A Pre-Attentive Vision Model for Data Prospecting

Ivan A. GALKIN, Grigori M. KHMYROV, Alexander V. KOZLOV, Bodo W. REINISCH
Center for Atmospheric Research, University of Massachusetts Lowell
Lowell, MA 01851 USA

James C. TILTON, Shing F. FUNG
NASA Goddard Space Flight Center
Greenbelt, MD 20771

and

Antonio PLAZA
University of Extremadura
Caceres, SPAIN

ABSTRACT

An increasing number of technical and academic applications rely on the intelligent systems for imagery data prospecting to cope with the information avalanche. An imagery data prospector pre-classifies input images in categories so as to reduce the unrealistic demand of the manual labor needed to analyze all of them. We studied applicability of the bottom-up pre-attentive vision models to the task of data prospecting. Generally, these models do not assume prior knowledge of the domain discipline and specific characteristics of signatures and merely implement the Gestalt perception principles to find features that stand out against the image background. We developed and tested such model against the RPI plasmagram dataset currently comprising 1.1 million images. The model implements a recurrent neural network that optimizes alignment of rotors placed on top of detected edge elements to identify whether the image contains contours of RPI signal reflections. The paper discusses model design, presents performance results, and outlines future efforts.

Keywords: Pre-attentive Vision, Data Prospecting, Intelligent Systems, Radio Plasma Imager.

1. INTRODUCTION

Intelligent systems for data prospecting have become a necessity in research areas employing imaging instruments with large information capacities. Such intelligent systems serve as automated data classifiers that reduce unrealistic amount of the manual labor needed to analyze all acquired data. We distinguish data prospecting from data mining where intelligent techniques are applied to a large dataset to find previously unknown, interesting dependencies between the elements that constitute the dataset. Data prospecting is about finding meaningful data among data that just takes up space and drawing attention of the human analysts to them.

Building an imagery data prospector is a significant task that involves modeling of the visual data perception by humans. Such model commonly includes a feature binding algorithm tailored to the specifics of the domain. Robustness of the feature binding approach to various real-life imperfections of the images is the key to successful solution of the prospecting task.

Biologically plausible models of visual data perception are drawing much attention recently for their potential in replicating the most sophisticated, adaptive, robust, and intelligent image analysis system that we know of. Reverse engineering of the eye retina architecture and computations in the visual cortex of the brain has brought a suite of important technical solutions (e.g., artificial neural networks) to the task of image recognition. Of particular interest to the data prospecting are models of so-called "pre-attentive" vision system that is found in many living organisms, including human. The pre-attentive vision is a perceptual system whose responsibility is to "pop-up" visual cues in the field of view without willful concentration of attention. It can be described as a task that runs in the background and requests switch of attention to objects that present potential danger or special interest. Pre-attentive vision is especially effective in rapid detection of salient objects by identifying their contours.

Existing studies of the pre-attentive vision suggest that it adheres to a "bottom-up" analysis strategy that assumes no prior knowledge of the features to be discovered in the analyzed image. The bottom-up model is only aware of a general perceptual quality of the image features that makes them stand out against the background. These perceptual considerations are known in the literature since 1930s as the Gestalt principles [1] that merely capture apparent preferences in human perception of visual patterns. For example, when presented with a binary image containing points and bars, it turns out that this perceptual process prefers to single out subset of points lying on some long, smooth and dense curve (corresponding Gestalt principles are good continuation, smoothness, and proximity). Bottom-up feature extraction algorithms [2]-[6] seek detectable low-level image features (dots, bars) that can be grouped together in salient contours under restrictions of the Gestalt. Because each stage extracts elements of higher perceptual strength from its input, this model is often referred to as the Marr's pyramid of perception [7].

2. PRE-ATTENTIVE VISION AND DATA PROSPECTING

We are interested in a particular scenario of automated image analysis where the vision model identifies images of interest and draws attention of human analysts to them. This approach is consistent with the pre-attentive vision framework. A variety of technical and academic applications can benefit from such pre-classification. Our study was originally inspired by the need of getting an insight into the depths of the NASA data archive holding over 1 million of plasmagram images acquired by the Radio Plasma Imager (RPI) [8] onboard the IMAGE spacecraft [9]. Similar "information avalanche" situations can be found in many other Earth and space physics projects where sheer volume of data precludes their manual analysis (e.g., IVOA alliance virtual astronomical observatories, EOSDIS Earth observing sensor network, etc.). There are at least two other application domains where ability of the pre-attentive model to call attention to particular imagery features is useful, namely, real-time 24/7 monitoring and warning systems that base their decisions on imagery data, and autonomous and collaborative explorers such as the Mars rover. In all these applications the pre-attentive vision model solves scientifically viable, significant problem of replicating intelligent tasks of feature extraction.

Our major research interest was then to explore the potential and limitations of the strictly bottom-up perceptual analysis of images where lower levels of the Marr's pyramid are unaware of the considerations happening at higher levels. We tailored our studies to a real-life imagery dataset collected by the RPI instrument.

3. RADIO PLASMA IMAGER DATASET

The RPI instrument on IMAGE spacecraft is currently obtaining radio remote-sensing data about the density distribution of magnetospheric plasmas. RPI's chief product is the plasmagram, as shown in Fig. 1. The figure shows received signal strength (color scale) as a function of echo delay (range in vertical scale) and radio-sounder frequency (horizontal scale) of the radar pulses. Radar echoes from important magnetospheric structures, such as the magnetopause and the plasmapause, appear as traces on plasmagrams (thin lines observed above 500 kHz in Fig. 1). Plasmagram traces are intermixed with vertical signatures corresponding to the locally excited plasma resonances (e.g., intensification near 450 kHz in Fig. 1) and various natural emissions propagating in space. Less than 20% of all plasmagrams contain echo traces because RPI is a radar of opportunity: for its 10 Watt signal to reflect at a remote location as far as 40,000 km away, return to the spacecraft location, and appear above the noise level to be detected, a number of conditions needs to be satisfied. Plasmagram data exploration is a major exercise requiring a substantial manual labor expense.

Traces, the key plasmagram signatures of interest, have the following properties that make their recognition difficult:

- traces are x-monotonic, thin (often just 1 pixel wide) lines of varying shape and length,
- traces can be faint and sketchy; their brightness varies with frequency (x ordinate),
- leading edge of a trace can jitter,
- traces can intersect and come close to each other; weak traces can be found near strong traces.

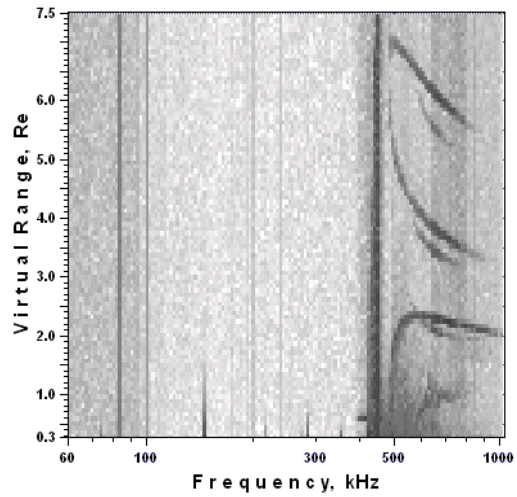

Fig. 1. IMAGE RPI sample plasmagram recorded at 03:22:33 UT on Sep. 17, 2001, showing signal reflections from remote plasma locations (dark traces) intermixed with stimulated resonances in the local plasma (vertical lines) and natural radio emissions (vertical bands).

4. PRE-ATTENTIVE MODEL FOR PLASMAGRAMS

Perceptual grouping under restriction of the Gestalt principle of good continuation (Fig.2) is the appropriate choice for this task.

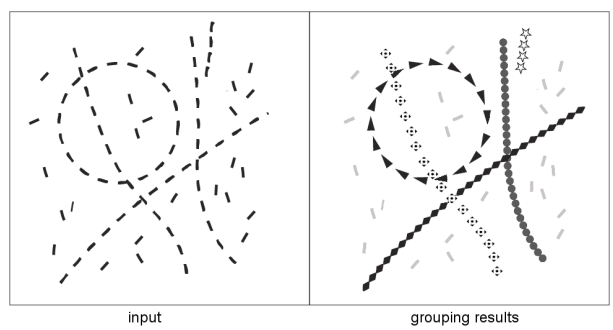

Fig. 2. Perceptual grouping under restriction of the Gestalt principle of good continuation. Input pattern (left) consists of oriented edge elements found by a detector sensitive to intensity gradients in the image. Grouping results (right) show found salient contours.

Left panel of Fig 2 shows a synthesized pattern of *oriented edgels*, edge elements found by locating sharp intensity gradients in the image and evaluating their local orientation. Right panel shows edgel grouping results that identify 5 contours in the input pattern. The grouping procedure is governed by calculations of the "saliency", first coined in [3] as a particular measure of length and smoothness of a contour. Biologically plausible models of saliency calculations [4]-[5] suggests that interaction between edgels depends on their mutual orientation and distance, and this dependence has different characteristics in three zone types (Fig.3).

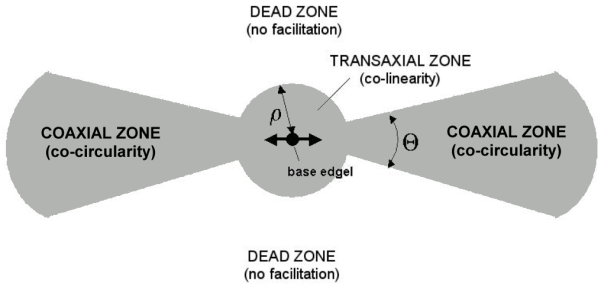

Fig. 3. Zones of edgel interaction for saliency calculations.

In the coaxial zone the saliency is highest for edgels that lie on the same arc (in agreement with the Gestalt constraint of co-circularity). The co-circularity constraint frequently appears in the bottom-up visual perception models [2]-[6], [13], [14]. Edgels outside the coaxial zone sectors do not contribute to the saliency measure, thus constituting two dead zones. Edgel interaction at close distances in the transaxial zone follows principle of co-linearity instead of co-circularity (i.e., saliency is highest for edgels that are parallel to the base edgel). Summary contribution from all edgels within the gray area of the pattern in Fig. 3 constitutes the likeliness for the base edgel to be a part of a contour.

After saliencies are calculated for all edgels in the image, the resulting saliency map is analyzed for presence of contours by a technique that is often referred to in the literature as "synchronization-desynchronization" [10] to reflect observed neural activity in the cortical networks of the brain. This task requires elements of attention-driven analysis if the image contains multiple features. Feature binding schemes [2] can replicate operations of attention switching and inhibition by introducing multiple layer models where separate features fall into different layers and thus become isolated from each other.

Early experimenting with the pre-attentive model on the RPI plasmagrams showed that they create a few problems for the classic approach:

1) Local calculations of edgel orientation are frequently wrong because of the low signal to noise ratio and distortions of the echo shape.
2) Jitter of the leading edge of contours smears the saliency map.
3) Short traces are indistinguishable in the saliency map if there are stronger nearby traces.
4) Application of smoothing filters for protection of the edge detector from excessive false alarms damages thin traces.

A few improvements to the pre-attentive model were suggested previously [11] to help with plasmagram processing. Because locally evaluated orientations for edgels can be wrong for various real-life reasons, the decision is made to let them change their orientation under collective facilitation of edgels in a larger context area, because these local errors are only visible in a larger context. The oriented edgels are then called *rotors* to denote their ability to change orientation. Similar to other rotor models [12]-[14], the optimal orientations are found iteratively by means of a recurrent neural network that evolves into the global minimum of its energy.

Allowing rotors in a large context area to interact creates a problem of distinguishing a weak signature in the presence of a strong one. Saliency of the weak signature can be overwhelmed by off-optimal but cumulatively larger contributions from the stronger signature. Although this situation is not the scenario for which the pre-attentive vision is best, use of the bio-plausible pattern for interaction of rotors shown in Fig. 3 helps to counteract the problem by narrowing down interacting rotors to the coaxial zone sectors. Another advantage of using this pattern is a higher robustness to the leading edge jitter because of the co-linearity processing in the close-distance transaxial zone. In summary, we combined two concepts, recurrent optimization of rotor orientation from physics and pattern for rotor interaction from biology to create our technique. Using this combination, we developed the Cognitive Online Rpi Plasmagram Ranking Algorithm (CORPRAL) for prospecting the plasmagram archive. We briefly discuss our test results on synthesized and actual data next.

5. PERFORMANCE TESTS

All of available 1,1 million plasmagrams were processed to detect traces; over 200,000 of them are now labeled in the RPI mission database as containing echo signatures. We shall note that our model was developed under requirements of only a few seconds of processing time and a small memory footprint of a flight computer, so that we could target implementation in a spaceborne warning system. Thus two important simplifications were made: (1) raw image is subjected to adaptive thresholding of intensity to reduce the number of detected edgels entering the saliency calculations, and (2) only one analysis scale (i.e., size of the context area) is processed instead of a more appropriate and biologically plausible parallel multi-scale analysis. These simplifications allowed us to meet the requirements, but at the same time they must have become responsible for additional analysis errors. Typical samples of the false negative and false positive errors are shown in Fig 4, where left panel shows a faint trace that CORPRAL missed, and right panel illustrates a typical false trace that is not substantiated by the underlying image.

Robust performance of computer vision models against imagery data of highly variable nature is a great challenge. It is not unusual for human analysts to dedicate more than a few seconds to find the visualization settings that highlight the faint RPI echo signatures best. Remarkably, our database of derived RPI scientific products [15] holds these visualization settings as a valuable information. So, whereas testing on the synthesized edgel patterns [16] was very satisfactory, the real-life imagery data were processed with variable degree of success, ranging from the accuracy as high as 97% down to occasional low 80-ies and less. Both sensitivity and specificity could drop to 50-60% for selected mission periods. In part, this uneven performance is explained by a dramatic variability of science data stream that RPI produced over 4 years of its mission.

Closer analysis reveals that a large part of the CORPRAL problems is caused by the errors of edgel detection that propagate up the Marr's pyramid uncompensated. Current algorithm for edgel detection, based on adaptive 1D local context thresholding of image intensity, is not optimal for low signal to noise ratios. Our present and future work is concentrated on development of

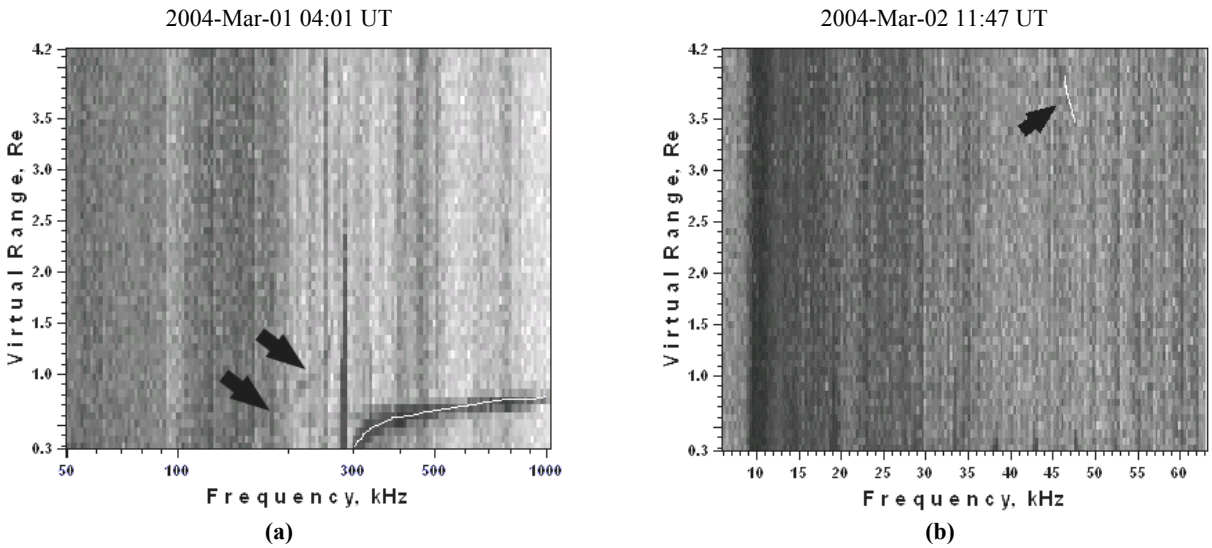

Fig. 4. Typical false negative (a) and false positive (b) errors of pre-attentive vision model.

fast and reliable 1D and 2D algorithms for robust detection of edge elements.

6. FUTURE WORK

It has been noticed that a typical RPI plasmagram has a noise background that is relatively constant with sounding range, but that can change fairly abruptly with frequency. We suggest that the performance of our pre-attentive model could be improved on RPI plasmagrams if this noise background could be detected and subtracted out from the data.

In our first attempt at finding the noise background we modeled the noise background as the median value for each frequency, i.e., the noise background was constant in range with a value equal to the median value over frequency at each range value. Somewhat better results were achieved when the 25% median value[†] was used instead of the usual 50% median value. Fig. 5b shows the result of subtracting the 25% median value from each column of an RPI plasmagram.

While this approach works well in preserving traces while reducing the noise background, it tends to remove resonances. Note in particular the resonance near 300 kHz is noticeably suppressed in Fig. 5b. To reduce this problem we turned to mathematical morphology. The usual use of mathematical morphology is to enhance a feature that is modeled by a specified "structuring element." In our case we wanted to do just the opposite. To better characterize the noise background, we wanted to suppress the traces and resonances found in the RPI plasmagram data.

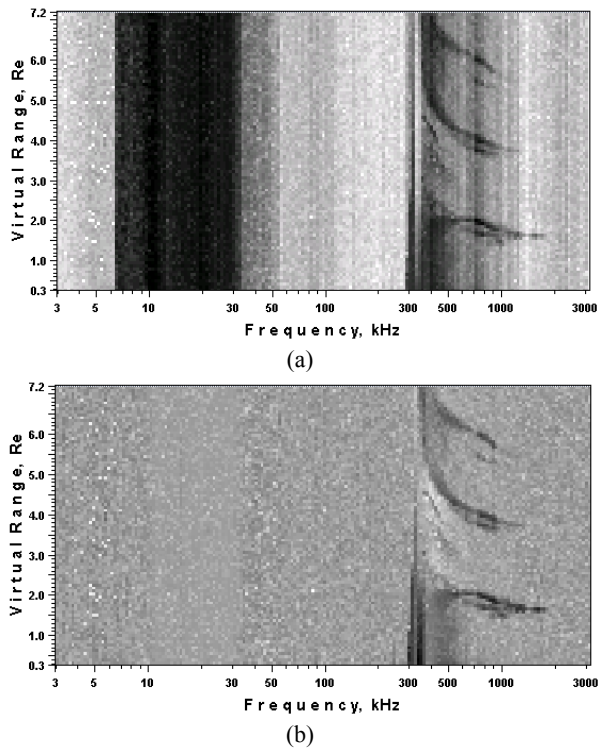

Fig. 5. Use of 25% median columnar filter to remove noise background from plasmagram images. (a) input image of RPI plasmagram taken on 2000-Oct-01 00:48 UT, (b) noise subtraction results.

We suppress traces and resonances by opening the data with a nxm structuring element in four different directions: 0º, 45º, 90º, and 135º. We then modeled the noise background as the 25% median value over frequency over all four directions. For RPI Plasmagrams of the type shown in Fig. 5, we found n=1 and m=15 works best. The noise background modeled in this manner

[†] The P% median value is the value of the P% largest value of a set of data sorted in order from smallest to largest. The usual median value of a data set is then the 50% median value.

is shown in Fig. 6a. The 25% median value selects values out of the morphologically opened images from directions that do NOT correspond to the best matches of the morphological structuring element with the RPI plasmagram resonance and traces. If a 75% to 100% median value was instead chosen, selection would have been made from directions that DO correspond to the best matches. The image that results from subtracting noise modeled in this way is shown in Fig. 6b. Note that the resonance near 300 kHz is much better preserved.

Even better preservation of resonances can be achieved by segmenting the 25% median value morphologically processed image. A segmentation serves to smooth out the small-scale fluctuation in the background noise image across frequency resulting in a good model of the large-scale background noise. A Hierarchical Segmentation (HSEG) approach developed by Tilton [17]-[18] is well suited for this purpose.

HSEG is based on hierarchical step-wise optimization (HSWO) [19] and produces a hierarchical set of image segmentations based on detected convergences. HSWO is an iterative approach to region growing segmentation in which the optimal image segmentation is found at N regions, given a segmentation at N+1 regions. The segmentation hierarchy produced by HSEG hierarchy is a set of several segmentations of the same image at different levels of detail in which the segmentations at coarser levels of detail can be produced from simple merges of regions at finer levels of detail. This is useful for applications that require different levels of image segmentation detail depending on the particular image objects segmented.

HSEG detects convergences in the segmentation process by monitoring *tratio*, the ratio of the merge threshold at the current iteration to the previous iteration. If the value of *tratio* exceeds a user-specified threshold, the segmentation result from the previous iteration is selected to be a level in the segmentation hierarchy.

The large-scale background noise can be modeled as the region mean image from the segmentation level in the segmentation hierarchy that has the largest *tratio*. Such a region mean image is shown in Fig. 7a and the image that results from subtracting noise modeled in this way is shown in Fig. 7b. Note that not only is the resonance near 300 kHz is well preserved, but other more subtle frequency fluctuations are better preserved. However, there appears to be some artifacts introduced around 30 and 60 kHz. Further work is required to address this issue. Perhaps all that is needed is an approach for selecting a more appropriate level from the segmentation hierarchy.

7. DISCUSSION

Many information-rich scientific projects spawn the Intelligent System applications that establish an automated clearinghouse for the dispersed and disorganized data. The computer plays powerful and enabling role in these projects where the size of data set precludes manual processing. We introduced a data prospecting tool for images based on a pre-attentive vision model that has been strengthened for improved robustness to the real-life imperfections of smoothness and good continuation of the

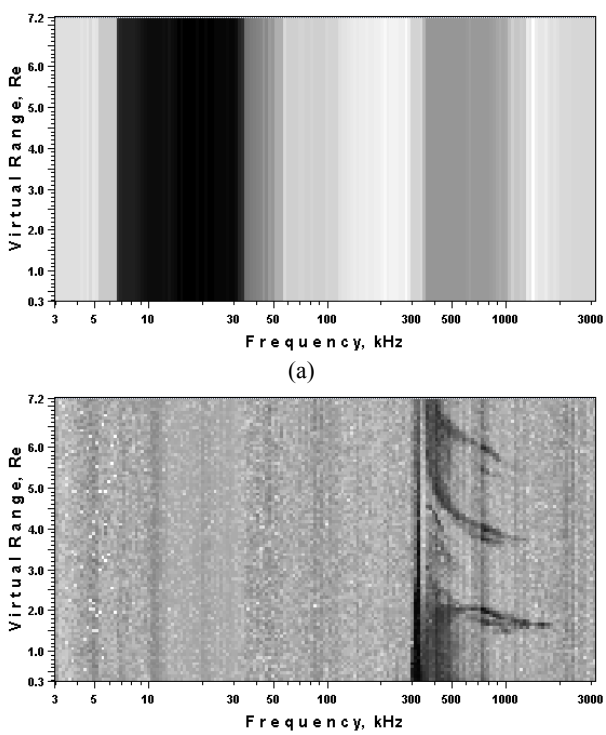

Fig. 6. Noise background determined by morphologic suppression of linear features in the image prior to application of 25% median columnar

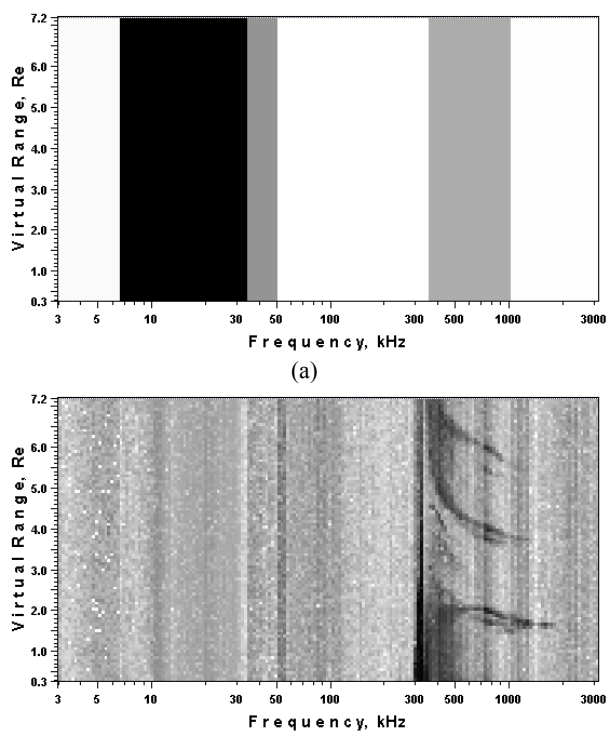

Fig. 7. Noise background determined by segmenting noise background in Fig. 6(a) by HSEG algorithm. (a) evaluated noise background, (b) noise subtraction results.

signatures to detect. Beside clear benefits of the automated exploration of RPI data that has dramatically reduced labor and time needed from the experts that browse plasmagram images to look for a new knowledge, the pre-attentive model provides new ways of getting new insights in the RPI scientific mission and instrument performance. It stands as a well-defined algorithm providing a uniform, consistent, repeatable view of the collected data of highly variable content and quality. This view is computationally tractable and, assuming that accuracy of the automated analysis is sufficiently high, it still captures the essence of data. This data model can be studied in explanatory, exploratory and predictive modes, as well as mined for previously unknown associations with other measurement variables such as the instrument orbital position and Geospace conditions.

REFERENCES

[1] I. Rock and S. Palmer, "The legacy of Gestalt psychology", **Scienific American,** Vol. SCA9012, 1990, pp. 84-90.

[2] H. Wersing, J.J.Steil, and H. Ritter, "A competitive layer model for feature binding and sensory segmentation of features", **Neural Computation**, Vol. 13, 2001, pp. 357-387.

[3] A. Shashua and S. Ullman, "Structural saliency: the detection of globally salient structures using a locally connected network", **2nd International Conference on Computer Vision**, 1998, pp. 321-327

[4] S.C.Yen and L.H. Finkel, "Extraction of perceptually salient contours by striate cortical networks", **Vision Research**, Vol. 38, 1998, pp. 719-741.

[5] G. Guy and G.Medoni, "Inferring global perceptual contours from local features", **International Journal of Computer Vision**, Vol. 20, 1996, pp. 113-133.

[6] P. Parent and S.W. Zacker, "Trace inference, curvature consistency, and curve detection", **IEEE Transactions**, Vol. PAMI-11, 1989, pp 823-839.

[7] D. Marr and H.K. Nishihara, "Visual information processing: Artificial intelligence and the sensorium of light", **Technology review**, Vol. 81, 1978, pp.2-23.

[8] B. W. Reinisch, et al, "The Radio Plasma Imager investigation on the IMAGE spacecraft", **Space Science Reviews**, Vol. 91, 2000, pp. 319-35.

[9] J. L. Burch (ed), **The IMAGE Mission**. Dordrecht: Kluwer Acad. Pub., 2000.

[10] Yen, S.C., E.D. Menschik, and L.H. Finkel, "Perceptual grouping in striate cortical networks mediated by synchronization and desynchronization", **Neurocomputing**, Vol. 26-27, 1999, pp. 609-616.

[11] I. A. Galkin, B.W. Reinisch, G. Grinstein, G. Khmyrov, A. Kozlov, and S. F. Fung, "Automated exploration of the Radio Plasma Imager data", **Journal of Geophysical Research**, Vol. 109, 2004.

[12] C. Peterson, "Track finding with neural networks", **Nuclear Instruments and Methods in Physics Research,** Vol. A279, 1989, pp.537-545.

[13] S. Baginyan, A. Glazov, I. Kisel, E. Konotopskaya, V. Neskoromnyi, and G. Ososkov (1994) "Tracking by a modified rotor model of neural network", **Computer Physics Communications**, Vol. 79, 1994, pp. 165-178.

[14] I.A. Galkin, B.W. Reinisch, G.A. Ososkov, E.G. Zaznobina, and S.P. Neshyba, "Feedback neural networks for ARTIST ionogram processing", **Radio Science**, Vol. 31, 1996, pp. 1119-1129.

[15] I.A. Galkin, G. M. Khmyrov, A. Kozlov, B.W. Reinisch, X. Huang, and G. Sales, "New tools for analysis of space-borne sounding data," **Proc. 2001 USNC/URSI Nat. Radio Sci. Meeting**, Boston, MA, 2001, p. 304.

[16] I.A. Galkin, "A pre-attentive vision model for automated data exploration applied to space plasma remote sensing data", **Ph.D Thesis**, University of Massachusetts Lowell, Lowell, MA, USA, 2004.

[17] J. C. Tilton, Analysis of hierarchically related image segmentations, **Proc. IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data,** Greenbelt, MD, USA, October 27-28, 2003.

[18] J C. Tilton, "A method for recursive hierarchical segmentation which eliminates processing window artifacts," **Disclosure of Invention and New Technology (Including Software): NASA Case No. GSC 14,681-1 (revised)**, NASA's Goddard Space Flight Center, Jan. 24, 2003.

[19] J-M. Beaulieu and M. Goldberg, "Hierarchy in picture segmentation: A stepwise optimal approach," **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Vol. 11, No. 2, pp. 150-163, Feb. 1989.