A NOVEL HEURISTIC ALGORITHM FOR FINDING SPARSE, THIN CURVED LINES IN IMAGERY WITH STRUCTURED NOISE.

N. T. Dutton[1], M. A. Mendlovitz[1], F. S. Turner[1], G. W. Patterson[1], [1]The Johns Hopkins University Applied Physics Laboratory; Laurel, MD (nicholas.dutton@jhuapl.edu).

Introduction: Many problems in digital signal processing, specifically image processing and computer vision, require object identification of curved lines. Examples include medical imaging applications such as x-ray angiograms, Intelligence Surveillance and Reconnaissance applications for finding roads and rivers, and – in the particular case of this work – finding the direct-path signal in bistatic Synthetic Aperture Radar (SAR) pulse-compressed data (i.e., waterfall plots).

Signal localization of lines in noisy imagery [1] is a broad topic. Edge detection techniques such as the Sobel or Canny operators, and feature extraction techniques such as the Hough transform exist and are well represented in literature, however with large datasets they can often be computationally expensive and also require customization for the specific application. This fine tuning can become cumbersome and time consuming especially when the imagery content significantly varies between datasets.

To mitigate some of these issues, the authors developed a general purpose solution to isolating thin, curved lines of variable shape and intensity in imagery that includes complexly structured and/or similarly shaped noise. We follow with an example application using bistatic SAR data in which the approach was used with excellent results.

Problem Statement: The problem presented is to create an algorithm that will completely automate finding and fitting of a vertical, thin, curved line of varying intensity and shape in the presence of structured noise. For this application, the vertical, thin, curved line represents the direct-path signal from a ground station transmitting through the back of the Mini-RF antenna onboard the NASA Lunar Reconnaissance Orbiter (LRO). The following is a list of constraints on the waterfall plots:

1) The structured noise in the imagery may cross the direct-path throughout the image. When this occurs the direct-path intensity may be less than, equal to, or greater than the crossing noise source at different pixels within the crossing region. 2) A minimum of two structured noise sources are present in the imagery, but there may be others. 3) The direct-path may not be visible in portions of the image. 4) The imagery may have horizontal strips where no data exists. 5) The structured noise may also be similar in shape and intensity to the direct-path. When this occurs, we cannot use traditional outlier rejection techniques such as the Hampel filter.

Solution Approach: The direct-path is usually distinguishable from the structured noise sources to the trained eye. One is inclined to recreate a similar process by which a human locates this thin line in the imagery.

First, we seek an algorithm to find the "sharpest" peak in intensity for each row (or column if the signal is generally horizontal in nature) of the imagery. The intensity values for each row across the sample (column) axis $I(x)$ are collected from the imagery. For each row, an average intensity count is then computed in a buffer window, $I_{avg}(x \pm \Delta x_{buffer})$, on each side of every element in $I(x)$. The index, x, with the highest ratio of the peak intensity value to the average buffer intensity, $I(x)/I_{avg}(x \pm \Delta x_{buffer})$, is the sharpest peak. This is the signal's estimated location on the sample axis for the given row. Using this buffered average approach reduces the likelihood that a signal will be selected purely because it is bright and selects thin vertical lines in the image over thicker ones. This process is repeated for every row in the image and the sharpest peak location is stored for later use.

Second, the mode of all of the sharpest peak x-coordinate locations provides the best estimate of the signal's centroid location on the sample axis.

Third, a rectangular subsection of the image is taken, centered on the centroid sample provided by the mode, and at the top row in the image. This subsections size is chosen such that the signal will always be contained within it. The signal's sample location is then estimated within the rectangular subsection by summing samples across rows and using the algorithm discussed in the first step. The estimated coordinates are then placed at the center row and sharpest peak sample of the subsection. This process steps through rows and is repeated until the full image has been spanned in 1 dimension.

Fourth, once the estimated x-y coordinates of the signal are stored for the full (vertical) image strip, outliers must be checked to ensure the estimates do not drift radically. These outliers can occur if one of the structured noise sources in the image crosses or is in close proximity to the signal. The outliers are removed by checking that the angle between adjacent signal pixel x and y coordinate estimates is approximately 90 degrees – i.e. a locally vertical signal. This initial filter is applied twice to minimize the amount of jump discontinuities in the sample space.

Finally, in several cases there happen to be similarly structured data point estimates that are not part of the signal, but ultimately show up due to the noise sources having a similar linear structure as shown in Figure 1.

When this occurs, we apply a newly developed heuristic algorithm we refer to as *Maximum Segment, Minimum Residuals* (MSMR).

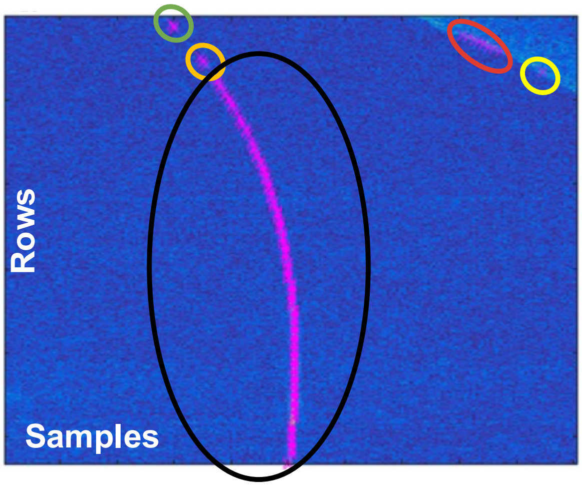

Figure 1. Signal estimated locations. Similarly structured noise causes false estimates in the upper right. MSMR segments are shown as ellipses.

The MSMR heuristic recognizes the need to maximize the data available for the current fit, otherwise regions could become sparse and fitting may be poor. It also ensures we are not including data in a problematic region (i.e. samples that are not signal). The following steps define the MSMR approach:

1) Categorize "segments" of data where the end of a segment is a jump discontinuity in the dependent variable (sample dimension).
2) Generate all possible combinations of segments, assuming the largest segment is always present (the largest ellipse in Figure 1). This results in $2^{N-1} - 1$ combinations of segmented estimates in x-y pixel space.
3) For each combination of segments, fit a low-order polynomial that just under-represents the nature of the curvature and store the residual norm of the fit.
4) The segment combination that maximizes the ratio of the number of segments used in the test to the residual norm for that test is the best estimate for the correct signal estimates.

The signal estimated locations can then be fit to a higher order polynomial and reapplied to the image as necessary. We see the MSMR algorithm works perfectly in the given example as shown in Figure 2.

Application: The LRO Miniature Radio Frequency (Mini-RF) dual-polarized synthetic aperture radar (SAR) instrument is operated in concert with the Arecibo and DSS-13 observatories to collect bistatic radar data on the nearside of the lunar surface [2]. When processing the bistatic radar data, the direct-path signal is the signal received by Mini-RF directly from the Earth-based transmitter and is used to improve timing and phase reconstruction of the chirped signal.

Several factors make automating the isolation of the direct-path in pulse-compressed data difficult. The Arecibo and DSS-13 observatories operate at different transmit powers and frequencies; 200kW in the S-band and 80kW in the X-band respectively. This leads to different signal strengths received by the radar and in general the direct-path may not be the strongest signal in the radar waterfall plot. The geometry and signal characteristics of the bistatic collect also influence the specular forward scatter and diffuse backscatter signal returns. Situations can arise where the direct-path signal approaches (or overlaps) the forward and/or backscatter. The algorithm described above is applied to the Mini-RF waterfall plots in order to locate the direct-path.

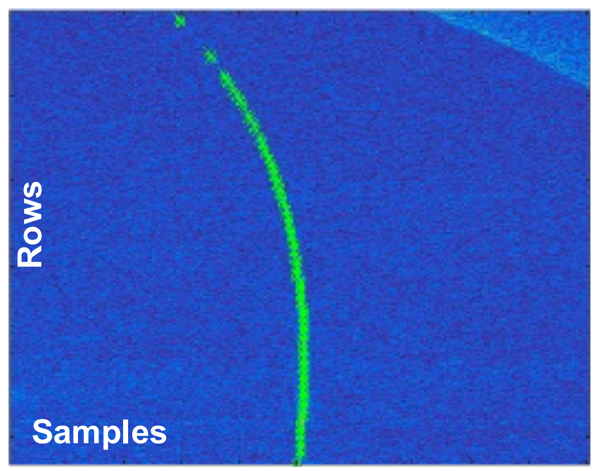

Figure 2. Direct-path Estimate and Fit after MSMR is applied

Conclusions: This work follows as an improvement to the previous techniques described in [1]. Previously, we were able to successfully identify the direct-path in >95% and ~80% of the S-band and X-band collects respectively. With the new algorithmic approaches discussed in this paper, we're now able to achieve a 100% success rate in finding and fitting the direct-path for both S-band and X-band collects. This achievement allows the radar processing workflow to remove the human in the loop.

References:

[1] Dutton, et al. (2019) LPICo2151.7078D
[2] Patterson G. W. et al. (2017) Icarus, *283,* 2–19.