Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia Computer Science 9 (2012) 1159 – 1166

Procedia
Computer Science

Science
Procedia Computer Science 00 (2009) 000–000
A

www.elsevier.com/locate/procedia

International Conference on Computational Science, ICCS 2012

Adaptive optical sensing in an object tracking DDDAS

Anthony Vodacek[a,]*, John P. Kerekes[a], Matthew J. Hoffman[b]

[a]Chester F. Carlson Center for Imaging Science, 54 Lomb Memorial Dr., Rochester, NY, 14623, USA [b]School of Mathematical Sciences, Lomb Memorial Dr., Rochester, NY, 14623, USA

Abstract

The generalized optical remote sensing tracking problem for an object moving in a dynamic urban environment is complex. Two emerging capabilities that can help solve this problem are adaptive multimodal sensing and modeling with data assimilation. Adaptive multimodal sensing describes sensor hardware systems that can be rapidly reconfigured to collect the appropriate data as needed. Imaging of a moving target implies some ability to forecast where to image next so as to keep the object in the scene. Forecasts require models and to help solve this prediction problem, data assimilation techniques can be applied to update executing models with sensor data and thereby dynamically minimize forecast errors. The direct combination of these two capabilities is powerful but does not answer the questions of how or when to change the imaging modality. The Dynamic Data-Driven Applications Systems (DDDAS) paradigm is well-suited for solving this problem, where sensing must be adaptive to a complex changing environment and where the prediction of object movement and its interaction with the environment will enhance the ability of the sensing system to stay focused on the object of interest. Here we described our work on the creation of a modeling system for optical tracking in complex environments, with a focus on integrating an adaptive imaging sensor within the system framework.

Keywords: Dynamic Data Driven Application Systems; DDDAS; remote sensing; object tracking; adaptive sensing

1. Introduction

The general problem of remote airborne object tracking is to detect, identify, and track vehicles and pedestrians of interest in an urban environment. Constant surveillance through video imagery interpreted by human analysts is successful in this task, but is resource intensive and limited by electro-optical (EO) visible and infrared (IR) imaging phenomenology. For example, these systems require high spatial resolution to perform, which limits the achievable field of view and ground coverage. A cluttered urban environment is also particularly challenging due to the presence of man made structures casting shadows during the day and serving as confounding thermal sources for infrared systems operating at night. It is generally recognized that this challenging problem can be better addressed through the use of multiple modalities, such as multi- or hyperspectral and polarization imaging (MSI, HSI, and PI), which bring to bear additional phenomenological features to enhance robustness across varying environments.

* Corresponding author. Tel.: +1-585-475-7816; fax: +1-585-475-5988. *E-mail address*: vodacek@cis.rit.edu.

doi:10.1016/j.procs.2012.04.125

However, just using these multiple modalities in parallel does add to data volumes and system complexities that are generally unnecessary and not an optimal use of sensors.

Two emerging capabilities that can help solve this general problem are adaptive multimodal sensing and modeling with data assimilation. Adaptive multimodal sensing describes sensor hardware systems that can be rapidly reconfigured to collect the appropriate data, which will change over time. This adaptive sensing capability can drastically reduce data volumes by enabling the collection of only the most appropriate data given the time and space-dependent conditions for imaging. However, imaging of a moving target also implies some ability to adjust the imaging to object movement, that is, to forecast where to image next so as to keep the object in the scene and to simultaneously adjust the detection algorithm. Forecasts require models and to help solve this prediction problem, data assimilation techniques can be applied to update executing models with sensor data and thereby dynamically minimize forecast errors. The direct combination of these two capabilities is powerful but does not answer the questions of how or when to change the imaging modality, nor how to process the data to most efficiently identify and track the object of interest. Further, while data assimilation can provide an optimized answer for a given set of input data, a more accurate answer might be found with a different data set.

A comprehensive solution that integrates and expands on these imaging and data assimilation capabilities exists in the Dynamic Data Driven Applications Systems (DDDAS) paradigm. In this paradigm an executing application updates using data assimilation techniques and this new solution gives direction to an adaptive sensing system to collect a new set of data that may specifically reduce forecast uncertainty. Here we describe our work to develop such a system with a focus on the integration of the required image processing and analysis necessary to incorporate the adaptive imaging sensor into the DDDAS

2. Approach

Our general framework for this work is independent of the particular application and consists of the three main functions: data collection, forecasting, and sensor tasking; along with the three interfaces between these functions: the model objective function, the data needs assessment, and the modality and acquisition requirements (Figure 1). The six components represent the focus areas of research necessary for the development of the DDDAS. The following sections describe further detail about each of the six components.

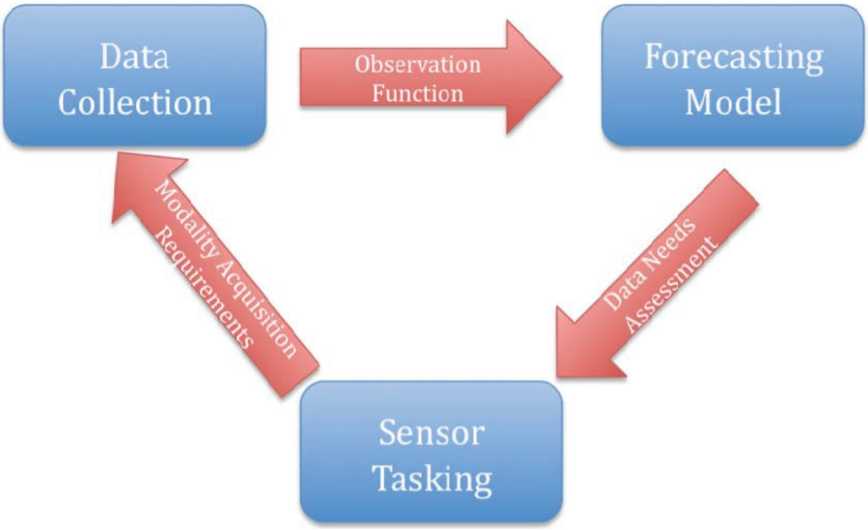

Figure 1. General modeling framework for the system.

2.1. Data Collection

Many optical imaging systems are currently used to track vehicles or dismounts, with the majority being singleband, single-modality systems. These include panchromatic visible (EO) or midwave infrared (IR) sensors. However, there are situations where bringing multiple modalities to bear on the problem would enhance the ability to maintain track and reduce false or spurious tracks. For example, spectral imaging can key on unique spectral reflectance characteristics of the targets of interest, as well as use the additional information from measurements of the background. Also, it is well known polarimetric imaging can enhance the contrast of man-made versus natural backgrounds due to the frequent polarizing specular reflections off of metal or glass components of a vehicle.

Through ongoing and related research on multi-modal imaging, we have been developing detailed models of adaptive multimodal sensors. One such sensor concept is shown in Figure 2 as an array of 2x2 superpixels, each of which contains a tunable single-pixel Fabry Perot spectrometer [1], as well as three wire-grid polarizers oriented such to collect polarization intensity imagery [2,3]. On the right of Figure 2 is one approach to providing the vertical motion on a pixel-by-pixel basis for the Fabry Perot devices with semi-transparent mirrors on top and bottom, and place directly over a detector. This sensor concept allows collection of tunable spectral and Stokes vector imagery for each pixel in the field-of-view.

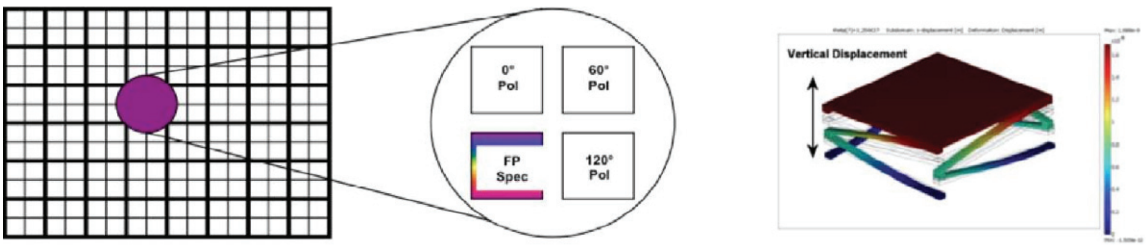

Figure 2. Concept for integrated MEMS single-pixel spectrometer and imaging polarimeter.

The flexibility of collecting different wavelengths across the array, along with selecting different wavelengths versus time for a given pixel, creates the opportunity for adaptive sensing driven by the application task at hand and derived data requirements. Also, given the relatively limited field of view of a single focal plane system, there is the opportunity for pointing or commanded data acquisition within an area of reconnaissance. This adaptive sensing offers the promise of optimally collecting only the necessary data (thus optimizing the allocation of sensor assets as well as reducing data communication and storage requirements) but requires sophisticated automated intelligence to task properly. Thus, the adaptive multimodal sensor is a prime candidate for use as part of a DDDAS system.

Another version of an adaptive sensor that was studied as part of a previous PhD thesis at RIT [4] combines polarization imaging with multiple pixel spectrometry. This instrument is a conceptual follow-on to the existing RIT Multi-Object Spectrometer (RITMOS) built at RIT [5]. In this case, the imaging channel has the same 2x2 superpixel design with 3 wire-grid polarizers as above, but the Fabry-Perot spectrometer is replaced with a simple digital micromirror. The micromirror can be selectively commanded to reflect light for that spatial pixel through the spectrometry channel to be dispersed across one dimension of a two-dimensional focal plane. While only one mirror can be used along that dimension, mirrors corresponding to the other dimension can reflect light from other parts of the scene into the spectrometry channel, thereby achieving adaptive multi-object spectrometry. These instrument models will serve as the starting point for simulated data collections as part of the DDDAS. In addition, thermal infrared imagers will be implemented as we add modalities to the study.

To support the development of scenarios for tracking in a complex environment, we will use imagery generated by the Digital Imaging and Remote Sensing Image Generation (DIRSIG) model. This first-principles synthetic image generation model can render very realistic image, EO, multispectral, hyperspectral, IR, polarimetric, lidar, and SAR imagery to support sensor design trade studies and to generate image sets for algorithm testing [5-8]. This modular code performs specific tasks such as predicting bi-directional reflectance functions (BRDF), computing time and material dependent surface temperature values, incorporating the atmospheric contribution (MODTRAN), and computing the dynamic viewing geometries of scanning instruments on moving platforms.

Over the past few years, the DIRSIG model has been used increasingly as a hardware simulator, allowing the user to create a very detailed description of the sensing platform. These features allow the users to simulate the raw data produced by the imaging system, which can then be used to exercise the processing pipelines used to generate final image products (Figures 3 and 4). Recently, DIRSIG has been integrated with a traffic mobility simulator SUMO (Simulation of Urban Mobility) to create dynamic moving imagery [9,10]. We plan to build on these scenarios and include not only moving vehicles but also moving pedestrians, together with simulated imagery collected at varying resolutions and from multiple platforms including the multimodal sensor. Thus, DIRSIG lies at the core of our capability for developing and testing a system that can operate with a variety of imaging modalities.

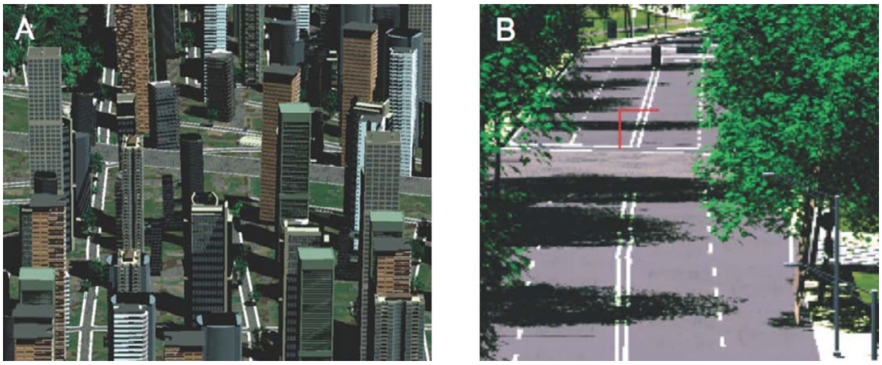

Figure 3. DIRSIG simulations of urban scenes. The rendering is RGB, but the objects have full spectral properties. A. Oblique overhead view of urban center with shadows and obscurations. B. Street level view showing clutter such as vehicles, road paint, light poles, signs, stop lights, and trees.

Figure 4. DIRSIG simulation of a desert industrial scene. The rendering is RGB, but the objects have full spectral properties.

2.2. Observation Function

The observation function interface reconciles new data with the model state variables of the forecast model in the data assimilation step. In prior work with remote sensing data we considered two possible approaches to implementing the observation function step [11]. In one case the approach is inversion of the remote sensing image data through an algorithm that classifies or otherwise results in a data product that describes the environment, such as temperature or a class map. This output is compared directly to the corresponding variable in the forecast model state in the data assimilation step. This approach is heavily reliant on signal processing. In the second approach the model state variables are used in a forward model to create a synthetic observation, which is compared directly to the remotely sensed image data. The forward modeling approach has advantages in maintaining physical information dependent on the 3-dimensional nature of the field of regard. In contrast, the output of most inversion algorithms for remotely sensed image data produce mapped 2-dimensional output with a corresponding loss of 3dimensional information. For this project we use the forward modeling approach and leverage the synthetic image generation capabilities of DIRSIG in creating the observation function similar to our previous work [12].

2.3. Forecasting Model

Object tracking generally uses particle filtering [13,14]. We will assess a variety of particle filter models for suitability within our framework. Particle filters (PFs) are one method of Monte Carlo state estimation. They share many features with ensemble Kalman filters, most notably that both use an ensemble of model realizations (called particles) that describe the PDF of the state. In a Kalman filter the PDF is assumed to be Gaussian and the ensemble is used to parameterize the distribution, while no assumption of the PDF is made for PFs and the distribution is approximated by some parameterization such as a sum of delta functions or sum of Gaussians, where N is the number of particles used.

As with all data assimilation, the PF combines observations with a prior estimate of the state by applying Bayes' rule to create a posterior estimate of the state and its associated PDF. This gives new weights to all of the particles and writes the PDF of a state given a set of observations. These new particles are then moved forward in time by applying a forecast model. At this new time, new observations are then used to improve the state estimate. For many target tracking applications, a very simple random walk model is used to update the particles. Here, we will utilize sensor and previously known information about road networks and other infrastructure to develop a more sophisticated model for the motion of the target. Once applied to all of the particles, the result is an estimate of the PDF of the signal emitted by the target. In this work, we will incorporate this information as another component of the utility function for collecting hyper spectral data. In this way, when the particle filter PDF is telling us that we have reasonable certain knowledge of the location of the target, we can continue observing with less expensive methods. One the other hand, when the filter tells us that we have greater uncertainty in the location of the target, the hyperspectral sensor can turned on to collapse the PDF.

Given an appropriate forecast model for the target tracking application, the forecast is the process of estimating the kinematic state of single or multiple, agile ground vehicles or pedestrians in the presence of clutter, dropped measurements, confusable vehicles, environmental occlusion, and ambiguous movement. A critical phase in the tracking process is the association of new measurements with existing tracks. The tracking system can employ various high-level association constructs to allow for statistics-based gating, multidimensional assignment, and deferred decision-making within the forecast decision. However, the fundamental cost $C_{i,j}$ to associate a track i with a measurement j is the key, as shown below [4].

$$C_{i,j} = \mu_K C^K_{i,j} + \Sigma \mu_F C^F_{i,j}$$

Here, $C^K_{i,j}$ *is a normalized kinematic cost based on the Mahalanobis distance, and* $C^F_{i,j}$ are likewise normalized costs based on statistical distances in an n-dimensional feature-space. The weighting terms μ establish the relative importance of the kinematic K and feature association F costs. The critical importance of this cost function is that it is a mechanism for incorporating multi-modal information. For example, it is well known that HSI instruments provide high-saliency feature measurements for many classes of ground vehicles. Hence, an HSI feature-aided tracking system has the potential to more accurately associate measurements with and subsequently perform with longer overall track life and higher track purity metrics. These assumptions are predicated on the availability of feature data, i.e., full spectral information, for both measurements and track state. While many realizable instruments always collect full HSI information throughout their fields of view, there is generally a design-time tradeoff such as ground sample distance or scan rate that makes tracking difficult. Our approach focuses on adaptive modality selection of an instrument that collects high-rate panchromatic or polarization imagery for the sake of tracking, but allows some pixels to collect HSI data as required. Thus, the cost function provides a means to make the critical step of a true DDDAS where the examination of the model state leads to the near real-time (rather than design-time) trade-off assessment of when, where, and what data to collect constrained to the capabilities of the adaptive sensor.

2.4. Data Needs Assessment

For the generalized tracking framework, the data needs assessment interface has two critical components that are at the heart of the DDDAS paradigm. The first data needs assessment component is an assessment of the errors in the model state or the forecast. In the tracking case these errors might include errors in detection, in time, or in space. The error in detection case refers to the target object itself – has the tracking system managed to maintain focus on the target of interest in the presence of other objects that have similar features or when temporary obscurations of the target have occurred? The errors in time are critical during obscurations – how well does the forecast model predict when the target will return to view? The errors in space include all standard errors related to the spatial image – primarily errors in georegistration and errors in target detection or classification that assign incorrect target or background pixels. Methods for assessing all of the above errors are known and we will apply these to our problem, although current work continues on understanding data assimilation errors in particular [15,16].

The second data needs assessment component is unique to our adaptive system. In the case of an adaptive sensing system, an assessment has to be made of the movement of the target from one observation scenario where a given imaging modality provides the best tracking to another observation scenario where a different imaging modality would provide better tracking. Here, the critical data assessment outcome is to make the decision to change the imaging modality. For example, consider the case of a tracked object that moves from an open area with few obstructions into an area with taller buildings casting shadows. A sensor operating in the reflective portion of the spectrum will suddenly have compromised image quality due to loss of signal to noise in shadows. Another example might be a vehicle moving from a paved road surface to an unpaved surface with aerial dust raised by the background traffic producing an obscuration. What are the consequences of these imaging scenarios with regard to imaging modality? A valid system should be capable of addressing these types of scenarios as they develop and tasking a different imaging modality that will better track the target of interest.

Successfully demonstrating this capability can leverage information already in hand and information collected at the time of the tracking exercise. Information already in hand may include static data in a GIS, meteorological observations, and low spatial resolution multispectral images. Or we can consider data from the tracking sensor itself, which could be particularly effective for a RITMOS type sensor [5]. Here, the multi-object spectrometer could be used to assess the background just as readily as it could assess the target signature. The analysis of the observation scenario can be considered a traditional image classification effort to describe the background that the target moves through.

A variety of current research at RIT is devoted to understanding the data content of MSI or HSI images and designing classification methodologies applicable to complex urban scenes [17,18]. One approach that is very applicable to our problem is the use of tiled images to reduce computation while focusing mainly on the parts of the background where the target is forecasted to move toward [19]. Our other applicable work includes detection algorithms with background suppression [20] and characterization of spectral spaces and change detection methods based on graph theory and implemented with tiling [21,22].

2.5. Sensor Tasking

Based on the data needs assessment derived from the system application model, a mechanism is required to determine how best to task the sensor to collect the necessary data. This functional block in the DDDAS needs to prioritize and rank data acquisitions across modalities, wavelengths, and locations within the focal plane and the field-of-regard. One way to pursue this optimization is through the use of a Sensor Resource Manager (SRM) using various quantifiable metrics to obtain the sensor allocation priorities.

Previous research [4,23] has explored the use of utility functions in the context of metrics for hyperspectral pixel data collection. As described in [4], the utility $U_{ij}(t)$ of collecting hyperspectral data at pixel *ij* at time *t* can be defined as:

$$U_{ij}(t) = C^D U^D(t) + C^N U^N(t) + C^A U^A(t) + C^M U^M + C^T U^T(t)$$

such that

$$U^{\Phi}_{ij}(t) \in [0,1], \;\; \Phi \in \{D,N,A,M,T\}, \;\; \Sigma C^{\Phi} = 1, \;\; C^{\Phi} \geq 0 \, \forall \Phi$$

The values of *C* are the relative importance or weighting of the different components of utility, U, where the components are defined as:

- D -Default value that every target of interest receives which gradually decreases towards 0 as we consider pixels farther from the predicted location of the target track.
- N -New model utility which is a function of the appearance of new or reacquired targets that need to be sampled in order to build a target feature model.
- A -Association utility defined for closely spaced targets where track state and the related uncertainty provide a measure of association ambiguity.
- M -Missed measurement utility which is a function of the number of missed detections for the kinematic tracker due to occlusion or shadow.
- T -Model age which is a function of the time since the last spectral model measurement was incorporated.

Previous research has used genetic algorithms with representative training data to obtain values for the component weighting coefficients, *C* [23].

2.6. Modality and Acquisition Requirements

An interface is necessary between the sensor tasking SRM and the actual sensor system to properly execute the prioritized data acquisitions within the DDDAS update cycle. This interface will have to take the prioritized sensor data needs and develop a sequential set of commands for sensor modality, sensor settings, platform location, platform orientation, and coverage. For example, the sensor tasking module may say polarization imagery is required from a zenith view angle of 30º and azimuth of 120º spanning an area of 1 square kilometer centered on a specific location. This execution of this requirement must take into account the current position of the sensor and the capabilities of the platform to determine the feasibility of acquiring the data within the necessary time frame. If not feasible, then the interface must feed that information back to the sensor tasking module and consider the next priority. Alternatively, it may be that the highest priority sensor tasking is only feasible after satisfying a lower priority tasking. Automated algorithms must be developed to consider and arbitrate between these various task implementation scenarios. General approaches for this will be explored, along with specific implementations for selected demonstrations. One approach that will be investigated is the adaption of previous multisensor tasking research [24] to our research domain including multiple views collected by a single sensor.

3. Summary

The general framework for object tracking described above, with its modules and interfaces provides a logical description of the necessary work required to construct an adaptive tracking system. Previous and ongoing work has addressed many aspects of the modules and interfaces, setting the stage for assembly and demonstration of a functioning framework that then can be tested under increasingly realistic and complex tracking scenarios by leveraging DIRSIG simulations.

The anticipated outcomes of the proposed work are centered on a creating a generalized modeling system for tracking in complex environments. In overview, the work plan will specify effort on designing data processing algorithms appropriate to adaptive sensors and creating a method for feedback control from the executing application to the adaptive sensor. These development tasks will leverage the DIRSIG model rather than relying on field data to demonstrate the ability of the modeling system to control adaptive imaging within a dynamic synthetic image dataset.

The flexibility of collecting different wavelengths across the array, along with selecting different wavelengths versus time for a given pixel, creates the opportunity for adaptive sensing driven by the application task at hand and derived data requirements. Also, given the relatively limited field of view of a single focal plane system, there is the opportunity for pointing or commanded data acquisition within an area of reconnaissance. This adaptive sensing offers the promise of optimally collecting only the necessary data, with direction from the model, in a truly DDDAS fashion.

Acknowledgements

This work is supported by AFOSR grants FA9550-11-1-0348 and FA9550-08-1-0028.

References

1. A. Rivas, J. Kerekes, and A. Raisanen. 2011. Tunable Single Pixel MEMS Fabry-Perot Interferometer. Proc. of the OSA Topical Meeting on Computational Optical Sensing and Imaging, Toronto, Canada.
2. A. Raisanen, M. Presnar, Z. Ninkov, K. Fourspring, L. Meng, and J. Kerekes. Simulation of Practical Single-Pixel Wire Grid Polarizers for Superpixel Stokes Vector Imaging Arrays. Submitted to SPIE J. Micro/Nanolithography, MEMS and MOEMS.
3. L. Meng and J. Kerekes, Appl. Opt. 50 (2011) 1925-1932.
4. M.D. Presnar, 2010. Modeling and Simulation of Adaptive Multimodal Optical Sensors for Target Tracking in the Visible to Near Infrared. Ph.D. Dissertation, Rochester Institute of Technology. 2010.
5. R.D. Meyer, K.J. Kearney, Z. Ninkov, C.T. Cotton, P. Hammond, and B.D. Statt, SPIE 5492 (2004) 200–219.
6. S.D.Brown, , N.J. Sanders, A.A. Goodenough, and M. Gartley, SPIE 8048 (2011) 804-814.
7. R.R. Burton, J.R. Schott, and S.D. Brown, SPIE 6214 (2002) 144-155.
8. M. Gartley, A. Goodenough, S. Brown, R.P. Kauffman, SPIE 7699 (2010) 76990N.
9. M.D. Presnar, J.P. Kerekes, D.R. Pogorzala. 2010. Proc. 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS) (2010).
10. M. Presnar, A. Raisanen, D. Pogorzala, J. Kerekes, and A. Rice, SPIE 7672 (2010) 76720T.
11. Z. Wang, A. Vodacek, and J.L. Coen, Int. J. of Wildland Fire. 18 (2009) 302-309.
12. J.A. Mandel, L.S. Bennethum, J.D. Beezley, J.L. Coen, C.C. Douglas, M. Kim, and A. Vodacek, Math. Comp. Simul. 79 (2008) 584-606.
13. K. Nummiaro, E. Koller-Meier, and L. Van Gool, LNCS 2449 (2002) 353–360.
14. Y. Rathi, N. Vaswani, and A. Tannenbaum, IEEE Trans. Image Proc. 16 (2007) 1370-1382.
15. P.F.J. Lermusiaux and A.R. Robinson, Mon. Weather Rev. 127 (1999) 1385-1407.
16. D. Zupanski and M. Zupanski, Mon. Weather Rev. 134 (2006)1337-1354.
17. J. Herweg, J. Kerekes, E. Ientilucci, and M. Eismann, SPIE 8040 (2011) 80400G.
18. A. Rice, J. Vasquez, M. Mendenhall, and J. Kerekes. 2009. Proc. 1st IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS) (2009).
19. K. Canham, A.A. Schlamm, B. Basener, and D.W. Messinger, SPIE 8048 (2011) 80841O.
20. Y. Li, A. Vodacek, R.L. Kremens, A. Ononye, and C. Tang, IEEE Trans. Geosci. Remote Sens. 43 (2005) 2115-2126.
21. D. Messinger and J.A. Albano. 2011. Proc. 3rd IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS) (2011).
22. J.A. Albano, D.W. Messinger, A. Schlamm, and B. Basener, SPIE 8048 (2011) 804809.
23. B.R. Secrest and J.R. Vasquez, SPIE 6964 (2008) 69640I.
24. R.T. Collins, A. Lipton, H. Fujiyoshi, and T. Kanade, Proc. IEEE, 89 (2001) 1456-1477.