Refining the histogram based segmentation of hyperspectral data

J. Silverman[a], S.R. Rotman[b], K.L. Duseau[c], P.W. Yip[c], and B. Bukhel[b]
[a] Solid State Scientific Corp.; [b] Ben Gurion University of the Negev
[c] Air Force Research Laboratory, Hanscom AFB

jerry.silverman@hanscom.af.mil; phone 781 377-3295; fax 781 377-4814; Solid State Scientific Corp., 27-2 Wright Rd., Hollis NH 03049. srotman@ee.bgu.ac.il; phone +972-8-646-1539; fax+972-8-647-2949; Ben-Gurion University of the Negev, Electrical Engineering Department, Beer-Sheva, 84105 ISRAEL. karen.duseau@hanscom.af.mil; phone 781 377-4047; fax 781 377-4814; Sensors Directorate, Air Force Research Laboratory, AFRL/SNHI, 80 Scott Rd., Hanscom AFB, MA USA 01731; pearl.yip@hanscom.af.mil; phone 781 377-8438; fax 781 377-4814; Sensors Directorate, Air Force Research Laboratory, AFRL/SNHI, 80 Scott Rd., Hanscom AFB, MA USA 01731. batya@ee.bgu.ac.il; Ben-Gurion University of the Negev, Electrical Engineering Department, Beer-Sheva, 84105 ISRAEL.

ABSTRACT

A recently-developed technique of histogram-based segmentation of hyperspectral data allows for a plethora of segmentations. The user can specify the desired number of levels of segmentations, minimum number of pixels defining a peak, and degree of non-linearity in mapping from principal component floating values to histogram bins, all of which affect the derived segmentation. In the present work, we seek to extend previous work which arrives at a small range of clusters or segmentation levels from the image itself. We seek within this range to find "better" segmentations or possibly a unique representative segmentation. The method employed to achieve this goal starts with an over-fine segmentation, i.e. more segmentation levels than needed, and uses quantitative metrics to measure the "quality" of that segmentation and to guide a compression into a reduced segmentation. If the method has merit, different starts should compress down into comparable segmentations. Therefore a measure to establish the similarity of two or more segmentations was developed. Different quantitative metrics were studied and several modes of compression were examined. Some impressive results are presented but the methods are still not robust with respect to segmentation starts and are image dependent as to the best modes of compression.

Keywords: Hyperspectral images, histogram-based segmentation, segmentation levels, compression, segmentation metrics, similarity matrix, K-means algorithm

1. INTRODUCTION

A new histogram-based segmentation of hyperspectral data was described in past conferences[1,2] and is briefly summarized here. Typically, the histogram is produced from the values of the first two or three principal components of a principal component analysis. Peaks (local maxima) in the histogram are located and labeled from 1 to N. The histogram space is then templated to N levels by calculating the distance of any point in the histogram from each peak: that point is then labeled with the label of the closest peak. The closest peak is established using the Gaussian parameters of each peak which are estimated from the co-occurrence histograms of the individual principal components. Each pixel in the data cube is then assigned the digital label of its corresponding point in the histogram and segmentation is accomplished. In order to generate the histograms, one must map the floating point values of the principal components into integer bins. The number of segmented levels increases, up to a point, with the chosen number of integer bins leading from coarse to finer segmentations. A desired number of segmentation levels is requested and an entropy measure is used to guide a non-linear histogram mapping[2] into that final set of integer bins that achieves roughly the desired number of levels.

As indicated, the scaling into integer bins provides a gradation of segmentations. In previous work[3], plots of number of segmentation levels versus number of integer bins with peak_min (the minimum number of pixels required to establish a peak) as a parameter often identify a natural range of clusters in the data. A typical range is N_0 to $N_0 + \Delta N$ with $\Delta N = 2 - 4$. Within that range, many segmentations can be generated. The goal which motivates the current work is to identify best or most representative segmentations in this range in the absence of ground truth and in an objective fashion.

Our approach to achieving this goal is to start with over-fine segmentations, i.e. more segmentation levels than needed, and to use quantitative metrics to measure the "quality" of that segmentation in order to guide a compression into a reduced segmentation. If the method has merit, different starts should compress into comparable segmentations. Therefore a measure to establish the similarity of two or more segmentations was developed. Different quantitative metrics were studied and several modes of compression were examined.

Our presentation is organized as follows: Section 2 explains the quantitative metrics, the similarity measure, and the modes of compression; Section 3 examines simulated data and real images in the VIS/NIR and MWIR; Section 4 discusses the results and the direction of continued research.

2. TOOLS AND APPROACH

We initially tested pair-based metrics derived from the ACTS Algorithm[4]. The ACTS method of segmentation works in the context of a descending tree structure in which each parent cluster is potentially split into two child clusters based on the scatter in the parent cluster. The pair-based metrics test the combining of two segments based on the average scatter measure of the original segmentation and the average scatter measure after combining every combination of two segments. That combination which made the greatest reduction in the scatter measure was effected and the process of pair combination repeated until a minimum scatter measure or metric value was achieved. The pair-based process of segmentation compression was limited in that two segments were forced to combine or not to combine; there was no way to disseminate pixels to different segments. Typically, minimums occurred at very coarse (too few levels) final segmentations. Analysis of these metrics will be published at a later date.

Better results were afforded by the metrics described next which form the main thrust of the present work. This group of metrics was based on assigning each segment a measure of scatter spread; both Euclidian distance and angular measures were considered. The four variations of these metrics, which represent, in effect, an average scatter per segment over the segmentation, change with distance measure and how the segments are averaged. A summary is given in the table below.

For a given segmentation, N_p is the total number of pixels. $n_1, n_2, n_3, \cdots n_s$ is the number of pixels in each segment; s is the number of segments. $N_p = n_1 + n_2 + n_3 + \cdots + n_s$. $\bar{\bar{p}}_1, \bar{\bar{p}}_2 \cdots \bar{\bar{p}}_s$ are the average vector profiles of each segment over the number of bands. $\vec{x}_{ij}$ is the vector profile of the i[th] pixel in the j[th] segment. M represents the final metric value.

Metric	Metric formula	Distance measure used	Weighting of averages
PE	$M = \frac{1}{N_p} \sum_{j=1}^{s} \sum_{i=1}^{n_j} \left\| \bar{\bar{p}}_j - \vec{x}_{ij} \right\|$	Euclidian	Equal pixel weighting
SE	$M = \frac{1}{s} \sum_{j=1}^{s} \frac{1}{n_j} \sum_{i=1}^{n_j} \left\| \bar{\bar{p}}_j - \vec{x}_{ij} \right\|$	Euclidian	Equal segment weighting
PA	$M = \frac{1}{N_p} \sum_{j=1}^{s} \sum_{i=1}^{n_j} \cos^{-1}\left(\frac{\bar{\bar{p}}_j \bullet \vec{x}_{ij}}{\left\|\bar{\bar{p}}_j\right\| \cdot \left\|\vec{x}_{ij}\right\|} \right)$	Angular	Equal pixel weighting
SA	$M = \frac{1}{s} \sum_{j=1}^{s} \frac{1}{n_j} \sum_{i=1}^{n_j} \cos^{-1}\left(\frac{\bar{\bar{p}}_j \bullet \vec{x}_{ij}}{\left\|\bar{\bar{p}}_j\right\| \cdot \left\|\vec{x}_{ij}\right\|} \right)$	Angular	Equal segment weighting

We evaluate our metrics by how well it compresses multiple over-segmented starts into a common final segmentation. We have used each of these four metrics to compress an N level segmentation to an (N-1) level segmentation as follows. N different (N-1) segmentation variations are generated by taking each segment in turn and assigning each of its pixels to one of the other segments (the one whose average profile is closest to that pixel). The new metric of each (N-1) segmentation variation is computed. The variation with the lowest metric is taken as the compressed (N-1) segmentation.

As one of our goals is to increase the commonality of different starting segmentations generated from the histogram technique, we require a quantitative measure of similarity between pairs of segmentations. Given two segmentations to N_1 and N_2 levels with $N_1 \leq N_2$, we can define a similarity matrix, $\mathbf{S}_{ij}$, with i running from 1 to N_1 and j running from 1 to N_2. Each matrix element, $\mathbf{s}_{ij}$, is the total number of pixels labeled i in the N_1 segmentation and j in the N_2 segmentation. We leave as an exercise to the reader to show that $M_1 + M_2 + M_3 + \cdots + M_{N_1}$, the sum of the N_1 largest values in this matrix, under the constraint that one value per row and per column only is used, is the total number of pixels segmented identically in both segmentations. We will present these results as a percent of total pixels.

In using our metrics, five modes of compression were identified and tested for reducing over-segmented data. All modes of compression are based on the first mode. The first mode compresses segmentations from N to (N-1) to (N-2), to (N-final) in the fashion described above stopping at the (N-final) segmentation with the lowest metric value. Mode 1K finishes mode 1 with the standard K-Means algorithm[5] applied to the final compressed segmentation from mode 1. The K starting centers are taken as average profiles of the segments. We use a Euclidian distance measure and iterate the K-Means process ten times. Mode 2 interleaves K-Means with mode 1 compressions as: compression, K-Means, compression, K-means, etc. to a minimum metric value. Mode 3 reduces a group of starting segmentations, N_1, N_2, N_3 where $N_1 < N_2 < N_3$, as follows. The N_1 segmentation is compressed to N_F levels, where $N_F \leq N_1$, in mode 1 compression, then the N_2 and N_3 segmentations are compressed to N_F levels as well (typically requiring going beyond the mode 1 minimum). Mode 3K applies a K-means algorithm to these three N_F segmentations. A capsule summary of the modes is given in the table below.

Mode	Method
1	Compression to a minimum metric value
1K	Final K-Means applied to mode 1 result
2	Compression, K-Means, compression, K-Means, etc. to a minimum metric value.
3	Smallest start compress to N_F by mode 1: other starts compressed to the same N_F.
3K	K-Means finish to mode 3

Our modus operandi in evaluating these techniques of metric-guided compression is to generate a group of starting segmentations and to characterize the evolution of their similarity matrices under various modes of compression. In all the examples shown (excluding the Vis/NIR example) in the following section, the starting recipe was as follows. From the characteristic roll-off plots of segmentation level versus number of integer bins with peak-min as a parameter (see figures 2 to 5 of reference 3), one chooses a representative upper bound value of the natural number of clusters in the particular image: N_0. We then use the interactive, entropy-guided form of histogram based segmentation and seek starting segmentations at N_0, N_0+5, and N_0+10 levels. For some images such as the Vis/NIR example in the next section, roll-off is not observed and a different technique of generating the starting segmentations was needed and is described in treating that example.

3. RESULTS

Example 1: Mid-wave simulated data with Gaussian noise.

In a preliminary examination of our metrics and our modes of compression, we created semi-simulated data to check the efficacy of our algorithms. The simulated data consisted of twelve equal regions, each of which was generated from a single spectral profile extracted from a real set of mid-wave data[6]. Gaussian noise was then added to the simulated data to introduce some variation (a constant sigma of 5.5 in radiance). The spectral profiles from the mid-wave data came from pixels on: the field stop, two different sections of foliage, windows on a building, sky, an unidentified man-made object, a metallic object, a flag pole, two different sections on a building, and two different sections of grass. These profiles are shown in Figure 1.

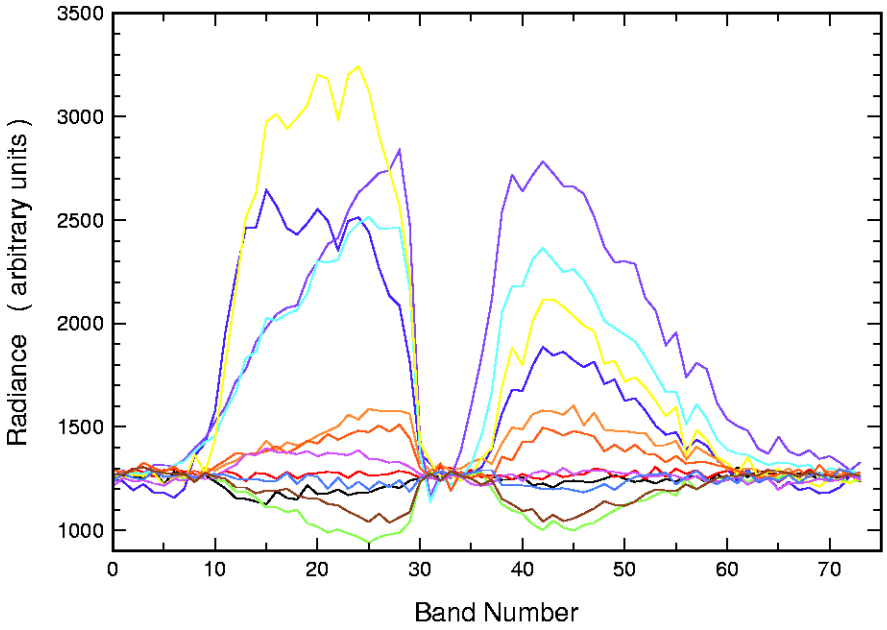

Fig. 1. Representative spectral profiles of the simulated data regions

The roll-off plots of the simulation data are shown in fig 2a. (There are a number of roll-off plots with different peak_min values, but each exactly overlaps the other, so they can not be visually differentiated.) Clearly, the plots point to the presence of twelve major clusters, therefore we set $N_0 = 12$. We requested 12, 17, and 22 levels from the entropy-guided histogram segmentation algorithm which returned starting segmentations of 12, 17, and 20 levels shown in fig 2d, 2e, and 2f (designated A, B, and C). Figure 2b shows the PE and SE metric values as one compresses down from the 20 level starting segmentation. The metric values using PE weights minimize nicely at 12 levels, but the metric values rise for SE weights. SE weighting seems to favor small segments split-off from the larger segments. Here, the SE weighting gave rise to noisy outliers as segments. The PE weighting almost always gave better results for the real data as well. Figures 2g, 2h, and 2i show the mode 1 compressed segmentations of the A, B, and C starts. All starting segmentations compressed nicely into 12 segments and all collapse to the optimum segmentation (Fig. 2c) upon applying the K-Mean algorithm (mode 1K). Mode 2 works for this data set as well but was generally poorer than mode 1K on real data.

In this case as well as the later examples, the angular metrics showed only minor differences from the Euclidean ones and hence we will show only PE and SE results in what follows.

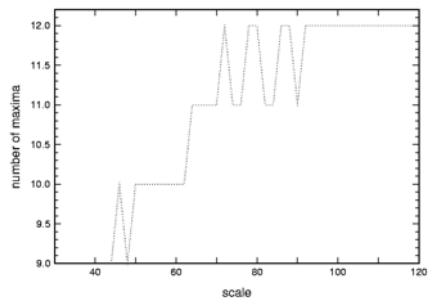

Fig. 2a. Segmentation levels versus scale for noise simulation

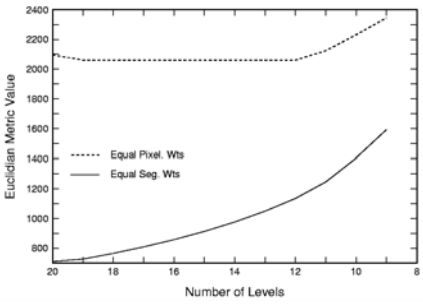

Fig. 2b. Range of metric values

Fig. 2c. Optimum segmentation

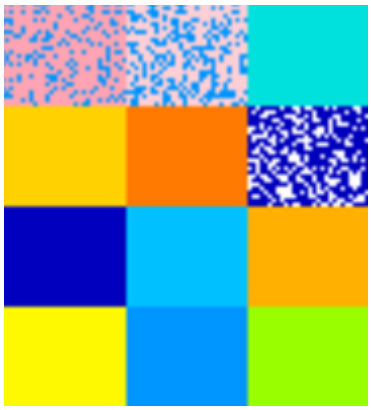

Fig. 2d. 12 level start - A

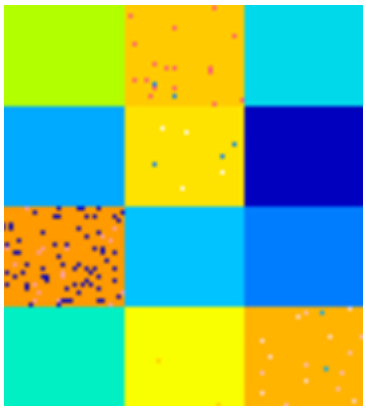

Fig. 2e. 17 level start - B

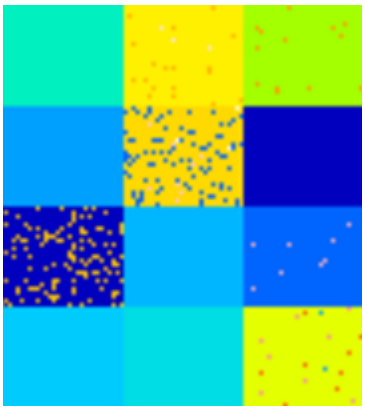

Fig. 2f. 20 level start - C

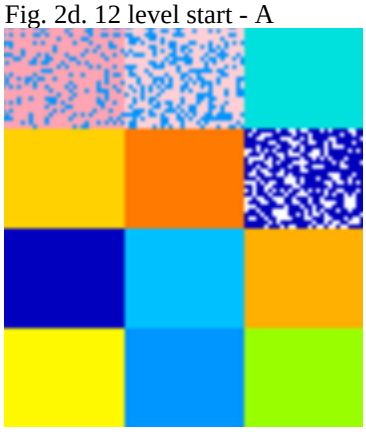

Fig. 2g. Mode 1 PE compression (A)

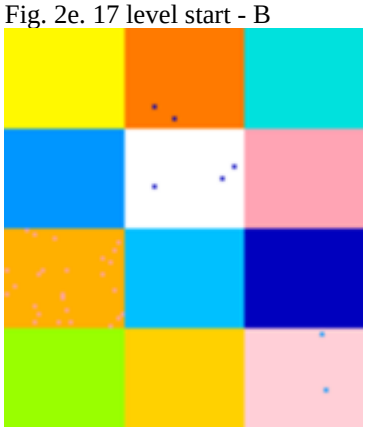

Fig. 2h. Mode 1 PE compression (B)

Fig. 2i. Mode 1 PE compression (C)

Example 2: Mid-wave data from CTHIS[6] camera.

The roll-off plots of the data in example 2 indicate the appropriate number of large-scale clusters to be 8 to 10. Hence, we requested 10, 15, and 20 levels from the entropy-guided histogram segmentation algorithm which returned starting segmentations of 10, 14, and 18 levels (designated A, B, C) as shown in Figure 3. The three starts had 80 to 86 percent similarity of the pixels. Figures 3d through 3o show our compression results using modes 1, 1K, 3, and 3K with PE weights. The similarity table below indicates how close we came to reaching a common segmentation. Mode 1K was very successful and mode 3K extremely successful. For this case, these modes with SE weights (not shown) separate start C from starts A and B.

Similarity table (Fig. 3)	Similarity of A and B	Similarity of A and C	Similarity of B and C
Starts	83 %	86 %	80 %
Mode 1 (PE Weights)	87	87	84
Mode 1K (PE Weights)	97	96	95
Mode 3 (PE Weights)	87	88	85
Mode 3K (PE Weights)	99.5	99	98.6

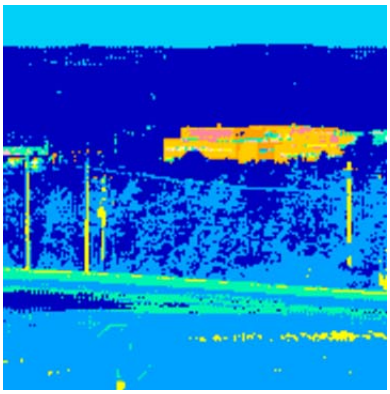
Fig. 3a. 10 level start (A)

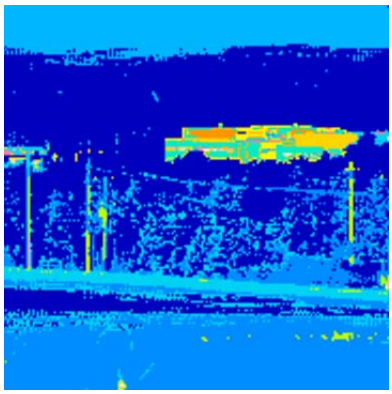
Fig. 3b. 14 level start (B)

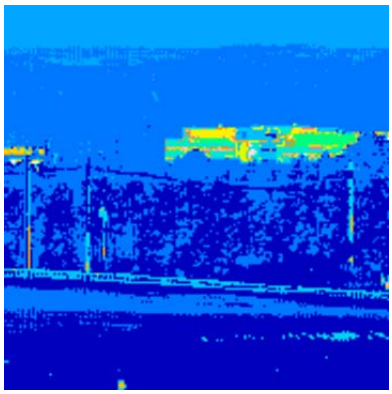
Fig. 3c. 18 level start (C)

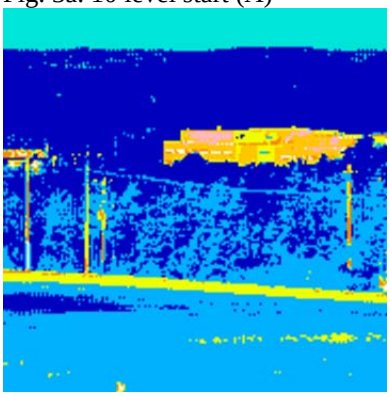
Fig. 3d. Mode 1 (A) 8 levels

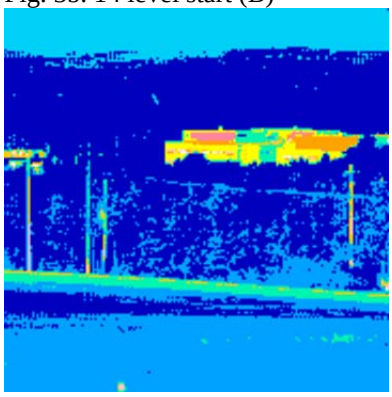
Fig. 3e. Mode 1 (B) 10 levels

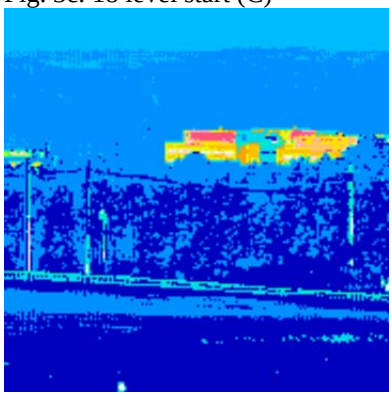
Fig. 3f. Mode 1 (C) 13 levels

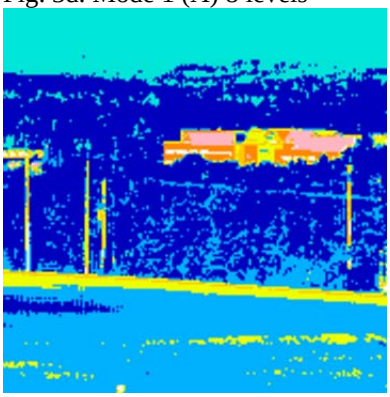
Fig. 3g. Mode 1K (A)

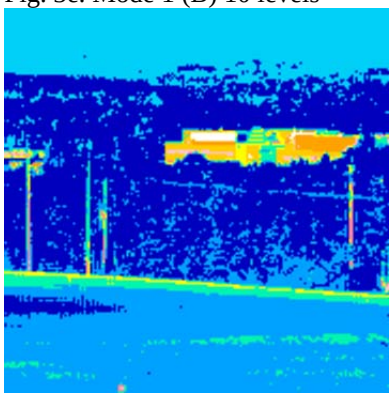
Fig. 3h. Mode 1K (B)

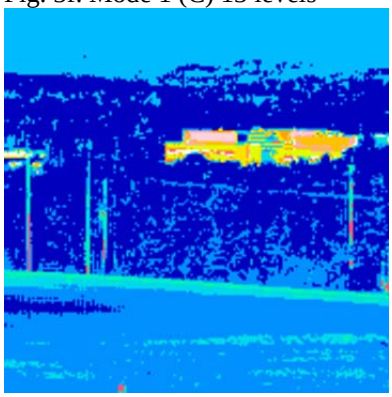
Fig. 3i. Mode 1K (C)

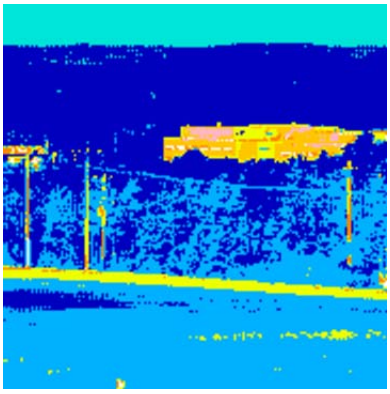

Fig. 3j. Mode 3 (A) 8 levels

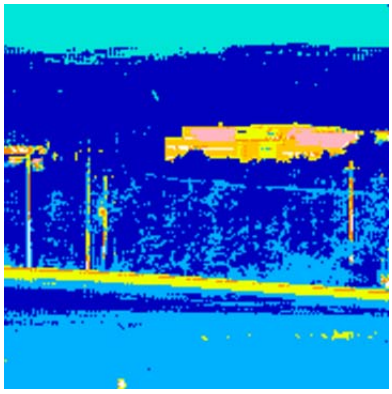

Fig. 3k. Mode 3 (B) 8 levels

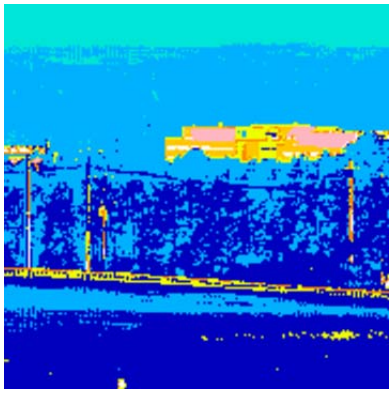

Fig. 3l. Mode 3 (C) 8 levels

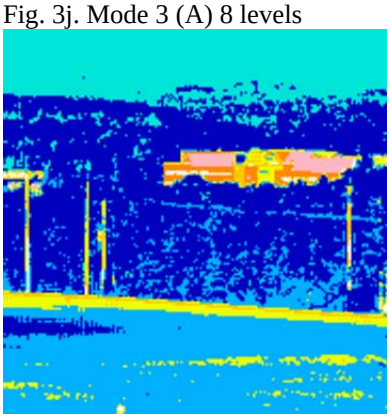

Fig. 3m. Mode 3K (A)

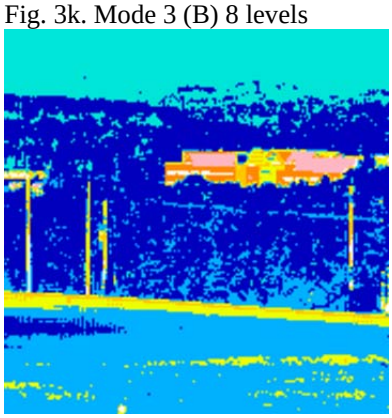

Fig. 3n. Mode 3K (B)

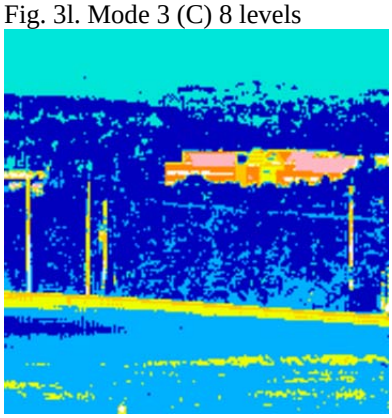

Fig. 3o. Mode 3K (C)

Example 3: Mid-wave data from CTHIS[6] camera.

For example 3 we requested 10, 15, 20 and 30 levels from the entropy-guided histogram segmentation algorithm which returned starting segmentations of 8, 13, 17, and 28 levels. The 8 level segmentation compressed into 5 levels, which appears to be too few and too concentrated on the building. These results are shown separately in Figure 4.

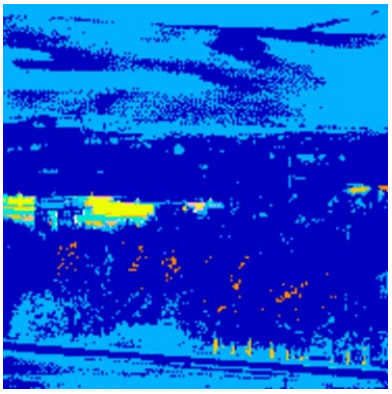

Fig. 4a. 8 level start

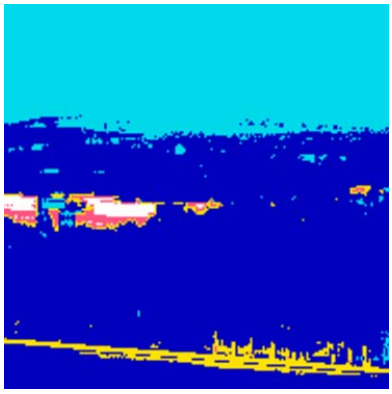

Fig. 4b. Mode 1K (PE wts) 5 levels

Figure 5 shows our starting segmentations of 13, 17 and 28 levels (designated A, B, and C). The similarity table below for example 3 is atypical in that SE weights are somewhat better than PE weights and mode 1K segmentations are , in some cases, less similar than mode 1. The buildings and the near foliage constitute numerous small segments leaving the results quite sensitive to SE versus PE weights. In Figure 5, we show mode 1 and 1K results with SE; the similarity table includes the PE results as well.

Similarity table (Fig. 5)	Similarity of A and B	Similarity of A and C	Similarity of B and C
Starts	78 %	90 %	76 %
Mode 1 (PE Weights)	67	92	67
Mode 1 (SE Weights)	85	91	84
Mode 1K (PE Weights)	84	84	75
Mode 1K (SE Weights)	94	82	79

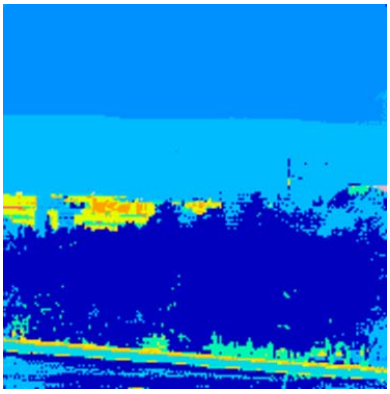

Fig. 5a. 13 level start (A)

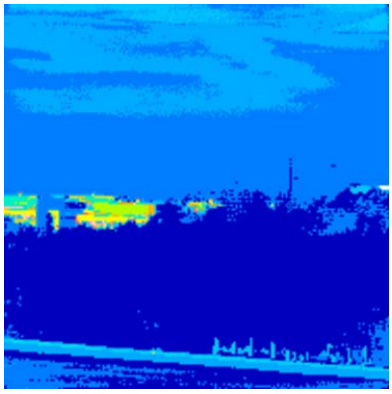

Fig. 5b. 17 level start (B)

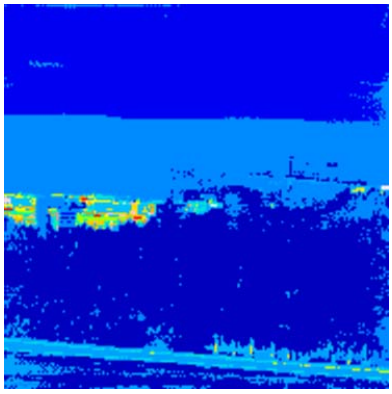

Fig. 5c. 28 level start (C)

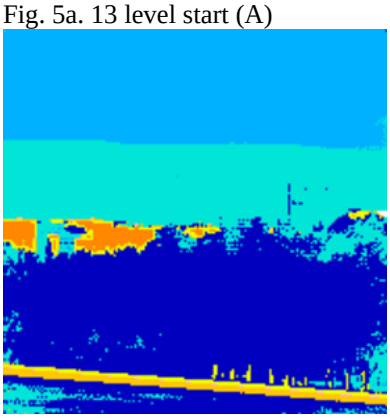

Fig.5d. Mode 1 SE wts (A) 8 levels

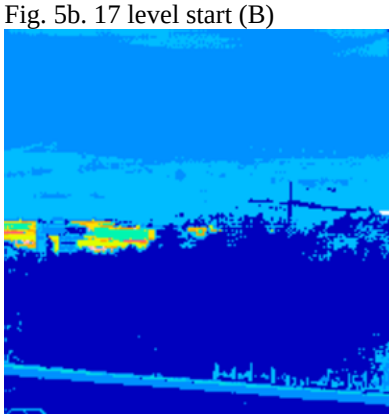

Fig.5e. Mode 1 SE wts (B) 13 levels

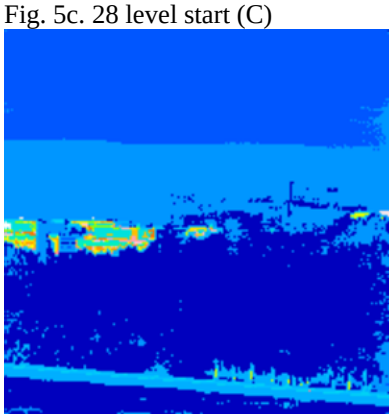

Fig.5f. Mode 1 SE wts (C) 23 levels

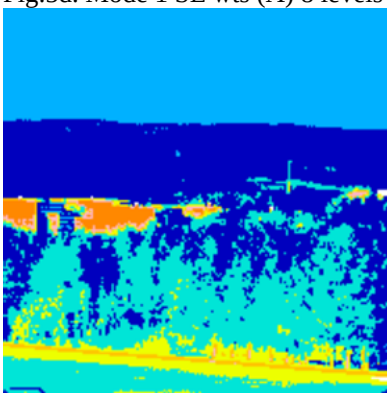

Fig. 5g. Mode 1K SE wts (A)

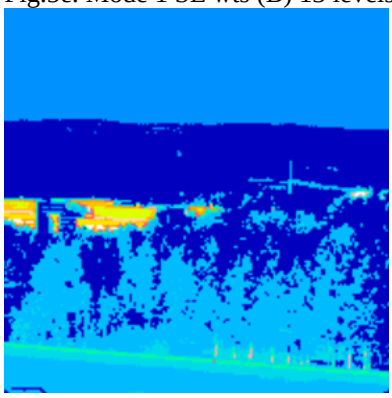

Fig. 5h. Mode 1K SE wts (B)

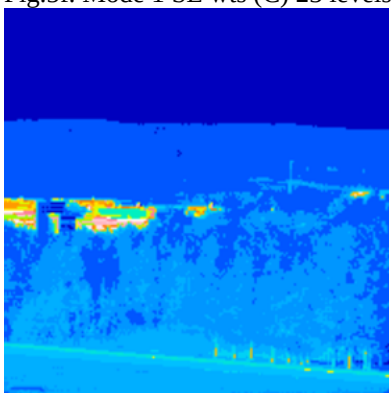

Fig. 5i. Mode 1K SE wts (C)

Example 4: Vis/NIR data from CTHIS[7] camera.

In this example there was no true roll-off of our plots. We looked for stability in the various segmentations generated. We called for an array of 10 segmentations ranging from 10 segments to 30 segments. The returned array of segmentations is then examined for persisting segmentation levels, or "islands of stability". In this data set, there were recurring segmentations found at 10, 14, and 19 levels (designated A, B, C), so we used these as our

starting segmentations (Fig. 6a, b, c). Modes 1 and 2 did not prove successful for this example. As seen in the similarity table below, mode 3 with PE weights and then mode 3K collapses the three starts into a common nine-level segmentation.

Similarity table (Fig. 6)	Similarity of A and B	Similarity of A and C	Similarity of B and C
Start	85 %	78 %	78.5 %
Mode 3 (PE weights)	92	86	90
Mode 3K	98	97	98

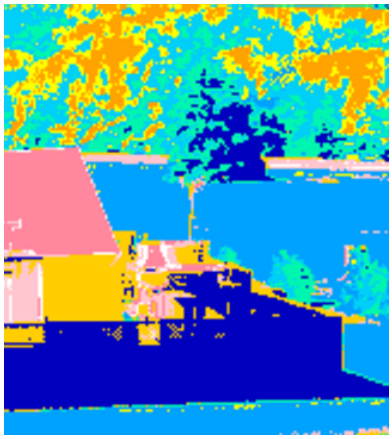
Fig. 6a. 10 level start (A)

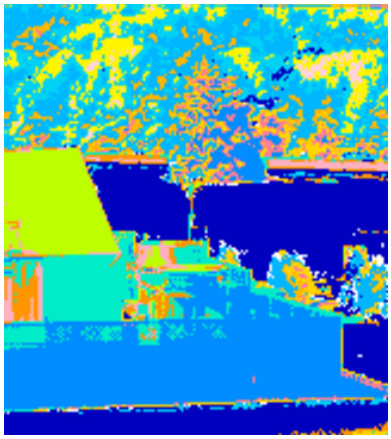
Fig. 6b. 14 level start (B)

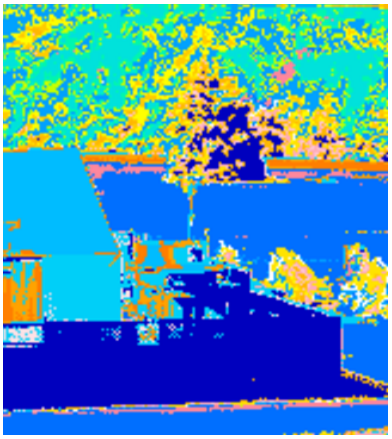
Fig. 6c. 19 level start (C)

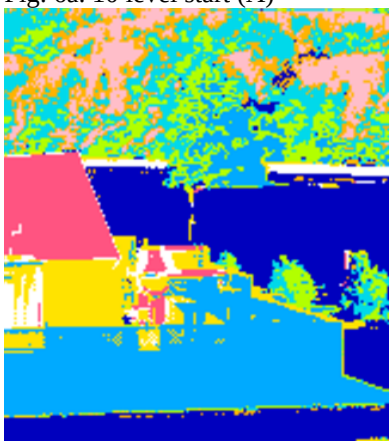
Fig. 6d. Mode 3 (A) 9 levels

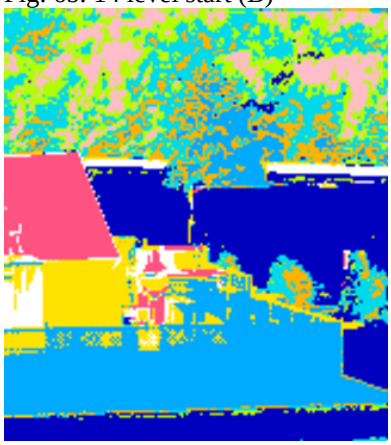
Fig. 6e. Mode 3 (B) 9 levels.

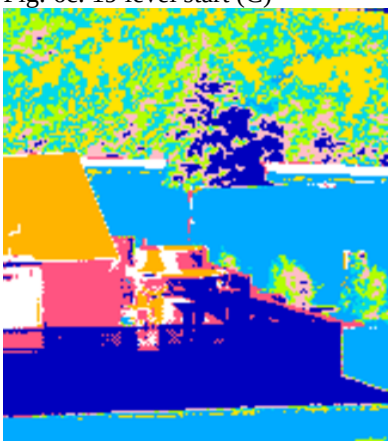
Fig. 6f. Mode 3 (C) 9 levels.

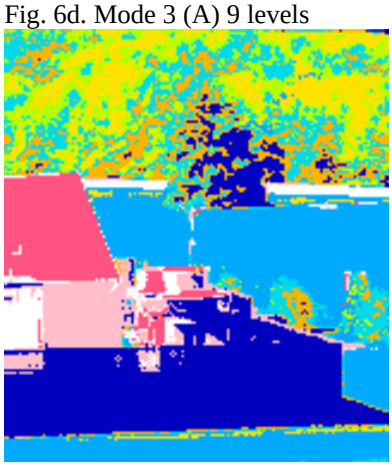
Fig. 6g. Mode 3K (A) 9 levels

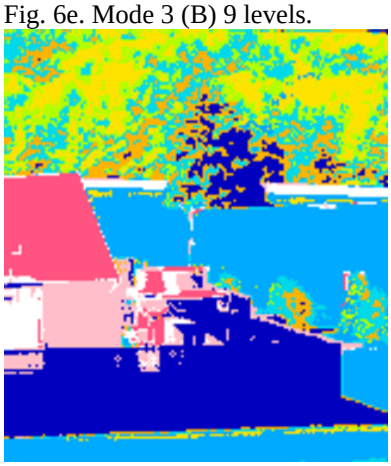
Fig. 6h. Mode 3K (B) 9 levels

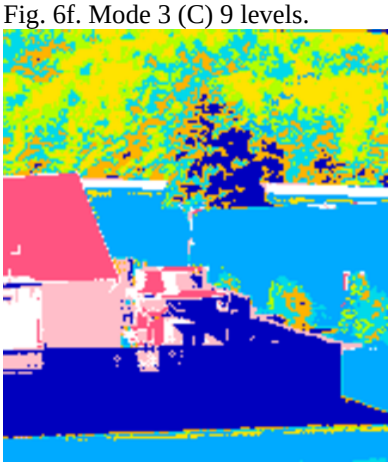
Fig. 6i. Mode 3K (C) 9levels

4. DISCUSSION/FUTURE WORK

In the examples shown, metric compression has extracted a more similar and representative segmentation from several diverse starting segmentations. In particular, mode 3 and mode 3K appear to be promising in compressing different starting segmentations into final segmentations with the same number of levels. However, the robustness of these techniques needs to be improved: the best way to generate starting segmentations as well as the best mode of compression appears to be data dependent. The sought-after best final segmentation is often not forthcoming and indeed may not be a meaningful entity in certain scenes.

The similarity matrix introduced in Section 3 could be used in a more sophisticated way. We reduce our comparison of two segmentations to a single fractional percent, F%, indicating the fraction of pixels "identically" segmented in the pair compared. However, the remaining (100-F)% fraction could, at one extreme, be found in one entry in the matrix, or at the other extreme, be splintered into many small entries. An observer, viewing the images reflecting these extreme cases, would judge the former case as a much more similar segmentation pair than the latter case.

To continue to improve upon this work, testing should continue as new data sets and new generation sensors become available. Data sets from HYDICE example proved to be problematic, perhaps due to the small segments inherent in the data. Data sets with available ground truth should enable more definitive conclusions about the utility of these metrics. In addition, the present idea of over-segmenting and compressing into a common representative segmentation should be joined to and compared with a complementary process such as used in the ACTS algorithm, where one starts with a few, or even just one segment and splits according to metric value tests.

ACKNOWLEDGMENTS

The cameras used to collect the visible and MWIR data were designed and fabricated by William Ewing, Toby Reeves and Steven DiSalvo of our laboratory. This work was carried out under Air Force Task 2305BN00.

REFERENCES

1. J. Silverman, C. E. Caefer, J.M. Mooney, M.M. Weeks, and P. Yip, "An automated clustering/segmentation of hyperspectral images based on histogram thresholding", in *Imaging Spectrometry VII*, Michael R. Descour, Sylvia S. Shen, Editors, Proceedings of SPIE Vol. 4480, 65-75 (2002).

2. J. Silverman, S.R. Rotman and C.E. Caefer, "Segmentation of hyperspectral images from the histograms of principal components", in *Imaging Spectrometry VIII*, Sylvia S. Shen, Editor, Proceedings of SPIE Vol. 4816, 270-277 (2002).

3. J. Silverman, S.R. Rotman, K.L. Duseau, and C.E. Caefer, "Issues in segmenting hyperspectral imagery from histograms", in *Imaging Spectrometry IX*, Sylvia S. Shen, Paul E. Lewis, Editors, Proceedings of SPIE Vol. 5159, 41-51 (2004).

4. N. Viovy, "Automatic Classification of Time Series (ACTS): a new clustering method for remote sensing time series", *Int. J. Remote Sensing* **21**, pp. 1537-1560 (2000).

5. R. A. Schowengerdt, *Remote Sensing: Models and Methods for Image Processing*, Chap. 9, Academic Press, San Diego (1997).

6. J. M. Mooney, V. E. Vickers, M. An, and A.K. Brodzik, "High throughput hyperspectral infrared camera", *J. Opt. Soc. Am. A*, **14(11)**, pp. 2951-2961 (1997).

7. J.E. Murguia, T.D. Reeves, J.M. Mooney, W.S. Ewing, F.D. Shepherd, A.K. Brodzik, "A compact visible/near infrared hyperspectral imager", in *Infrared Detectors and Focal Plane Arrays VI*, Eustace L. Dereniak, Robert E Sampson, Editors, Proceedings of SPIE Vol. 4028 457-468 (2000).