

Mixture Modeling of M51

Paul E. Johnson¹, David M. Waddill, Leisa K. Townsley, Timothy N. Titus, and Earl J. Spillar

Dept. of Physics and Astronomy, University of Wyoming, Laramie, WY 82071, USA

Abstract. One of the major problems with 3D spectroscopy is the adequate and meaningful display and analysis of large image cubes. We use linear mixture modeling to analyze multispectral CCD images of M51 in the visible and near-IR. We find that M51 can be modeled within the noise limits of the data as a linear combination of six components taken three at a time.

1. Introduction

Mixture modeling of remotely sensed images of the Earth and other solid-surfaced planets has only recently gained wide acceptance as the premier technique for extracting surface mineralogy from remotely sensed data (Adams *et al.* 1989, Pinet *et al.* 1993); we apply it to extragalactic images for the first time. In this method, pixels in a remotely sensed image are modeled as a linear combination of “pure” endmember spectra. Two of the major problems with application of mixture modeling to multispectral images are the convergence on a unique set of endmembers and the robust calculation of fractional abundance maps of these endmembers. We have developed a new mixture modeling methodology that we have applied to M51. Images were taken at seven wavelengths, between 0.4 and 1.6 μm , with visible and near-IR CCD cameras at the 2.3 m telescope of the Wyoming Infrared Observatory. The spatial resolution is 1.3 and 2.2'', for the visible and H-band images respectively.

2. Our Approach

We find these M51 images to contain six statistically significant endmembers. The problem with conventional mixture modeling is the application of least-squares fitting of six or so endmember spectra, simultaneously, to the spectrum of a pixel at seven wavelengths, given that the intensity at each wavelength is not independent of those at the other wavelengths. Our approach is unique and more robust compared to previous approaches in that:

¹Observatoire Midi-Pyrénées, 31400 Toulouse, France

1. We assume that each pixel is spectrally dominated by at most three endmembers. We then find the endmembers that fit the pixel spectra to within the estimated measurement error. Not every pixel is spectrally governed by the same three endmembers. We choose two or three endmembers at a time, model that part of the image that can be modeled satisfactorily by this endmember choice, and then iterate with another endmember set on the remaining image.
2. Endmember candidates are chosen from pixels that are spectrally extreme and are distributed over spatially contiguous areas in the image. We apply Principal Component Analysis (PCA) to the galaxy data cube to select these pixels.
3. A χ^2 analysis is used to evaluate the selection of endmembers at the end of each iteration. Further iterations are made until it is judged that remaining pixels are: foreground stars, pixels at the nucleus affected by image registration, cosmetic defects on the CCD, or are not spatially contiguous and are small enough in number to be consistent with pixels falling outside 2σ in a Gaussian distribution. The endmember selection is considered satisfactory and no further iterations are performed.
4. To examine the uniqueness of our solution the order of application of the endmember suites is permuted. Pixels that satisfactorily fit more than one endmember combination can be included in a pixel ambiguity map.
5. At the end of the analysis any endmembers that have been chosen for more than one iteration have multiple fractional abundance maps that need to be coadded. For M51, a red-star component has been chosen for all four iterations.
6. The resulting endmember spectra are calibrated and compared to template spectra of stars, clusters, and nebulae, from the literature.

The three primary advantages of this technique are: hyperspectral images can be reduced to a comprehensible display, linear modeling of the data into multiple components is allowed even when the number of components exceeds the number of bands, and the technique uses a minimum of preconceived ideas about the physical constituents of the image.

3. Analysis of M51

Images of M51 were made at 425, 500, 600, 700, 800, and 900 nm with ~ 100 nm bandpasses, and with a broadband H filter. These images were dark subtracted, sky corrected, and flattened using standard techniques. The visible data were sub-pixel coaligned to an accuracy of approximately $0.2''$ and smoothed to the lower resolution of the H-band data. Suspected foreground stars and pixels within $6''$ of the center of the nucleus were removed from the analysis. (Pixels near the nuclear brightness cusp potentially suffer adverse photometric effects from image misregistration.) The 425 and 500 nm filters were used to determine approximate B-V colors by calibrating five $8.9''$ diameter spots measured

by Schweizer (1976). Our 500 nm data have a readout noise limited surface brightness at V (3σ) of ≈ 26 magnitude/ \square'' .

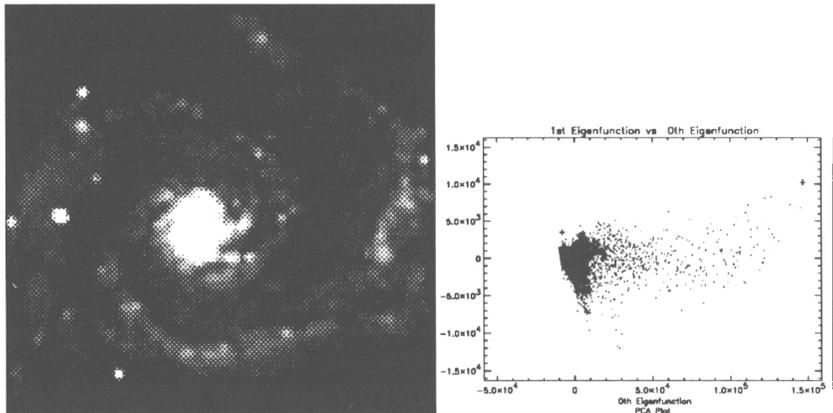


Figure 1. On the left is shown is a CCD image of M51 at 425 nm smoothed to a resolution of $2.2''/\text{pixel}$ with a field of view of $4.5' \times 4.2'$. On the right the PCA data cloud is plotted for the first two principal axes of variation. Crosses designate chosen endmembers: the upper left is sky, the lower left represents a bright patch of spiral arm, and the rightmost cross is a pixel near the nucleus.

Figure 1 shows an M51 image and its representation in PCA space (where each pixel undergoes a linear transformation into a new space defined by the principal axes of variation in the data). The spectrally most extreme pixels in the data are extrema in the PCA data cloud. These extrema are used as candidates for endmember pixels.

The entire image cube was first modeled as a linear combination of HII/star formation region, sky, and red stars. A noise spectrum is estimated for each pixel assuming that the data are limited by photon statistics of the galaxy plus CCD readout noise, in the visible, and of the background, at H-band. 560 out of 14,022 pixels fit within 2σ (Figure 2). The pixels that were satisfactorily fit were used to generate fractional abundance maps for each endmember (Figure 2). Only those pixels that fit our linear model to within 2σ were removed from the analysis. For the remaining iterations, sky and red stars were also used as endmembers. The third endmember was spiral arm, interarm, and opaque dust, for the second, third, and fourth iterations, respectively. After four iterations 13,416 pixels were successfully fit, 61 had residuals between 2σ and 3σ , while 562 had residuals greater than 3σ . Almost all of the latter can be attributed to foreground stars, cosmetic defects on the CCD (high dark-current pixels), and a corner of the field for which complete filter coverage was not obtained.

Endmembers have been labeled based on their brightness, spatial distribution, and B-V colors: sky, red star population ($B - V \approx 0.75$), spiral arm ($B - V \approx 0.41$), interarm ($B - V \approx 0.49$), HII/star formation region ($B - V \approx 0.24$), and opaque dust. We find that spiral arm, interarm, and opaque dust

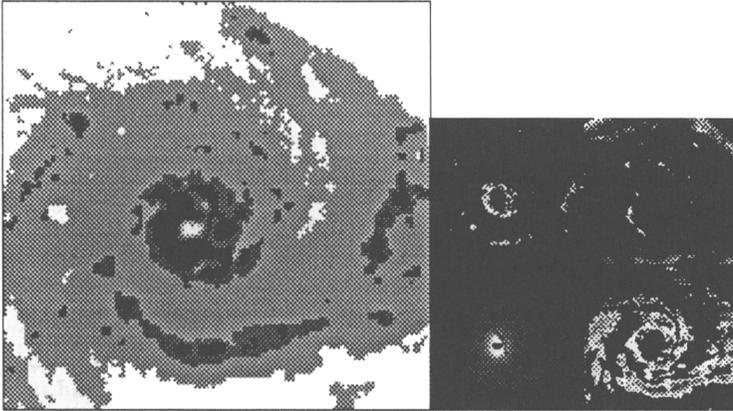


Figure 2. On the left is the residual map for the first iteration. Pixel grey level corresponds to goodness of fit. The darkest pixels are those outside 4σ , white pixels represent a fit within 2σ . On the right the fractional endmember abundance maps are shown (from top left, clockwise): HII/star formation region, opaque dust, spiral arm, and red star component. The lightest pixels correspond to the highest fractional abundances. The central $\sim 6''$ of the galaxy was not satisfactorily fit with any endmember combination.

have well-defined spiral patterns, whereas the red star component has a relatively smooth, elliptical distribution. HII/star formation regions are concentrated near the nucleus as well as being distributed along spiral arms.

Acknowledgments. D.W. was supported by a NASA Space Grant Graduate Fellowship given through the Wyoming Planetary and Space Science Center. P.J. was partially supported by fellowships from Centre National de la Recherche Scientifique (France) and Université Paul Sabatier. We are grateful to Pat Broos for writing the mosaicing software.

References

- Adams J.B., Smith, M.O., and Gillespie, A.R. 1989, Proc. IEEE Int. Geosci. Remote Sensing Symp., I, 16.
 Pinet, P.C., Chevrel, S.D., and Martin, P. 1993, Science, 260, 797.
 Schweizer, F. 1976, ApJS, 31, 313.