of scientists who build instruments. We graduated 6 Ph.D. students between 2000 and 2002. We currently have 11 Ph.D. students (9 men and 2 women). Three Ph.D. thesis will be as a result of this proposed research effort. Three students will callobarate with fellow students, each will emphasize in a particular topic of his/her specialty.

Thesis 1: Automatic Detection of Solar Activities. This student will be supervised by Shi and Rees. Several methods will be applied to the detection of solar flares, filament eruptions and CMEs, which includes segmentation, neural network and SVM.

Thesis 2: Real-time Imaging Processing. This student will be supervised by Denker, Gerbesstiotis, Keller and Löfdahl. The research tasks include developing efficient codes for phase-diversity speckle codes; setting up phase-diversity observations; parallel computing implementation and optimization.

Thesis 3, Space Weather Forecasting. This student will be supervised by Wang, Keller and Denker. The research tasks include development of data interfaces for VSO and users; development of space weather tools based on the products of theses 1 and 2; statistical studies among flares, filament eruptions, EFRs, CMEs and geomagnetic storms to advance space weather forecasting.

Several new courses will be developed and offered during this 5-year period, such as,“Innovative Pattern Recognition for Dynamic Features”; “Implementation of Parallel Computing in Imaging Processing”, “Space Weather Forecasting”. These courses do not currently exist at NJIT.

Each year, we budget two slots for undergraduate research. Each student will work on the project for a period between 1 to 3 years. They will assist Ph.D. students and faculty members in the above three research areas. Their tasks would include obtaining and processing selected data sets; writing part of the computer codes and part of the web design. The new courses will be flexible enough, so juniors and seniors can take these graduate courses. Two summer schools will be organized during the 5-year period to invite experts in imaging processing, parallel computing and artificial intelligence to give lectures.

We will enhance our public outreach program to improve the public awareness of both space weather research and innovative technologies. The BBSO public tours/lecture programs consist of three parts: 1) summer tours, offered every Saturday afternoon to the general public; 2) private tours given mostly to school groups and 3) on- and off-site lectures. The summer observatory tours are very popular partly because Big Bear is a resort town located close to the very populous Los

Angeles metropolitan area. We will develop materials suitable for general audience to understand the solar activity, the impact on earth and innovative technologies to forecast them.

Table 1: Team Members and Responsibilities

Investigator	Institute	Responsibilities
Haimin Wang	BBSO/NJIT, Professor	Overall Coordination and SW Forecasting
Carsten Denker	BBSO/NJIT, Assist. Professor	Speckle Imaging and Parallel Computing
Christoph Keller	NSO, Assoc. Astronomer	Phase Diversity Speckle and VSO
Frank Shih	CIS/NJIT, Professor	Feature Detection and Characterization
Alex Gerbessiotis	CIS/NJIT, Assist. Professor	Parallel Optimization and Data Management
David Rees	CSIRO, Scientist	Feature Detection and Characterization
Mats Löfdahl	RSAS, Scientist	Phase Diversity Speckle Codes

Table 2: Milestones

Category	Pattern Recognition	Real-time Imaging	Space Weather
FY 2003	Investigation of Research Methods Initial Codes	Speckle and Phase Diversity Codes	Statistical Studies of Solar Activities
FY 2004	Code Development Segmentation	Phase Diversity Speckle Codes	Web Design
FY 2005	Code Development Neural Network	Phase Diversity Speckle Codes	Data Structure User Interface
FY 2006	Code Development SVM and RBF	Observation Data Analyses	Space Weather Forecasting
FY 2007	Implementation	Implementation	Space Weather Forecasting

F. Management

F.1 Team Members and Responsibilities

In Table 1, we list the investigators and break-down of responsibilities. Wang is the PI, who is responsible for general coordination. Other team members have clear responsibilities in three basic areas: pattern recognition, real-time imaging and space weather. However, most research tasks have collaborative nature. In addition three Ph.D. students will work in these three major areas.

The majority part of the budget is for supporting Ph.D. students. Some summer salaries are requested for faculty members. It covers the 15% yearly effort of Drs. Rees and Löfahl, in the form of participant supports.

The initial development of various codes will be based on NJIT campus, with collaboration with three outside institutes (NSO, Royal Swedish Academy of Sciences and CSIRO). Phase-Diversity will be implemented at three locations, BBSO, NSO and Swedish 1-m Solar Telescope in Canary Island. These three sites are regarded as best sites for high resolution solar observations in the world. The experimental setting will be initially tested at Swedish 1-m Solar Telescope under the direction of Löfdahl. However, the final implementation will be at BBSO as parallel cluster will be located at BBSO.

As we discussed in the proposal, the phase diversity speckle will be used behind the corrected light beams by AO. NSO and BBSO AO systems are expected to be on line in 2003.

F.2 Research Milestones

The three proposed research components are closely related. We give milestones by years and by the projects in Table 2. FY 2003 starts on September 1, 2003.

Bibliography

Berger, T. E., Löfdahl, M. G., Shine, R. A. and Title, A. M.: 1998a, Measurements of solar magnetic element dispersal, *Astrophysical Journal* **506**, 439–449.

Berger, T. E., Löfdahl, M. G., Shine, R. A. and Title, A. M.: 1998b, Measurements of solar magnetic element motion from high resolution filtergrams, *Astrophysical Journal* **495**, 973–983.

Beveridge, R.: 1989, Segmenting images using localized histograms and region merging, *InternationalJournal of Computer Visio* **2**, 311–347.

Broomhead, D. and Lowe, D.: 1988, Multivariate functional inter- polation and adaptive networks, *Complex Systems* **2**, 321–355.

Burns, G., Daoud, R. and Vaigl, J.: 1994, Lam: An open cluster environment for mpi, *Proceedings of Supercomputing Symposium '94*, University of Toronto, pp. 379–386.

Dainty, J. C.: 1984, *in* J. C. Dainty (ed.), *Laser Speckle and Related Phenomena*, Vol. 9 of *Topics in Applied Physics*, Springer, p. 255.

de Boer, C. R. and Kneer, F.: 1992, Speckle observations of abnormal solar granulation, *Astronomy & Astrophysics* **264**, L24–L26.

Denker, C. and Wang, H.: 1998, High spatial resolution observations of a small delta spot, *Astrophysical Journal* **502**, 493–497.

Denker, C., Yang, G. and Wang, H.: 2001, Near real-time image reconstruction, *Solar Physics* **202**, 63–70.

Engl, H. W., Hanke, M. and Neubauer, A.: 1996, *Regularization of Inverse Problems*, Vol. 375 of *Mathematics and Its Applications*, Kluwer Academic Publishers, Dordrecht, Netherlands.

Er, M., Wu, S., Lu, J. and Toh, H.: 2002, Face recognition with radial basis function (rbf) neural networks, *IEEE TRansactions on Neural Networks* **13**, 697–710.

Fernandez Borda, R. A., Mininni, P. O., Mandrini, C. H., Gomez, D. O., Bauer, O. H. and Marta, G.: 2002, Automatic solar flare deteciton using neural network techniques, *Solar Physics* **206**, 347–357.

Feynman, J. and Martin, S.: 1995, The initiation of coronal mass ejections by newly emerging magnetic flux, *JGR* **100**, 3355–3367.

Gallagher, P., Moon, Y. and Wang, H.: 2002, Active-region monitoring and flare forecasting - i. data processing and first results, *Solar Physics* **209**, 171–183.

Gao, J.: 2002, *Development of an Automatic Filament Disappearance Detection System*, PhD thesis, New Jersey Institute of Technology.

Gao, J., Wang, H. and Zhou, M.: 2002, Development of an automatic filament disappearance detection system, *Solar Physics* **205**, 93–103.

Gerbessiotis, A. V.: 2002a, Parallel option price valuations with the explicit finite difference method, *Technical report.* Submitted to Journal of Parallel Computing.

Gerbessiotis, A. V.: 2002b, Parallel option price valuations with the explicit finite difference method, *Technical report.* Submitted to Journal on Cluster Computing.

Gerbessiotis, A. V.: 2002c, Trinomial-tree based parallel option price valuations, *Technical report.* Submitted to Journal of Parallel Algorithms and Applications.

Gerbessiotis, A. V. and Siniolakis, C. J.: 1999, Efficient deterministic sorting on the bsp model, *Parallel Processing Letters* **9**, 69–79.

Gerbessiotis, A. V. and Siniolakis, C. J.: 2002a, Architecture independent parallel algorithm design: Theory vs practice, *Future Generation Computer Systems* **18**, 573–593.

Gerbessiotis, A. V. and Siniolakis, C. J.: 2002b, Architecture independent parallel selection with applications to parallel priority queues, *Theoretical Computer Science.* In press.

Gerbessiotis, A. V. and Valiant, L. G.: 1994, Direct bulk-synchronous parallel algorithms, *Journal of Parallel and Distributed Computing* **22**, 251–267.

Gerbessiotis, A. V., Siniolakis, C. J. and Tiskin, A.: 2002, Parallel priority queue and list contraction: The bsp approach, *Computing and Informatics* **21**, 1–32.

Gonsalves, R. A.: 1982, Phase retreival and diversity in adaptive optics, *Optical Engineering* **21**(5), 829–832.

Gonsalves, R. A. and Chidlaw, R.: 1979, Wavefront sensing by phase retrieval, *in* A. G. Tescher (ed.), *Applications of Digital Image Processing III*, Vol. 207 of *Proc. SPIE*, pp. 32–39.

Gropp, W., Huss-Lederman, S., Lumsdaine, A., Lusk, E., Nitzberg, B., Saphir, W. and Snir, M.: 1998, *MPI -The Complete Reference: Volume 2, The MPI-2 Extensions*, MIT Press, Cambridge, MA.

Gropp, W., Lusk, E. and Skjellum, A.: 1999a, *Using MPI: Portable Parallel Programming with the Message Passing Interface, 2nd Edition*, MIT Press, Cambridge, MA.

Gropp, W., Lusk, E. and Thakur, R.: 1999b, *Using MPI-2: Advanced Features of the Message-Passing Interface*, MIT Press, Cambridge, MA.

Harrison, R. A.: 1994, A statistical study of the coronal mass ejection phenomenon, *Advances in Space Research* **14**, 23–28.

Harrison, R. A.: 1995, The nature of solar flares associated with coronal mass ejection, *Astronomy & Astrophysics* **304**, 585–595.

Hill, F.: 2001, The virtual solar observatory - status and plans, *AGU Spring Meeting.*

Huang, J. and Wechsler, H.: 1999, Eye detection using optimal wavelet packets and rbfs, *International Journal of Pattern Recognition and Artificial Intelligence.* In press.

Keller, C. U.: 1995, Properties of solar magnetic fields from speckle polarimetry, *Reviews in Modern Astronomy* **8**, 27–60.

Keller, C. U. and von der Lühe, O.: 1992, Solar speckle polarimetry, *Astronomy & Astrophysics* **261**, 321–328.

Keller, C. U., Rimmele, T. R., Paxman, R. G., Seldin, J. H., Garrara, D. and Gleichman, K.: 2000, Evolution of small-scale magnetic fields from combined adaptive optics and phase-diverse speckle imaging, BAAS, 32, 833.

Kumar, V., Grama, A., Gupta, A. and Karypis, G.: 2001, *Introduction to Parallel Computing: Design and Analysis of Algorithms*, Benjamin Cummings.

Löfdahl, M. G.: 2002, Multi-frame blind deconvolution with linear equality constraints, *in* P. J. Bones, M. A. Fiddy and R. P. Millane (eds), *Image Reconstruction from Incomplete Data II*, Vol. 4792-21 of *Proc. SPIE.* In press.

Löfdahl, M. G. and Scharmer, G. B.: 1993, Phase-diversity restoration of solar images, *in* R. R. Radick (ed.), *Real Time and Post Facto Solar Image Correction*, Proc. 13th Sacramento Peak Summer Workshop Series, pp. 89–104.

Löfdahl, M. G. and Scharmer, G. B.: 1994, Wavefront sensing and image restoration from focused and defocused solar images, *Astronomy & Astrophysics Supplement Series* **107**, 243–264.

Löfdahl, M. G. and Scharmer, G. B.: 2002, Phase diverse speckle inversion applied to data from the Swedish 1-meter solar telescope, *in* S. Keil and S. Avakyan (eds), *Innovative Telescopes and Instrumentation for Solar Astrophysics*, Vol. 4853-54 of *Proc. SPIE.* In press.

Löfdahl, M. G., Berger, T. E. and Seldin, J. H.: 2001, Two dual-wavelength sequences of high-resolution solar photospheric images captured over several hours and restored by use of phase diversity, *Astronomy & Astrophysics* **377**, 1128–1135.

Löfdahl, M. G., Berger, T. E., Shine, R. A. and Title, A. M.: 1998, Preparation of a dual wavelength sequence of high-resolution solar photospheric images using phase diversity, *Astrophysical Journal* **495**, 965–972.

Mjolsness, E. and Decoste, D.: 2001, Machine learning for science: State of the art and future prospects, *Science* **293**, 2051–2055.

Nitta, N. V. and Hudson, H. S.: 2001, Recurrent flare/cme events from an emerging flux region, *GRL* **28**, 3801–3804.

Pacheco, P. S.: 1997, *Parallel Programming with MPI*, Morgan Kaufman.

Paxman, R. G. and Seldin, J. H.: 1994, Fine-resolution imaging of solar features using phase-diverse speckle imaging, *in* R. R. Radick (ed.), *Real-Time and Post-Facto Solar Image Correction*, National Solar Observatory. Proceedings of the Thirteenth Sacramento Peak Summer Workshop, 1992.

Paxman, R. G., Schulz, T. J. and Fienup, J. R.: 1992a, Joint estimation of object and aberrations by using phase diversity, *Journal of the Optical Society of America A* **9**(7), 1072–1085.

Paxman, R. G., Schulz, T. J. and Fienup, J. R.: 1992b, Phase-diverse speckle interferometry, *Signal Recovery and Synthesis IV*, Vol. 11 of *Technical Digest Series*, Optical Society of America, pp. 5–7.

Paxman, R. G., Seldin, J. H., Löfdahl, M. G., Scharmer, G. B. and Keller, C. U.: 1996, Evaluation of phase-diversity techniques for solar-image restoration, *Astrophysical Journal* **466**, 1087–1099.

Roddier, N.: 1990, Atmospheric wavefront simulation using Zernike polynomials, *Optical Engineering* **29**(10), 1174–1180.

Scharmer, G. B., Gudiksen, B. V., Kiselman, D., Löfdahl, M. G. and Rouppe van der Voort, L. H. M.: 2002, Dark cores in sunspot penumbral filaments, *Nature* **420**, 151–153.

Schulz, T. J.: 1993, Multi-frame blind deconvolution of astronomical images, *Journal of the Optical Society of America A* **10**, 1064–1073.

Schulz, T. J., Stribling, B. E. and Miller, J. J.: 1997, Multiframe blind deconvolution with real data: Imagery of the Hubble Space Telescope, *Optics Express* **1**(11), 355–362.

Schwenker, F., Kestler, H. and Palm, G.: 2000, Radial-basis-function networks: learning and applications, *Proc. Fourth International Conference on Knowledge-Based Intelligent Engineering System and Allied Technology*, pp. 33–43.

Seldin, J. H. and Paxman, R. G.: 1994, Phase-diverse speckle reconstruction of solar data, *in* T. J. Schultz and D. L. Snyder (eds), *Image reconstruction and restoration*, Vol. 2302 of *Proc. SPIE*.

Shih, F.Y. Moh, J. and Chang, F.: 1992, A new art-based neural architecture for pattern classification and image enhancement without prior knowledge, *Pattern Recognition* **25**, 533–542.

Snir, M., Otto, S. W., Huss-Lederman, S., Walker, D. and Dongarra, J.: 1998, *MPI - The Complete Reference: Volume 1, The MPI CORE, 2nd Edition*, MIT Press, Cambridge, MA.

Tanaka, K.: 1991, Studies on a very flare-active delta group - peculiar delta spot evolution and inferred subsurface magnetic rope structure, *Solar Physics* **136**, 133–149.

Tritschler, A. and Schmidt, W.: 2002, Sunspot photometry with phase diversity. I. Methods and global sunspot parameters, *Astronomy & Astrophysics* **382**, 1093–1105.

Turmon, M., Pap, J. M. and Mukhtar, S.: 2002, Statistical pattern recognition for labeling solar active regions: Application to soho/mdi imagery, *Astrophysical Journal* **568**, 396–407.

Tyler, D. W., Ford, S. D., Hunt, B. R., Paxman, R. G., Roggeman, M. C., Roundtree, J. C., Schulz, T. J., Schulze, K. J., Seldin, J. H., Sheppard, G. G., Stribling, B. E., Van Kampen, W. C. and Welsh, B. M.: 1998, Comparison of image reconstruction algorithms using adaptive optics instrumentation, *in* D. Bonaccini and R. K. Tyson (eds), *Adaptive Optical System Technologies*, Vol. 3353 of *Proc. SPIE*, pp. 160–171.

Valiant, L. G.: 1990, A bridging model for parallel computation, *Comm. of the ACM* **33**(8), 103–111.

van Ballegooijen, A. A., Nisenson, P., Noyes, R. W., Löfdahl, M. G., Stein, R. F., Nordlund, Å. and Krishnakumar, V.: 1998, Dynamics of magnetic flux elements in the solar photosphere, *Astrophysical Journal* **509**, 435–447.

Van Kampen, W. C. and Paxman, R. G.: 1998, Multi-frame blind deconvolution of infinite-extent objects, *in* L. R. Bissonnette (ed.), *Propagation and Imaging through the Atmosphere II*, Vol. 3433 of *Proc. SPIE*, pp. 296–307.

Vapnik, V.: 1995, *The Nature of Statistical Learning Theory*, Springer-Verlag, New York.

Vogel, C. R., Chan, T. F. and Plemmons, R. J.: 1998, Fast algorithms for phase diversity-based blind deconvolution, *in* D. Bonaccini and R. K. Tyson (eds), *Adaptive Optical System Technologies*, Vol. 3353 of *Proc. SPIE*.

von der Lühe: 1993, Speckle imaging of solar small structure: I. Methods, *Astronomy & Astrophysics* **268**, 374–390.

von der Lühe: 1994, Speckle imaging of solar small structure: II. Study of small scale structure in active regions, *Astronomy & Astrophysics* **281**(3), 889–910.

Wang, H., Tang, F., Zirin, H. and G., A.: 1991, Motions, fields, and flares in the 1989 march active region, *Astrophysical Journal* **380**, 282–286.

Zirin, H. and Wang, H.: 1993, Narrow lanes of transverse magnetic field in sunspots, *Nature* **363**, 426–428.