

Journal of  
**Applied Remote Sensing**

**Hyperspectral image analysis  
using artificial color**

Jian Fu  
H. John Caulfield  
Dongsheng Wu  
Wubishet Tadesse



# Hyperspectral image analysis using artificial color

Jian Fu,<sup>a</sup> H. John Caulfield,<sup>b</sup> Dongsheng Wu,<sup>c</sup> and Wubishet Tadesse<sup>d</sup>

<sup>a</sup> Alabama A & M University, Department of Computer Science, 4900 Meridian Street, Normal, AL 35762, USA

[jian.fu@aamu.edu](mailto:jian.fu@aamu.edu)

<sup>b</sup> Alabama A & M University Research Institute, Normal, AL 35762, USA

[John.Caulfield@cim.aamu.edu](mailto:John.Caulfield@cim.aamu.edu)

<sup>c</sup> University of Alabama in Huntsville, Department of Mathematical Sciences, Huntsville, AL 35899

USA [dongsheng.wu@uah.edu](mailto:dongsheng.wu@uah.edu)

<sup>d</sup> Alabama A & M University, Department of Natural Resources and Environmental Sciences, 4900 Meridian Street, Normal, AL 35762 USA

[wubishet.tadesse@aamu.edu](mailto:wubishet.tadesse@aamu.edu)

**Abstract.** By definition, HSC (HyperSpectral Camera) images are much richer in spectral data than, say, a COTS (Commercial-Off-The-Shelf) color camera. But data are not information. If we do the task right, useful information can be derived from the data in HSC images. Nature faced essentially the identical problem. The incident light is so complex spectrally that measuring it with high resolution would provide far more data than animals can handle in real time. Nature's solution was to do irreversible POCS (Projections Onto Convex Sets) to achieve huge reductions in data with minimal reduction in information. Thus we can arrange for our manmade systems to do what nature did – project the HSC image onto two or more broad, overlapping curves. The task we have undertaken in the last few years is to develop this idea that we call Artificial Color. What we report here is the use of the measured HSC image data projected onto two or three convex, overlapping, broad curves in analogy with the sensitivity curves of human cone cells. Testing two quite different HSC images in that manner produced the desired result: good discrimination or segmentation that can be done very simply and hence are likely to be doable in real time with specialized computers. Using POCS on the HSC data to reduce the processing complexity produced excellent discrimination in those two cases. For technical reasons discussed here, the figures of merit for the kind of pattern recognition we use is incommensurate with the figures of merit of conventional pattern recognition. We used some force fitting to make a comparison nevertheless, because it shows what is also obvious qualitatively. In our tasks our method works better.

**Keywords:** artificial color, margin setting, hyperspectral image, feature extraction, classification, projection onto convex sets.

## 1 INTRODUCTION

One of the main tasks of remote sensing is to analyze the received image in terms that the user needs (classes of objects, moisture of the soil, pollution in rivers, and the like. The HyperSpectral Image (HSI) contains such information implicitly. Analysis must make the desired information explicit in the final image.

Animals have to understand scenes with almost unimaginable spectral complexity and do so quite quickly, especially in view of how slow neurons are. In principle, animals could have done some sort of spectral analysis at each pixel, but did not. Instead, it projected that spectral information into two or more broad, spectrally-overlapping bands. They thus replace the actual spectrum with measures (discriminants) that relate to the complex spectra but are not

invertible. That is, the discriminants measured cannot be even crudely inverted to form a spectrum. It is a form of lossy bandwidth. Those discriminants are used to form more complex discriminants that we call colors. The set of spectra that give identical projections are called metamers. That set has infinitely many members. That is the price of irreversible loss of information. Accurate wavelength measurements are impossible this way, but that is not nature's goal. Finding food and not becoming food are more important to an animal than finding spectra. So nature seems quite content to use these few discriminants to make decisions about what to do next. So powerful is the color concept that, once it was "invented" perhaps 525 MYA [1], it has never been replaced. Indeed, it may well have triggered a predator-prey arms race that, in turn, triggered the Cambrian explosion [2].

The critical insights just summarized not only tell what color is but also provide a program technologists can emulate it – a program we have called Artificial Color [3-5].

Hyperspectral Imaging (HSI) too produces a vast overabundance of information: the reason nature adopted, spread, and maintained color as soon as it stumbled upon it [6].

A hyperspectral image can be considered as an image cube where the third dimension is represented by up to hundreds of contiguous spectral bands. A hyperspectral pixel is actually a column vector with dimensions equal to the number of spectral bands. Many conventional measures proposed in signal processing and pattern recognition can be used for this purpose. Several approaches for classification of hyperspectral pixels have been used such as some sort of minimum distance, maximum likelihood classifiers [7, 8], spectral signature matching [9, 10] and the spectral angle mapper [11]. To reduce the data volume, techniques for reducing the image dimensionality are often applied. Typically, the dimensionality of a hyperspectral image cube is reduced by applying a linear transformation [12], such as a Principal Components Analysis (PCA) and retaining only the significant components for further processing [13, 14]. The poor performance of PCA has been studied widely [15, 16]. Other approaches have been developed. In the orthogonal subspace projection approach, the data dimensionality is reduced to a user-prescribed level. Simultaneously, in the resulting images the presence of each signature of interest is selected [17, 18]. Various projection schemes have been compared [19]. These methods are developed to do spectral "unmixing" to decompose the spectrum in any pixel into a mixture of several known spectra. The discrete wavelet transform (DWT) has been utilized for dimensionality reduction of hyperspectral data and feature extraction [20]. They use the multiple range and low computational complexity of wavelets as a way to compute multistage discriminants which are shown to be quite useful in pattern recognition. Their analysis suggests that this approach has distinct advantages over other approaches such as PCA and Fourier transforms. Dutkiewicz et al have done extensive work on a scheme for browsing the data measured by Hyperspectral Imaging. Their vector-quantization-based compression serves that need well [21].

Projections Onto Convex Sets (POCS) is a widely used lossy compression method [22, 23]. If the convex sets are chosen carefully, the projections are often capable of giving great discrimination [24-26]. Artificial Color uses what we believe to be totally new kind of orthogonality none of the past work has used.

In this paper, Artificial Color is first proposed for and applied to hyperspectral image data dimensionality reduction and feature classification. Artificial Color irreversibly reduces the information content of the data cube but does it in a way that facilitates good discrimination. Compared to traditional classification methods such as minimum distance and maximum likelihood classifiers, and so on, Artificial Color results in a significant increase in the classification accuracy.

Too much detail not only requires too much processing, but also it distracts the pattern recognition by irrelevant noise inevitable at high resolution. Broad spectral bands rather than narrow ones help us attend to the essential information by averaging out the largely irrelevant variations at maximum resolution. Conversely, limiting the amount of information that the system uses, keeps it from remembering the nonrepeating details.

So it seems reasonable to apply Artificial Color to handling the same problem in technology. The Artificial Color program is:

1. Design two or three linear discriminants in spectral space that are broad and spectrally overlapping
2. Project the spectral data onto those discriminants at each pixel.
3. Apply powerful pattern recognition means to discriminate targets on the basis of those two or three discriminants
4. Attribute those discriminants to the image detected to facilitate discrimination by spectrum.

Subsequent papers will discuss and illustrate inclusion of the spatial and (when available) polarization information in producing a final segregated 2D image with targets identified, located, and even posed. For now, however, we concentrate on the spectral information.

We have been unable to find previous work on projection of observed hyperspectral data onto broad overlapping curves and the subsequent use of those projections in powerful statistical pattern recognizers and the use of those projections to accomplish fast image segmentation.

Myron J. Block of Optix LP has done some related work in a field he calls "kromoscopy" [27]. Two sentences from his paper indicate that our Artificial Color and his Kromoscopy attack the same general problem and are both biomimetic. "Considering that almost 20 million colors are resolved, this is an impossible feat for only three channels without spectral overlap. The method of kromoscopy and its improvements can be described as the simulation and extension of human color perception to the analysis of radiation modified by its interaction with substances." One other quotation indicates similarity: "In short, kromoscopic analysis is the real-time simultaneous detection of radiation in different but overlapping spectral regions. This method analyzes radiation by an extension of the principles similar to human color perception. Not only does color perception operate on kromoscopic principles, but also olfaction. Nobel laureate Linda Buck states for olfaction, the key kromoscopic principle "partial sensory overlap." Thereafter similarities become more difficult to find. Kromoscopy is active. Artificial Color, like animal color, is almost always passive. He says nothing about the centrality of the pattern recognition algorithm, but that is central to Artificial Color. And, we can find no application of kromoscopy to hyperspectral data. It is always used in looking at transmitted white light results.

The rest of the paper is organized as follows. Section II introduces the concept of Artificial Color and presents the novel classification algorithm: Margin Setting in detail. The application of Artificial Color to two test data (from ARL and from AVIRIS) is represented in Section III. Section IV compares quantitatively Artificial Color with other well-known hyperspectral image classifiers. Section V gives concluding remarks and future research direction.

## **2 SOME BACKGROUND ON ARTIFICIAL COLOR**

Artificial Color is derived from Biological Color. Biological Color is a discriminant computed in brains using stored information and currently-sensed data. In human vision, for example, we can say that each portion of the scene is sensed with three types of cone cells in daylight vision. Their spectral sensitivities overlap considerably as shown in Fig 1. Data are projections of the input spectrum onto multiple sensitivity curves (cone cells in animals). Converting raw data into discrimination information is what the brain must do in real time.

For example, in the RGB model, each color appears in its primary spectral components of red, green, and blue. These components can be represented by the brightness values of the scene obtained through spectral sensitivities based on the following equations:

$$R = \int_{\lambda} E(\lambda) S_R(\lambda) d\lambda \quad (1)$$

$$G = \int_{\lambda} E(\lambda) S_G(\lambda) d\lambda \quad (2)$$

$$B = \int_{\lambda} E(\lambda) S_B(\lambda) d\lambda \quad (3)$$

where  $S_R, S_G, S_B$  are the spectral sensitivity curves on the incoming light or radiance  $E(\lambda)$ , and  $\lambda$  is the wavelength.

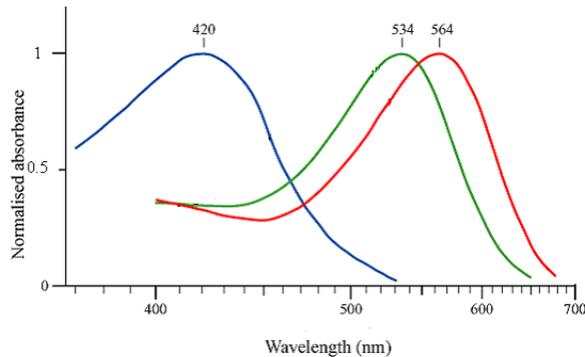


Fig. 1. Responses of the three cone types of most humans to different wavelengths of light.

What animals perceive (the “percept”) is a story the brain computes projected in the percept to an experience that distinguishes among things in the story at least partially on the basis of the experience of color that attributed to each object in the brain’s story.

Artificial Color also computes spectral discriminants using data taken with two or more spectrally overlapping sensitivity curves. Table 1 lists the comparison of Artificial Color and Biological Color. Clearly they are essentially identical. It should not be surprising that Artificial Color works extremely well.

Table 1. This compares biological color (a 440 million-years-ago “invention” of nature with our attempt to translate biology into technology with Artificial Color.

Mechanism	Biological Color	Artificial Color
Sensing	Two or more broad overlapping sensitivity curves	Same
Conversion to information	Computed using the detected signals at each pixel	Same
Attribution of the computed discriminants to the object	Done by a brain and used in that brain to help label objects	Done by a computer and used in that computer to help label objects
Stored information is combined with just-computed information to compute discriminants	Subjective experience of colors attached to objects in the percept	Attribution of a discriminant pattern to each pixel in the image

In principle, any good discrimination method can be used [28]. We chose to use an extremely powerful approach we invented and developed. This method is called Margin Setting [29, 30]. We have compared Margin Setting with several other methods later in this paper. It seems better than any of the others, as we have come to expect.

Here is a brief outline of the training of discriminants in Margin Setting. Start with numerous random points in the hyperspace of measurements (one dimension for each curve). Each point is viewed as a potential prototype for the class of the nearest member of the training set  $S$ . Find the nearest member of another class. That distance  $R_0$  is called the zero-margin radius as shown in Fig. 2. The figure of merit  $F$  of the potential prototype is given by the member number of  $S$  inside the  $R_0$  around that prototype. For instance, in Fig. 2,  $F = 6$ . Then new potential prototypes are chosen based on the old ones. A fixed number of these prototypes is employed. One of the prototypes is selected from the prior generation randomly from a probability distribution function governed by  $F$ 's of the various prototypes. Then the prototype is mutated by choosing a perturbation from a normal distribution centered at the selected prototype. This continues until no improvement in  $F$  is achieved or some other stopping condition is met. The highest scoring prototype is selected.

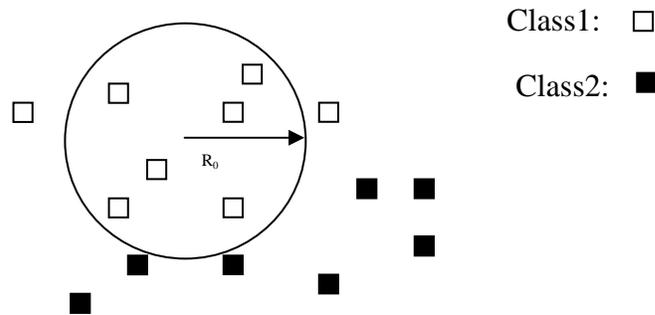


Fig. 2. Zero-margin radius  $R_0$ .

That process is repeated for each new stage. Each stage uses only those samples not classified in earlier stages in  $S$ .

In classification, new data are tested on the first-stage classifiers. If those classifiers indicate a class, it is accepted. If not, go to the second-stage classifiers, and so forth through all the predefined stages.

Margin Setting is represented in details in Fig. 3.

Sample feature vectors  $A = \{\bar{x}_1, \dots, \bar{x}_n\} \subset [0, 1]^q \subset R^q$  taking from the target class and the sample vectors  $B_l = \{\bar{y}_{l,1}, \dots, \bar{y}_{l,m_l}\} \subset [0, 1]^q \subset R^q$  ( $l = 1, \dots, k$ ) taking from the other  $k$  objects. Denote  $B = \bigcup_{l=1}^k B_l$ .

**Notation:**

Rand( $F$ ): random numbers taking from distribution  $F$ .

$\text{Unif}(D)$ : uniform distribution function on set  $D$ .  
 $\text{Card}(D)$ : the figure of merit of set  $D$ .  
 $[a, b]^q$ : a  $q$ -dimensional cube.  
 $O(\bar{z}, R)$ : a  $q$ -dimensional ball centered at  $\bar{z}$  with radius  $R$ .  
 $I$ : designed number of generations.  
 $\mathcal{E}$ : designed error tolerance.  
 $s$ : designed number of samples taken for each generation.  
 $\mathcal{D}$ : designed size of perturbation.  
 $L$ : designed number of mutations for each generation.

**Algorithm:**

1. Compute  $N \equiv \text{Card}(A)$ , and set  $i = 0$ .
2.  $i := i + 1$ . If  $i > I$  or  $\text{Card}(A)/N < \mathcal{E}$ , output  $\{(\bar{c}_j, r_j; N_j) : j = 1, \dots, i - 1\}$ , and stop!
3. Take  $\bar{c}_{i,1}, \dots, \bar{c}_{i,s}$  from  $\text{Rand}(\text{Unif}([0, 1]^q))$  and set  $t_i = 0$ . For  $j = 1, \dots, s$ , compute  $r_{i,j} = \min_{\bar{y} \in B} |\bar{y} - \bar{c}_{i,j}|$ . If  $N_{i,j} \equiv \text{Card}(A \cap O(\bar{c}_{i,j}, r_{i,j})) = 0$ , discard ball  $O(\bar{c}_{i,j}, r_{i,j})$ ; otherwise,  $t_i := t_i + 1$ , and record  $(\bar{c}_{i,j}, r_{i,j}; N_{i,j})$ .
4. If  $t_i = 0$ , record  $(\bar{c}_i, r_i; N_i) = (\bar{0}, 0; 0)$ , go to step 2; otherwise, for  $\{(\bar{c}_{i,j}, r_{i,j}; N_{i,j}) : j = 1, \dots, t_i\}$ , compute

$$w_{i,j} = \frac{N_{i,j}}{\sum_{j=1}^{t_i} N_{i,j}}, \quad j = 1, 2, \dots, t_i.$$

5. Take  $v_i$  from  $\text{Rand}(\text{Unif}([0, 1]))$ , if  $v_i \in [\sum_{j=1}^{\ell_{i,0}-1} w_{i,j}, \sum_{j=1}^{\ell_{i,0}} w_{i,j}]$  for some  $\ell_{i,0} \in \{1, 2, \dots, t_i\}$ , Set  $N_i = N_{i,\ell_{i,0}}$  and the corresponding ball as  $O(\bar{c}_i, r_i)$ ; and discard all the other balls. Set  $l = 1$ .
6. Take  $\bar{s}_l$  from  $\text{Rand}(\text{Unif}([-\mathcal{D}, \mathcal{D}]^q))$ , mutate  $\bar{c}_i$  to  $\bar{c}_i^1 = \bar{c}_i + \bar{s}_l$ , and compute  $r_i^1 = \min_{\bar{y} \in B} |\bar{y} - \bar{c}_i^1|$ . If  $\text{Card}(A \cap O(\bar{c}_i^1, r_i^1)) \leq N_i$ , record  $(\bar{c}_i, r_i; N_i)$  and set  $A := A \setminus O(\bar{c}_i, r_i)$ , go to step 2; otherwise,  $N_i := \text{Card}(A \cap O(\bar{c}_i^1, r_i^1))$ ,  $\bar{c}_i := \bar{c}_i^1$ ,  $r_i := r_i^1$  and  $l := l + 1$ . If  $l > L$ , record  $(\bar{c}_i, r_i; N_i)$  and set  $A := A \setminus O(\bar{c}_i, r_i)$ , go to step 2; otherwise, go to step 6.

---



---

Fig. 3. Margin Setting.

**3 INITIAL TESTS**

The process starts by picking sensitivity curves. It's assumed that there should be three Gaussian curves,  $G_i(\sigma_i, m_i)$ ,  $i = 1, 2, 3$ , so there are six parameters to be set – three means and

three standard deviations. The parameters chosen were almost random but for the for a few constraints: substantial breadth and overlap among them along with the requirement that each wavelength be registered in more than one curve.

Later the curves will be optimized for a specific discrimination problem. The choice here is one that seems appropriate if there is no *a priori* information. Fig. 4 shows the three sensitivity curves used.

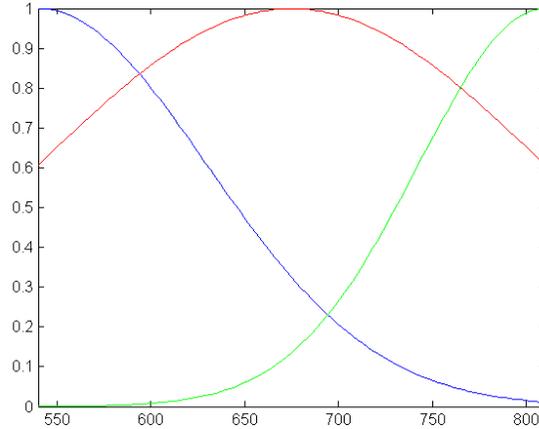


Fig. 4. If there is no *a priori* knowledge of what will be in the scene and the associated spectra, it is best to choose rather neutral curves such as the ones shown here. Those sensitivity curves were used in our experiments.

The next step is to project the spectral data at each pixel onto each of the three sensitivity curves. In an image at a specific wavelength  $\lambda_i$ , the illumination intensity of a pixel at  $P(x, y)$  is  $I_i(x, y, \lambda_i)$ . We will obtain measurements in three bands for  $\lambda_i$  – call it a measurement vector  $C(x, y)$ .

$$C(x, y) = (r, g, b) \tag{4}$$

Where

$$r(x, y) = \sum_{\lambda_{\min}}^{\lambda_{\max}} I_i(x, y, \lambda_i) G_1(\lambda_i), \tag{5}$$

$$g(x, y) = \sum_{\lambda_{\min}}^{\lambda_{\max}} I_i(x, y, \lambda_i) G_2(\lambda_i), \text{ and} \tag{6}$$

$$b(x, y) = \sum_{\lambda_{\min}}^{\lambda_{\max}} I_i(x, y, \lambda_i) G_3(\lambda_i) \tag{7}$$

The resulting irreversible three data at each pixel are then used to discriminate among targets.

### 3. 1 Using ARL (U. S. Army Research Laboratory) data

We obtained the spectral-spatial-polarization maps of a quite ordinary scene provided by the U. S. Army Research Laboratory. The data used here are hyperspectral enhanced as well by polarization images of a quite ordinary outdoor scene. The spectral sampling is done with an

AOTF (Acousto Optic Tunable Filters) every 10 nm between 540 nm and 810 nm using both vertical and horizontal polarization analyzers. Figure 5 shows an ordinary RGB digital camera image of the scene.



Fig. 5. This figure shows an ordinary RGB camera image of the scene for reference.

At any wavelength, two gray scale images are observed – one for each polarization state. Examples are shown in Fig. 6.

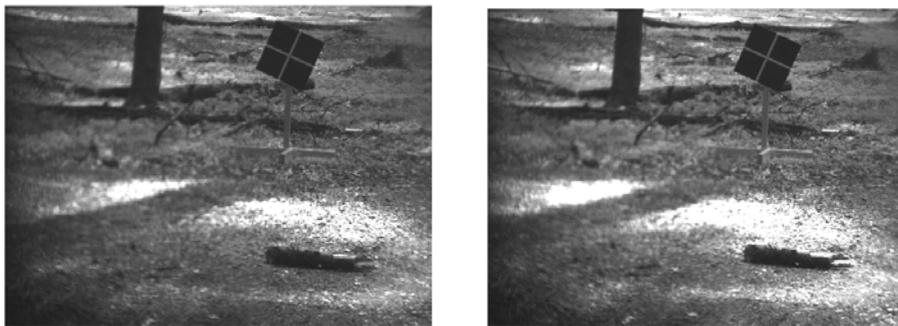


Fig. 6. At any wavelength, there are gray scale images such as these for wavelength 540 nm. The leftmost (rightmost) scene is imaged through a vertically (horizontally) oriented polarizer.

Only one polarization is employed here for simplicity. The polarization information is also useful as shown in [31]. The polarization alone provided reasonably good discrimination. Those two types of discriminants have not been merged yet.

Applying the Artificial Color method used before to this reduced data set, we trained filters that should recognize the red and the blue squares very efficiently.

Applying those filters to the reduced images led to the images shown in Fig. 7.

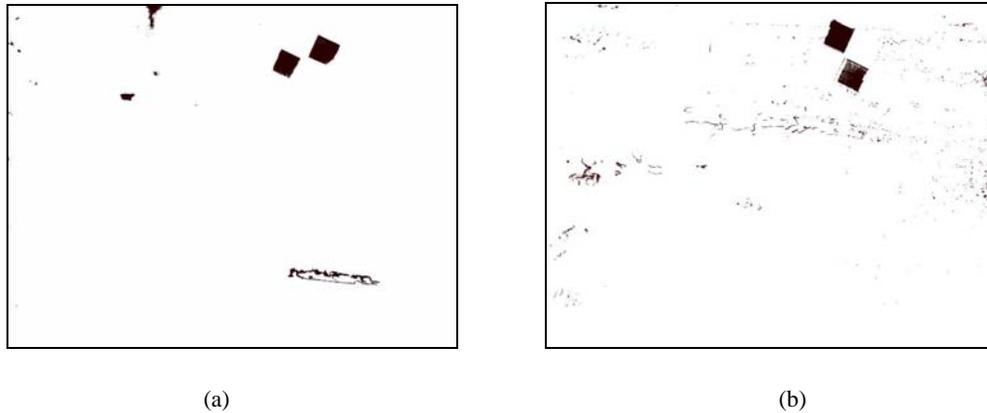


Fig. 7. (a) The resulting Artificial Color image to recognize the blue flag components is interesting because it also allows some of the information on the unidentified object (limb) to survive the operation to some extent. (b) shows the result of applying Artificial Color to recognize the red flag components. It does not pass the information on the unidentified object at all.

### 3.2 Using AVIRIS data

AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) standard data products were downloaded from Jet Propulsion Laboratory Web site. The AVIRIS data were acquired over Moffett Field and were converted to apparent surface reflectances with proper calibration and correction for atmospheric effects. In the reflected visible and near infrared spectrum, AVIRIS samples between 380 and 2500 nm in 224 spectral wavelengths (bands) of 10 nm width. Some of AVIRIS wavelengths (bands) had zero/negative signal values due to poor sensor response or other collection issues. In our experiment, these negative values are set to 0s.

Figure 8 shows one of AVIRIS hyperspectral images over the Moffett Field, CA, an area with water, vegetation, and urban structures. Ten training samples (pixels) are randomly selected from every class (water, vegetation, and urban structures). The three Gaussian sensitivity curves used in AVIRIS data are same in Fig. 3 in the range of 380-2500 nm. Figure 9 shows the segmentation of Artificial Color.



Fig. 8. One of the AVIRIS Hyperspectral Images representing a narrow spectral band

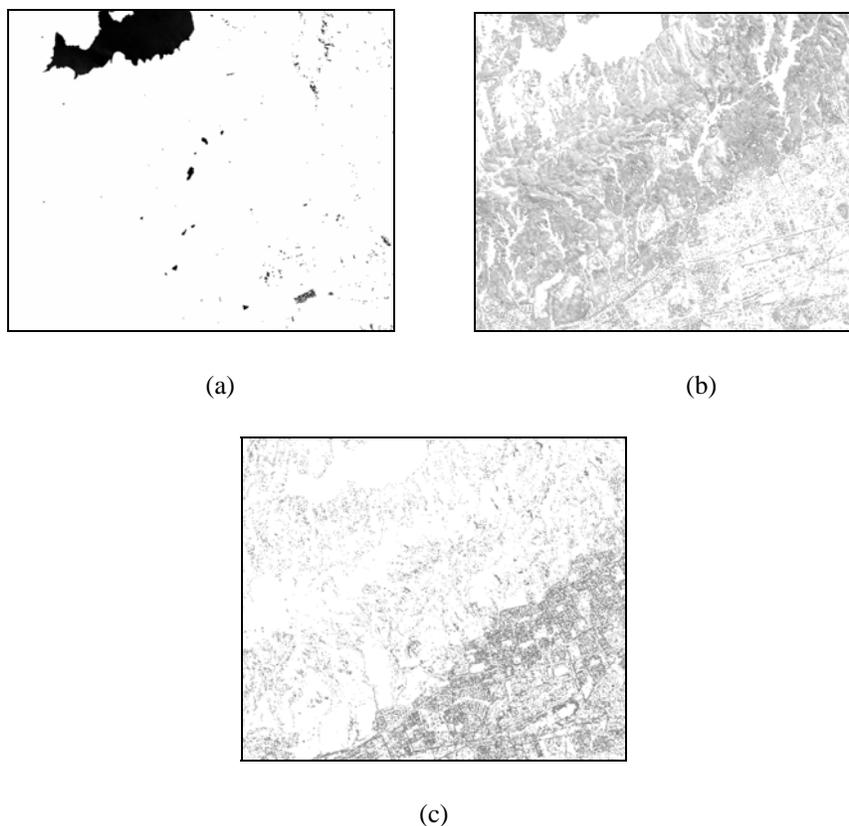


Fig. 9. The segmentation results of Artificial Color: (a) water (b) vegetation (c) urban structures.

Clearly, a great deal of information relevant to scene analysis can be obtained by projecting the hyperspectral data onto a few broad overlapping spectral sensitivity curves. This simplifies data analysis considerably and may even give results better than those we would obtain using all of the information.

#### **4 COMPARISONS WITH POPULAR OTHER PATTERN RECOGNITION METHODS**

Since the task is so easily described in the case of the red and blue squares, we use the ARL dataset. To make it possible to quantify our effects, we segmented the images in Fig. 6 by eye as best we could, see Fig. 10. This became our standard allowing us to compare quantitatively Artificial Color with other well-known hyperspectral image classifiers.

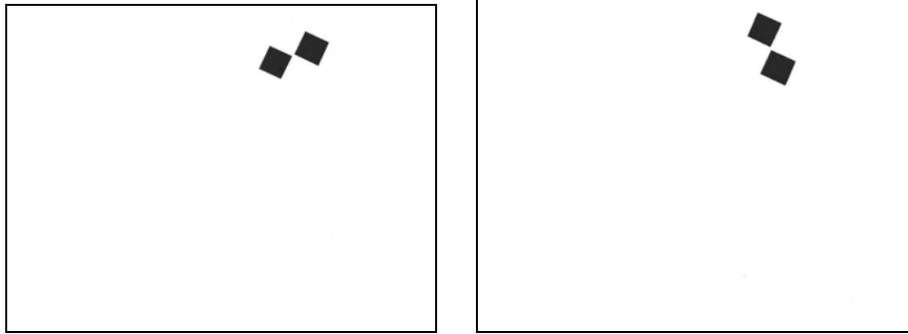


Fig. 10. These hand-segmented images were used as the ideal with which the quantification evaluation of some well-known algorithms was done.

But comparisons between conventional pattern recognition (that deals with two bad things – false positives and false negatives that can be traded off against each other. Margin setting produces a third bad thing – unclassifiable pixels. In our comparison we somewhat mutilated Margin Setting by testing only the zero-margin case and simply ignored the unclassified pixels. The number of pixels in the red and blue squares is easy to determine. That allows us to see what fraction of those that were labeled correctly. False negatives for this purpose are simply the complement (in percentages) of the number of positives recognized in the squares. Tables 2 and 3 summarize the results.

The performance of Artificial Color discrimination depends strongly on the margin that is chosen [24]. In this work we used zero margin, for no particular reason. It is not necessarily the best choice. That is true in all cases, because we have no numerical figure of merit for comparison. A simple, crude means to compare this particular Artificial Color system with conventional approaches can be done as follows. Use a Bayesian weighting of the percent errors of both types equally weighted. To that, add  $\frac{1}{2}$  of the Unclassified percent on the theory that not classifying is better than classifying incorrectly.

The columns in the next two tables have nonobvious meanings. In the blue square of the original image there were 13113 pixels. The column labeled true positives shows how many of these were found for each method in percent of pixels in the squares correctly identified. The number of pixels labeled blue that were outside the square is given in the column labeled False Positives and shown in the second column. This may be somewhat misleading as some of the points outside the square could actually be about the same color as the pixels in the square. Again, we give use relative values, that is, we took the number of those so-called false positives and divided them by the number of true positives. The third column gives the percent of false negatives. Because the percent of false positives is so low, this column is essentially the complement of the first column. The fourth column is a crude Bayesian costing with costs of both false positives and false negatives being set to 1 and the cost of an unclassified pixel being set to  $\frac{1}{2}$ . That is, we add the number of false positives to the number of false negatives, and half the number of unclassified pixels to obtain a cost.

These numbers will allow us to compare the results from all of the methods. With tuning of the Margin Setting parameters, it might very well perform much better. But even here, Artificial Color classifies 74% of the pixels correctly. The SVM (Support Vector Machine) comes in second with 65%.

Averaged over both red and blue squares, Artificial Color identified 76% of the cells properly, while SVM identified 70% right. And Minimum Distance also identified 70% correctly.

Table 2. The other methods are compared with our Artificial Color results here for the blue squares.

Name	True Positive Rate $\alpha$ (%)	False Positive Rate $\beta$ (%)	False Negative Rate $\gamma$ (%)	Unclassified Rate $\delta$ (%)	Bayesian Cost = $(\beta + \gamma) + \delta/2$
<b>Artificial Color</b>	74	67	26	1.9	<b>93.8</b>
<b>Support Vector Machine</b>	65	72	35	0	<b>107</b>
<b>Neural Net</b>	63	90	37	0	<b>127</b>
<b>Parallelepiped</b>	14	0	86	30	<b>101</b>
<b>Minimum Distance</b>	61	151	39	0	<b>190</b>
<b>Maximum Likelihood</b>	24	0	76	0	<b>76</b>
<b>Mahalanobis</b>	<b>60</b>	<b>30</b>	<b>40</b>	<b>0</b>	<b>70</b>

Table 3. The other methods are compared with our Artificial Color results here for the red squares.

Name	True Positive Rate $\alpha$ (%)	False Positive Rate $\beta$ (%)	False Negative Rate $\gamma$ (%)	Unclassified Rate $\delta$ (%)	Bayesian Cost = $(\beta + \gamma) + \delta/2$
<b>Artificial Color</b>	78	79	22	2.1	<b>102</b>
<b>Support Vector Machine</b>	74	91	26	0	<b>117</b>
<b>Neural Net</b>	73	0.9	27	0	<b>28</b>
<b>Parallelepiped</b>	10	0	90	52	<b>116</b>
<b>Minimum Distance</b>	78	121	22	0	<b>143</b>
<b>Maximum Likelihood</b>	3	0	97	0	<b>97</b>
<b>Mahalanobis</b>	<b>66</b>	<b>37</b>	<b>34</b>	<b>0</b>	<b>71</b>

## 5 CONCLUSIONS

Remote sensing needs high discrimination analysis and flexibility. The methods described above show how to convert a HSI data cube quickly and efficiently into a simple gray scale

image that represents the probability of any given pixel belongs to the class of interest. In addition the simplicity of the scheme allows the possibility of doing these operations in real time – also an important goal. The objective of this study was to show that Artificial Color

1. Is based firmly in biology. Animal spectral processing of images is almost exclusively through such projections onto broad, spectrally-overlapping curves (cone cell sensitivity. of course, at night only the rods produce substantial signals, so “color” can not be computed by the brain.
2. Can be used to simplify the discrimination among targets quite dramatically and still offer excellent discrimination. It does this just as Biological Color does: projecting the complex spectrum onto two or more broad, spectrally overlapping curves to produce discriminants not descriptions. In Biological Color, the spectra arise from the contents of the field but are much too complex to analyze in brains that must act in “real time” (real time should have been called biological reaction time) using very slow components (neurons).
3. As we do it, uses a set of predetermined inner products of reference and live projections. It can be very fast, as DSPs (Digital Signal Processors) all do this operation.
4. Capable of doing even better if we have significant *a priori* information about the targets and interferants.
5. Should be able to do even better by designing sensitivity curves for the specific task [26, 32].
6. Can be used to segment by polarization data. That should allow us transcription of polarization into analog polarization as we showed using only polarization to recognize the vertical pole in Fig. 5.

More importantly, Artificial Color can be used to reduce the many measured spectral data to a single number at each pixel. That number would be the probability that the observed hyperspectral values indicated the target or not. This turns the target identification and location measurement indicted by a data cube into merely a spatial analysis problem. The 2D image analyzed spatially has as its amplitude values likely to belong to the target. That task is greatly simplified from the task of directly analyzing the data cube. So it can be done quite rapidly relative to processing means working in the incommensurate dimensions of the data cube.

Figure 11 shows that the main difference between the way you analyze signals from your eyes and the way we propose to analyze data cubes from HyperSpectral Imaging cameras is the ways the separate spectral and spatial analysis. Nature does those tasks in parallel then merges them. We propose to do them sequentially, where there is no merging to do, nor is there any necessary speed loss. The sequential analyses can be pipelined to achieve roughly the same concurrency and we do not have to solve the merging problem.

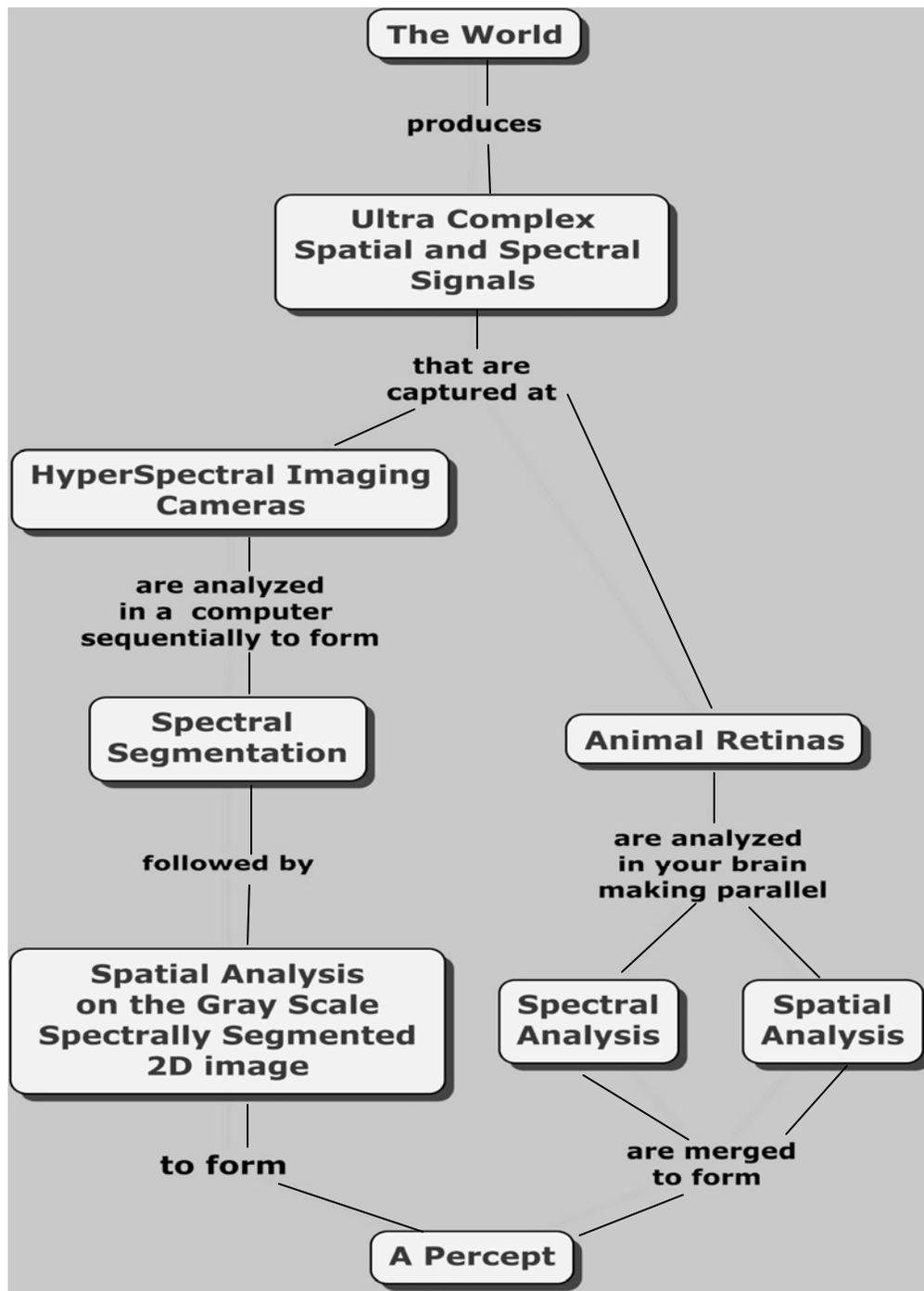


Fig. 11. Biology and this paper follow almost the same routes to image understanding. The primary difference is the type of concurrency used. Animals use parallelism, while we prefer pipelining.

### Acknowledgment

This work is supported by AFOSR under award No. FA9550-07-1-0061.

## References

- [1] A. Parker, 2003, *In the Blink of an Eye*, Perseus Publishing , Cambridge, MA. (2003).
- [2] M. A. S. McMenamin and D. L. McMenamin, *The Emergence of Animals*, Columbia University Press, New York. (1990).
- [3] H. J. Caulfield, “Artificial color,” *Neurocomputing* **51**, 463-465 (2003) [doi:10.1016/S0925-2312(02)00698-7]
- [4] J. Fu, H. J. Caulfield, and S. R. Pulusani, “Artificial color vision: a preliminary study,” *J. Electron. Imag.* **13**, 553-558 (2004) [doi:10.1117/1.1760081].
- [5] H. J. Caulfield, J. Fu, and S. M. Yoo, “Artificial color image logic,” *Inform. Sci.* **167**, 1-7 (2004) [doi:10.1016/j.ins.2003.09.027].
- [6] C. I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer Academic Publishers (2003).
- [7] C. Yang and J. H. Everitt, “Evaluating airborne hyperspectral imagery for mapping waterhyacinth infestations,” *J. Appl. Remote Sens.* **1**, 013546 (2007) [doi:10.1117/1.2821827].
- [8] X. Jia and J. A. Richards, “Efficient maximum likelihood classification for imaging spectrometer data sets,” *IEEE Trans. Geosci. Rem. Sens.* **32**, 274-281 (1994) [doi:10.1109/36.295042].
- [9] A. Mazer, M. Martin, M. Lee, and J. E. Solomon, “Image processing software for imaging spectrometry data analysis,” *Rem. Sens. Environ.* **24** 201-210 (1988) [doi:10.1016/0034-4257(88)90012-0].
- [10] J. P. Reitz, B. V. Brower, and A. Lan, “Hyperspectral compression using spectral signature matching with error encoding,” *Proc. SPIE* **2821**, 64 (1996) [doi:10.1117/12.257185]
- [11] H. Kwon and N. M. Nasrabadi, “Hyperspectral target detection using kernel spectral matched filter,” Conf. Computer Vision Pattern Recog. Work. 127- 127, (2004).
- [12] B. F. Merembeck and B. J. Turner, “Directed canonical analysis and the performance of classifiers under its associated linear transformation,” *IEEE Trans. Geosci. Rem. Sens.* **18**, 190–196 (1980) [doi: 10.1109/TGRS.1980.350272].
- [13] F. Tsai, E. K. Lin, and K. Yoshino, “Spectrally segmented principal component analysis of hyperspectral imagery for mapping invasive plant species,” *Int. J. Rem. Sens.* **28**, 1023–1039, (2007) [doi: 10.1080/01431160600887706].
- [14] X. Jia and J. A. Richards, “Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification,” *IEEE Trans. Geosci. Rem. Sens.* **37**, 538-542 (1999) [doi: 10.1080/01431160600887706].
- [15] A. Cheriyyadat and L. M. Bruce, “Why principal component analysis is not an appropriate feature extraction method for hyperspectral data,” *IEEE Int. Geosci. Rem. Sens. Symp.* **6**, 3420- 3422 (2003) [doi:10.1109/IGARSS.2003.1294808].
- [16] J. Li, L.M. Bruce, J. Byrd, and J. Barnett, “Automated detection of Pueraria Montana (kudzu) through Haar analysis of hyperspectral reflectance data,” *Proc. Geosci. Rem. Sens. Symp.* **5**, 2247-2249 (2001).
- [17] J. C. Harsanyi and C. I. Chang, “Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection,” *IEEE Trans. Geosci. Rem. Sens.* **32**, 779-785 (1994) [doi: 10.1109/36.298007].
- [18] H. Ren and C. I. Chang, “A generalized orthogonal subspace projection approach to unsupervised multispectral image classification,” *IEEE Trans. Geosci. Rem. Sens.* **38**, 2515 – 2528 (2000) [doi: 10.1109/36.885199].
- [19] S. Johnson, “Performance comparison between orthogonal subspace projection and the constrained signal detector,” *Opt. Eng.* **46**, 067005-1-5 (2007) [doi:10.1117/1.2749708].

- [20] L. M. Bruce, C. H. Koger, and J. Li, "Dimensionality reduction of hyperspectral data using discrete wavelet transform feature extraction," *IEEE Trans. Geosci. Rem. Sens.* **40**, (2002) [doi:10.1109/TGRS.2002.804721].
- [21] S. E. Qian, A. B. Hollinger, M. Dutkiewicz, H. Tsang, H. Zwick, and J. R. Freemantle, "Effect of lossy vector quantization hyperspectral data compression on retrieval of red-edge indices," *IEEE Trans. Geosci. Rem. Sens.* **39**, 1459–1470 (2001) [doi:10.1109/36.934077].
- [22] R. Y. Tsai and T. S. Huang, "Multiframe image registration and registration," *Advances in Computer Vision and Image Processing*, T.S. Huang, ED., pp. 317-339, JAI Press, Greenwich, CT. (1984).
- [23] L. Yaroslavsky and H. J. Caulfield, "Deconvolution of multiple images of the same object," *Appl. Opt.* **33**, 2157-2162 (1994) [doi:10.1364/AO.33.002157].
- [24] J. Fu, H. J. Caulfield, and T. Mizell, "Applying median filtering with artificial color," *J. Imag. Sci. Tech.* **49**, 498-504 (2005).
- [25] J. Fu, H. J. Caulfield, and A. Bandyopadhyay, "Pairing mathematical morphology with artificial color to extract targets from clutter," *J. Imag. Sci. Tech.* **51**, 148-154 (2007).
- [26] J. Fu and H. J. Caulfield, "Designing spectral sensitivity curves for use with Artificial Color," *Pattern Recogn.* **40**, 2251-2260 (2007) [doi:10.1016/j.patcog.2006.12.023].
- [27] M. J. Block, "Kromoscopy," *Military Medical Technology*, 12(7) – reprinted from the same journal in Volume **9**, Issue 5 (2005) <http://www.military-medical-technology.com/article.cfm?DocID=1062>
- [28] Y. V. Venkatesh and S. K. Raja, "On the classification of multispectral satellite images using the multilayer perceptron," *Pattern Recogn.* **36**, 2161-2175 (2003) [doi:10.1016/S0031-3203(03)00013-X].
- [29] J. Fu, "Joint exploration of artificial color and margin setting: an innovative approach in color image segmentation," Ph.D dissertation, Univ. of Alabama in Huntsville, (2005).
- [30] H. J. Caulfield and K. Heidary, "Exploring margin setting for good generalization in multiple class discrimination," *Pattern Recogn.* **38**, 1225-1238 (2005) [doi:10.1016/j.patcog.2005.01.009].
- [31] J. Fu and H. J. Caulfield, "Applying color discrimination to polarization discrimination in images," *Optic. Comm.* **272**, 362-366 (2007) [doi:10.1016/j.optcom.2006.11.058].
- [32] J. Fu, H. J. Caulfield, and A. J. Bond, "Artificial and biological color band design as spectral compression," *Image Vis. Comput.* **23**, 761-766 (2005) [doi:10.1016/j.imavis.2005.05.009].

**Jian Fu** is an associate professor in Computer Science Department in Alabama A & M University. He received PhD degree in Computer Science and Engineering from University of Alabama in Huntsville in 2005, the MS degree in Computer Sciences and Physics from Alabama A & M University in 1998 and 1996, respectively.

**H. John Caulfield** is a multiple awardee of SPIE (Gabor, Fellow, President's, Governors') and former editor of *Optical Engineering*. Among his publications are 11 books, 22 book chapters, over 200 refereed journal papers, 29 U. S. patents, and numerous popular articles.

**Dongsheng Wu** is an assistant Professor in Department of Mathematical Sciences at University of Alabama in Huntsville. He received his PhD in Mathematics and his MS degree in Statistics from Michigan State University in 2006 and 2005, respectively.

**Wubishet Tadesse** is an associate professor in Department of Natural Resources and Environmental Sciences in Alabama A & M University. He received PhD degree in Plant & Soil Sciences from Alabama A & M University in 2001.