AUTOMATIC GENERATION OF CHANGE INFORMATION FOR MULTITEMPORAL, MULTISPECTRAL IMAGERY

Morton J. Canty[1] **and Allan A. Nielsen**[2]

[1] *Institute for Chemistry and Dynamics of the Geosphere, Forschungszentrum Jülich, D-52425 Jülich, Germany, E-mail: m.canty@fz-juelich.de*

[2] *Informatics and Mathematical Modelling, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark, E-mail: aa@imm.dtu.dk*

ABSTRACT

A data-oriented and automatic approach to the extraction of change information from time series of multi- and hyperspectral imagery is presented . The method applies the iteratively re-weighted multivariate alteration detection (IR-MAD) algorithm to signal significant changes, recording the results in a highly compressed, georeferenced binary image. A prototype change extraction and archival system, implemented as an extension to the ENVI/IDL remote sensing data processing user interface, is illustrated using time series of LANDSAT TM and ASTER multispectral imagery over the Nevada Nuclear Test Site in the USA.

1. INTRODUCTION

The detection and classification of significant changes in image time series is one of the most important applications of remote sensing with earth observation satellites. Within the context of the Global Monitoring for Security and Stability (GMOSS) Network of Excellence, for example, anthropogenic changes are of particular interest: population movements, infrastructure modification, deviations from declared information, clandestine military activities, and the like. The monitoring process implies the acquisition and maintenance of a large database of satellite imagery from diverse sensors with different spatial, spectral, and temporal resolutions for the areas of interest. Manual examination of such a database for the identification of relevant changes will generally be impractical.

A data-oriented and fully automatic approach to the extraction of change information from multi- and hyperspectral imagery can be achieved with the *iteratively reweighted multivariate alteration detection* (IR-MAD) transformation [1, 2, 3, 4]. Based entirely on the second-order statistics of the no-change observations (which are extracted in an iterative scheme as described below), the image pixels may be labelled according to their change probabilities. The MAD transformation has been successfully applied to location of clandestine underground nuclear explosions, see [5].

By setting significance thresholds, binary images of significant changes or, alternatively, tables of the latitude/longitude of the changes, can be generated for all available pairs of co-referenced images without any human intervention whatsoever. Furthermore, since the IR-MAD transformation is invariant under linear, affine transformations of the pixel intensities of the images involved [1, 3], results are insensitive to instrument gain and (linear) atmospheric corrections to the input data. Prior processing for change detection is often unnecessary.

2. IR-MAD

The pixel intensities for two N-band multispectral images of the same scene acquired at different times t_1 and t_2 may be represented by random vectors $\boldsymbol{F}$ and $\boldsymbol{G}$, respectively. We can make a linear combinations of the intensities of the spectral bands for each acquisition time, creating scalar images characterized by the random variables $U = \boldsymbol{a}^\top \boldsymbol{F}$ and $V = \boldsymbol{b}^\top \boldsymbol{G}$ and then examine the difference $U - V$. This combines all of the change information into a single image, but one has of course still to choose the coefficients $\boldsymbol{a}$ and $\boldsymbol{b}$ in some suitable way. In [1] it is suggested that they be determined so that the correlation ρ between U and V is minimized subject to $\mathrm{var}(U) = \mathrm{var}(V) = 1$, implying that the resulting difference image $U - V$ will have maximum variance (maximum spread in its pixel intensities). Minimizing the correlation between the two linear combinations is achieved via standard canonical correlation analysis (CCA) and generates, through solution of a coupled eigenvalue problem, a sequence of transformed difference images

$$M_i = U_i - V_i, \quad i = 1 \ldots N, \tag{1}$$

referred to as the *MAD variates*. They have, by virtue of the chosen ordering of eigenvalues, successively decreasing variance. The MAD variates have

nice statistical properties which make them very useful for visualizing and analyzing change information [1, 4]. Thus for instance they are uncorrelated, with

$$\mathrm{cov}(M_i, M_j) = 0,\ i \neq j, \quad \mathrm{var}(M_i) = \sigma^2_{M_i} = 2(1-\rho_i), \tag{2}$$

where $\rho_i = \mathrm{corr}(U_i, V_i)$ and is determined by the square root of the ith eigenvalue.

If no physical reflectance changes have occurred in the scene, the MAD variates, being uncorrelated and nearly normally distributed, should obey a multivariate normal distribution with diagonal covariance matrix. Change observations would deviate more or less strongly from a normal distribution. In the presence of genuine change, we expect an improvement of the sensitivity of the MAD transformation if we place emphasis on establishing an increasingly better background of no change against which to detect change. This can be done in an iteration scheme in which observations are weighted by the probability of no change, as determined on the preceding iteration, when estimating the sample means and covariance matrices which determine the MAD variates for the next iteration [2]. The probability weights may be obtained by examining the MAD variates directly. Let the random variable Z represent the sum of the squares of the standardized MAD variates:

$$Z = \sum_{i=1}^{N} \left(\frac{M_i}{\sigma_{M_i}} \right)^2, \tag{3}$$

where σ_{M_i} is given by the last equality in Equations (2). Then, since the no-change observations are normally distributed and uncorrelated, their realizations should be approximately chi-square distributed with N degrees of freedom (distribution function $P_{\chi^2;N}(z)$). Change observations will have anomalously large values of Z. For each iteration, the observations can then be given weights determined by the chi-square distribution, namely

$$\Pr(\text{no change}) = 1 - P_{\chi^2;N}(z). \tag{4}$$

Thus Pr(no change) is the probability that a sample z drawn from the chi-square distribution could be that large or larger. A small z implies a large probability of no change. Other weighting schemes are possible.

3. IMAGERY AND PREPROCESSING

Since 1962, all nuclear tests in the USA have been underground and most of them have taken place at the Nevada Test Site (NTS). A moratorium on underground testing has been in effect since October, 1992. An exhaustive list of US nuclear tests from July 1945 through September 1992 has been published by the US Department of Energy [6] as well as by Springer et al. [7]. In order to test the proposed scheme, a series of Landsat TM and ASTER images covering all or portions of the NTS was acquired for the periods 1986-2003, see Table 1. A spatial/spectral subset of one of the LANDSAT images covering the Pahute Mesa test area is shown in Figure 1.

Table 1. Multispectral satellite imagery over the Nevada Test Site.

Date	Time (GMT)	Sensor
May 28, 1986	17:45:38	TM5
May 31, 1987	17:45:35	TM5
April 18, 1989	17:50:19	TM5
May 26, 1991	17:43:59	TM5
June 2, 2000	19:00:24	ASTER
October 1, 2001	18:52:58	ASTER
July 6, 2003	18:38:32	ASTER

Each image series (LANDSAT and ASTER) was co-registered by applying a contour matching algorithm [8] and using first-order polynomial, nearest-neighbor resampling. The RMS errors were less than 0.5 pixel. In addition, the six short wave infrared (SWIR) bands of the ASTER images were sharpened to the 15m ground resolution of the three visual near infrared (VNIR) bands with a wavelet fusion technique [9]. The processed ASTER images thus consisted of nine spectral bands each.

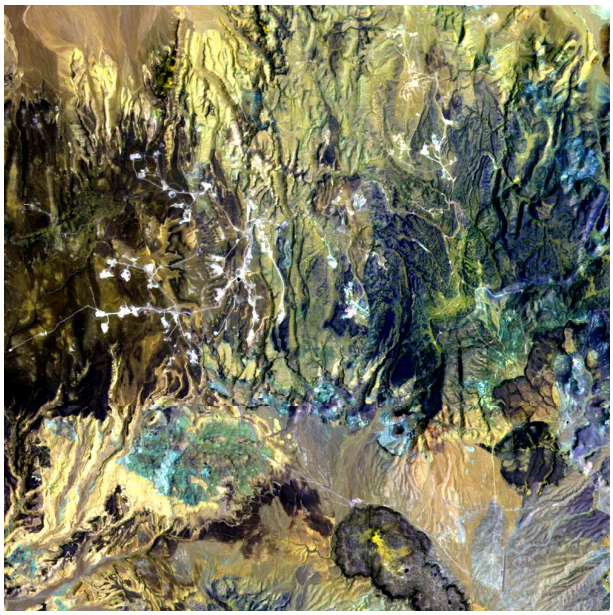

Figure 1. LANDSAT TM image over the Pahute Mesa area of the Nevada Test Site, acquired May 26, 1991. The three infrared bands 4, 5 and 7 are displayed in blue, green and red, respectively. The across- and along-track ground sample distance (GSD) is 30m.

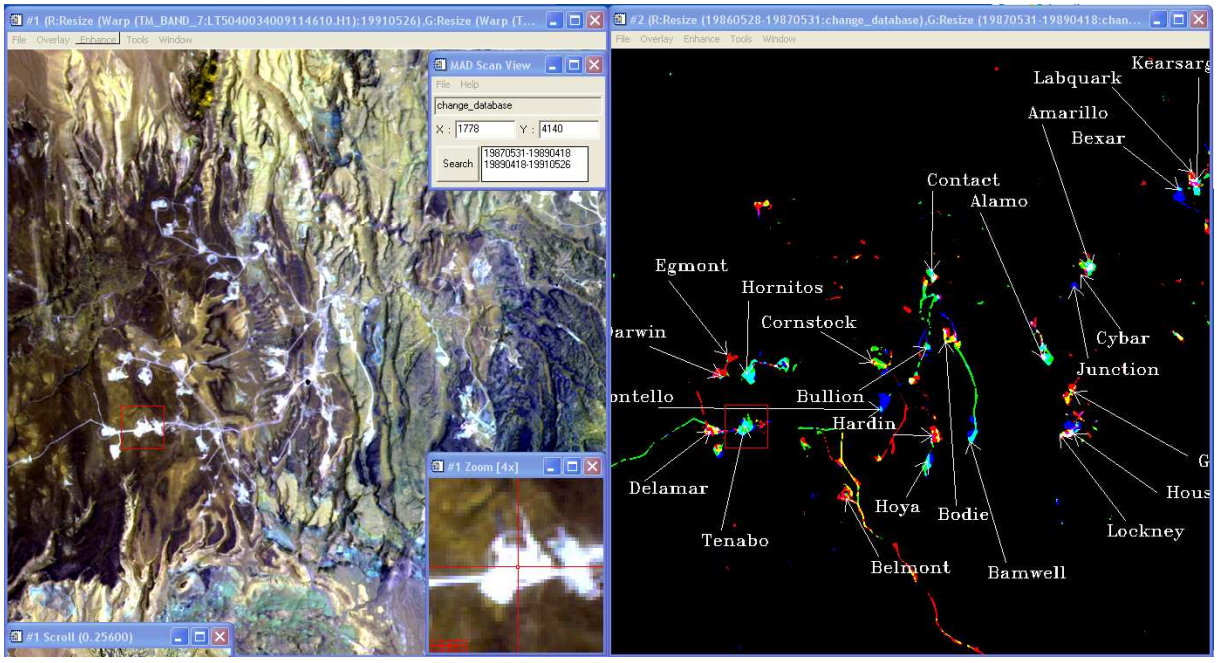

Figure 2. Screenshot showing a portion of the LANDSAT TM scene of Figure 1 (left) and corresponding change information over three time intervals as an RGB composite (right). The time intervals are: red: May 28, 1986 - May 31, 1987, green: May 31, 1987 - April 18, 1989, blue: April 18, 1989 - May 26, 1991. Changes occurring in two or three intervals have corresponding mixed colors(yellow, cyan, magenta or white). The locations and codenames of underground nuclear test explosions that took place in these intervals are marked. The top insert on the left is a widget (graphical interface) which displays the interval over which changes occurred at the chosen pixel location (bottom insert).

4. AUTOMATIC CHANGE EXTRACTION

In order to create a simple, compact database of changes in a time series of satellite images, the following strategy was adopted[1]:

1. A given directory is scanned to extract filenames of the co-registered images contained therein.
2. The images are sorted according to increasing acquisition time.
3. For each consecutive pair of acquisitions:
   (a) The IR-MAD algorithm is run to convergence.
   (b) The chi-square statistic, Equation (3), is determined at each pixel location.
   (c) A threshold is set for high probability of change and all pixels are labelled accordingly as change (1) or no-change (0).
   (d) A 3×3 median filter is run on the resulting binary array and the end result is added as a band to a multi-band image in standard ENVI format (constituting the database).
4. The change database is provided with the same georeferencing data as the co-registered images and stored on disk in compressed (GZIP) format.

[1]and implemented in a prototype IDL program running within the ENVI remote sensing graphical user interface.

4.1. Decision thresholds

After a single application of the CCA transformation, i.e., without iteration, the resulting MAD variates M_i given by Equation (1) are necessarily uncorrelated [1]. To the extent that they are also normally distributed, the statistic Z given in Equation (3) should be chi-square distributed with N degrees of freedom. After iteration to convergence, however, only the no-change observations determine the statistics for the CCA transformation, and the resulting IR-MAD variates, consisting as they do of both change and no-change pixels, will no longer be uncorrelated. Moreover, depending on the proportion of true change observations, their histograms may deviate more or less strongly from a Gaussian and even be multi-modal. For this reason the step 3(c) in the preceding strategy, i.e., setting a threshold on the chi-square image, is somewhat problematic. It nevertheless may be sufficient to choose a threshold based on the the assumption that *all* sums of squares of the standardized IR-MADs obey a chi-square distribution. That is, after iteration the quantities σ_{M_i}

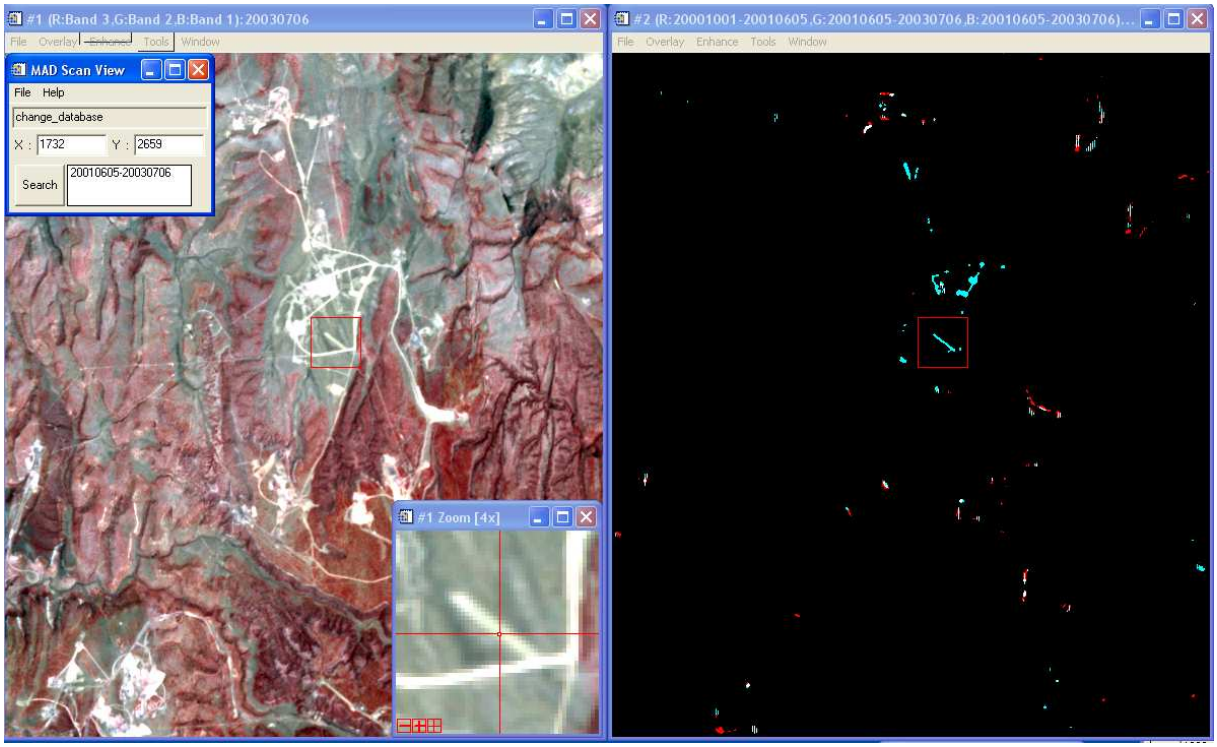

Figure 3. Similar to Figure 2 showing a portion of an ASTER scene over the Pahute Mesa region (left) and corresponding change information over two time intervals as an RGB composite (right). The time intervals are: red: June 2, 2000 - October 1, 2001, cyan: October 1, 2001 - July 6, 2003. No underground tests took place during these periods.

in Equation (3), hitherto determined according to (2) and corresponding to no-change observations only, are replaced by the standard deviations estimated from all of the IR-MAD variates. Then a threshold can be chosen to signify significant change. We chose the 99.9 percentile of the chi-square distribution.

4.2. Results

For part of the time series of Table 1, consisting of four 2000×2000-pixel LANDSAT scenes covering the entire Pahute Mesa and Yucca Flat areas of the NTS (3600 km^2), the compressed change database generated according to the scheme described above occupied about 80 kB of disk space. This is to be compared with the

$$4 \text{ images} \times 6 \text{ bands} \times 4,000,000 \text{ pixels} = 96MB$$

occupied by the image time series itself. Figure 2 shows the operation of a simple IDL widget (graphical interface) for querying the database for changes occurring at any given pixel location, here over the Pahute Mesa area. The widget can be invoked from the ENVI display menu of any georeferenced representation of the area of interest (e.g., a rasterized map). When accessed by the widget, the database is temporarily de-compressed and made available to ENVI until the widget is closed. This is illustrated on the right hand portion of the Figure, where codenames and locations of underground tests that took place have also been inserted manually. A similar screenshot is shown in Figure 3, this time for ASTER data. In this case, no underground tests took place during the time intervals spanned (moratorium). Significant changes in the vicinity of one of the previous test sites are nevertheless observed. Figure 4 similarly depicts a portion of the NTS LANDSAT change database over the Yucca Flat region, showing the recorded significant changes and the underground testing activities which took place over two time intervals in question. There is good temporal and spatial correspondence between recorded test locations and times and observed change signals, although many change signals are not directly associated with known underground tests events.

5. CONCLUSIONS

We have described an approach to extracting and archiving change information automatically from satellite imagery time series based on the IR-MAD algorithm. For arid/desert scenes in which significant anthropogenic changes are not masked or confused by seasonal changes in vegetation, the procedure works reliably and condenses the essential information on geographical locations and time intervals

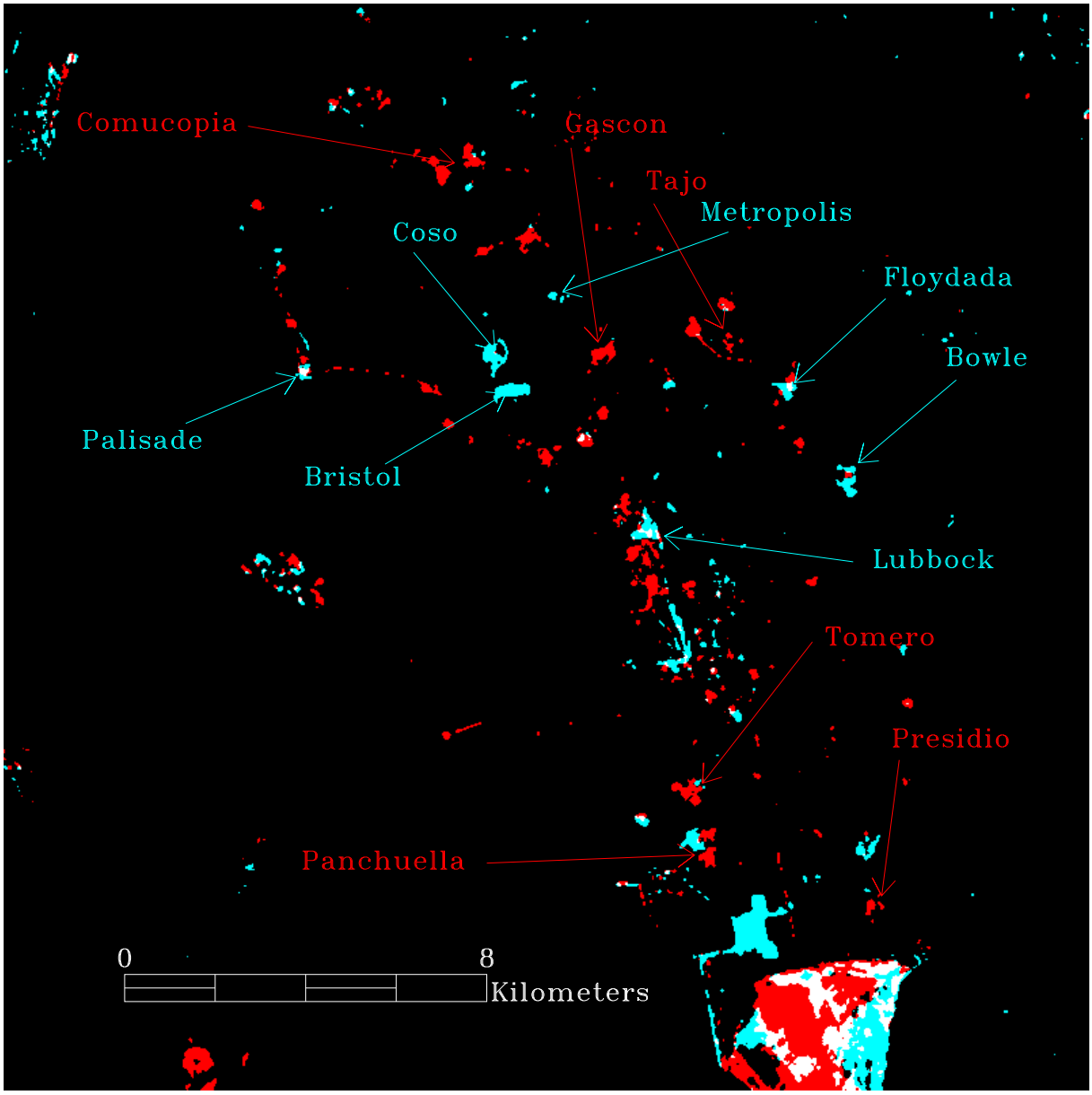

Figure 4. Color composite of a portion of a change database showing changes over the Yucca Flat area of the NTS between May 28, 1986 and May 31, 1987 (red) and between April 18, 1989 and May 26, 1991 (cyan). Regions where changes occurred in both intervals are white. The locations and codenames of underground nuclear test explosions that took place in the two intervals are marked and colored correspondingly. The large feature bottom left is a salt flat which is occasionally flooded.

of significant change into a tiny fraction of the storage space occupied by the source data.

REFERENCES

1. A. A. Nielsen, K. Conradsen, and J. J. Simpson. Multivariate alteration detection (MAD) and MAF post-processing in multispectral, bitemporal image data: New approaches to change detection studies. *Remote Sensing of Environment*, 64:1–19, 1998. Internet http://www.imm.dtu.dk/pubdb/p.php?1220.

2. A. A. Nielsen. The regularized iteratively reweighted MAD method for change detection in multi- and hyperspectral data. 2006. Accepted for *IEEE Transactions on Image Processing*. Internet http://www.imm.dtu.dk/pubdb/p.php?4695.

3. M. J. Canty, A. A. Nielsen, and M. Schmidt. Automatic radiometric normalization of multitemporal satellite imagery. *Remote Sensing of En-*

vironment, 91(3-4):441–451, June 2004. Internet http://www.imm.dtu.dk/pubdb/p.php?2815.

4. M. J. Canty and A. A. Nielsen. Visualization and unsupervised classification of changes in multispectral satellite imagery. *International Journal of Remote Sensing*, 27(18):3961–3975, 2006. Internet http://www.imm.dtu.dk/pubdb/p.php?3389.

5. M. J. Canty and J. Schlittenhardt. Satellite data used to locate site of 1998 Indian nuclear test. *Eos, Transactions, American Geophysical Union*, 82(3):25–29, 2001.

6. U. S. DOE. United States nuclear tests July 1945 through September 1992. U. S. Dept of Energy DOE/NV-209-REV 15, 2000.

7. D. L. Springer, G. A. Pawloski, J. L. Ricca, R. F. Rohrer, and D. K. Smith. Seismic source summary for all U.S. below-surface nuclear explosions. *Bulletin of the Seismological Society of America*, 92(5):1806–1840, 2002.

8. H. Li, B. S. Manjunath, and S. K. Mitra. A contour-based approach to multisensor image registration. *IEEE Transactions on Image Processing*, 4(3):320–334, 1995.

9. B. Aiazzi, L. Alparone, S. Baronti, and A. Garzelli. Context-driven fusion of high spatial and spectral resolution images based on oversampled multiresolution analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 40(10):2300–2312, 2002.