

Intrinsic Bayesian Algorithm for Shortwave Voice Noise Reduction(IBA-SVNR) version 1.0

By Joshua Rainstar(Pseudonym)

The Intrinsic Bayesian Algorithm or IBA employs novel approaches to the estimation of noise presence as well as novel statistical measures for the filtration of single-sideband amplitude-modulated bandwidth-limited voice signals as typically is demonstrated in radio communications environments ranging from VLF(very low frequency) to UHF(ultra-high frequency), including amateur radio bands, military communications, and first responder/law enforcement systems, although the use of high-power digital modes with error correction has greatly replaced the use of single-sideband voice communication.

The algorithm was developed by a single author without theoretical considerations or a formal academic basis and with an empirical approach to algorithmic optimization, as such it represents novel work which is computationally effective and optimized for the specific use case it is intended for, and cannot readily be generalized to other environments or uses.

The algorithm requires a specific set of pre-defined values as constraints:

Audio: Firstly, audio, sampled at 8192 samples per interval of algorithm execution with a rate of 48000 Hz must be provided as an input. The algorithm cannot be readily generalized to other sample rates, although it can be roughly approximated for 24khz. The algorithm returns a noise-filtered series of equal size.

Window size: The FFT windowing size must be the sample rate/93.75, or 512 samples at 48000 Hz.

STFT hop: The hop of the short time Fourier transform must be $1/4^{\text{th}}$ the windowing size.

Hann, Synthesis(Hann), and Logit Windows: Three windows must be supplied to the algorithm: A Hann or raised cosine bell window of 512 elements, an optimized inverse synthesis window derived from the Hann window such as that described in D. Griffin and J. Lim, Signal estimation from modified short-time Fourier transform, IEEE Trans. Acoustics, Speech, and Signal Process., vol. 32, no. 2, pp. 236-243, 1984., and a unique window which constitutes a sub-algorithm, normalized between 1 and 0, consisting of two logistic functions, one reversed, with the end-point values which normally constitute infinity being replaced with the reflected padded value of twice the next to last minus the second from last.

Noise Constant: A noise similarity constant, default to $1/10^{\text{th}}$ of the Euler–Mascheroni constant, and truncated to two decimal places for most intents and purposes.

Number of Frequency Bins: The bandwidth the algorithm effectively operates on, defaulting to 36 bins, or a filter width close to 3400hz, and effectively usable from 32:257 bins

Correlation Normalization Factor: a number generated for use in normalizing a comparison, which must be altered for the number of frequency bins, and which defaults to 0. 6091672572096941

Logistic Distribution: a normalized logistic distribution generated as if for the Logistic window, of size equal to the number of frequency bins, starting at 0 and ending with 1.

Sawtooth: an odd-sized filter, consisting a series of equal-spaced values from 0 to 1 and back to 0

[illegible][illegible]

correlation normalizations given from 32.257 to 1.059882432365164 (345993860083043658, 0.60275453747893, 0.606012453304209, 0.609167574492269, 0.6122169028112414, 0.616516769342915, 0.61825392102714, 0.620057280006214, 0.623482207895196, 0.626900419854999, 0.62826470335797106, 0.6315984029236633, 0.634996488886294, 0.637581380639618, 0.6380801831066941, 0.640353080758548, 0.642462407854978, 0.64468552579166, 0.64682534251987, 0.648631683297601, 0.650595286898764, 0.652507942335818, 0.654380460146234, 0.656127502310614, 0.65801748149876, 0.6598277203770326, 0.6617583963971474, 0.6637014836980941, 0.6648254089817681, 0.666440572737036, 0.668031649136775, 0.669846962876037, 0.67171025658374, 0.673677092134274, 0.675670896246433, 0.677635830446236, 0.679636632631673, 0.678352156685459, 0.679739477118866, 0.680170246825447, 0.682498457930021, 0.6839865848844801, 0.686246761293761, 0.687498611012182, 0.688845210000000, 0.690345210000000, 0.691845210000000, 0.693345210000000, 0.694845210000000, 0.696345210000000, 0.697845210000000, 0.699345210000000, 0.700845210000000, 0.702345210000000, 0.703845210000000, 0.705345210000000, 0.706845210000000, 0.708345210000000, 0.709845210000000, 0.711345210000000, 0.712845210000000, 0.714345210000000, 0.715845210000000, 0.717345210000000, 0.718845210000000, 0.720345210000000, 0.721845210000000, 0.723345210000000, 0.724845210000000, 0.726345210000000, 0.727845210000000, 0.729345210000000, 0.730845210000000, 0.732345210000000, 0.733845210000000, 0.735345210000000, 0.736845210000000, 0.738345210000000, 0.739845210000000, 0.741345210000000, 0.742845210000000, 0.744345210000000, 0.745845210000000, 0.747345210000000, 0.748845210000000, 0.750345210000000, 0.751845210000000, 0.753345210000000, 0.754845210000000, 0.756345210000000, 0.757845210000000, 0.759345210000000, 0.760845210000000, 0.762345210000000, 0.763845210000000, 0.765345210000000, 0.766845210000000, 0.768345210000000, 0.769845210000000, 0.771345210000000, 0.772845210000000, 0.774345210000000, 0.775845210000000, 0.777345210000000, 0.778845210000000, 0.780345210000000, 0.781845210000000, 0.783345210000000, 0.784845210000000, 0.786345210000000, 0.787845210000000, 0.789345210000000, 0.790845210000000, 0.792345210000000, 0.793845210000000, 0.795345210000000, 0.796845210000000, 0.798345210000000, 0.799845210000000, 0.801345210000000, 0.802845210000000, 0.804345210000000, 0.805845210000000, 0.807345210000000, 0.808845210000000, 0.810345210000000, 0.811845210000000, 0.813345210000000, 0.814845210000000, 0.816345210000000, 0.817845210000000, 0.819345210000000, 0.820845210000000, 0.822345210000000, 0.823845210000000, 0.825345210000000, 0.826845210000000, 0.828345210000000, 0.829845210000000, 0.831345210000000, 0.832845210000000, 0.834345210000000, 0.835845210000000, 0.837345210000000, 0.838845210000000, 0.840345210000000, 0.841845210000000, 0.843345210000000, 0.844845210000000, 0.846345210000000, 0.847845210000000, 0.849345210000000, 0.850845210000000, 0.852345210000000, 0.853845210000000, 0.855345210000000, 0.856845210000000, 0.858345210000000, 0.859845210000000, 0.861345210000000, 0.862845210000000, 0.864345210000000, 0.865845210000000, 0.867345210000000, 0.868845210000000, 0.870345210000000, 0.871845210000000, 0.873345210000000, 0.874845210000000, 0.876345210000000, 0.877845210000000, 0.879345210000000, 0.880845210000000, 0.882345210000000, 0.883845210000000, 0.885345210000000, 0.886845210000000, 0.888345210000000, 0.889845210000000, 0.891345210000000, 0.892845210000000, 0.894345210000000, 0.895845210000000, 0.897345210000000, 0.898845210000000, 0.900345210000000, 0.901845210000000, 0.903345210000000, 0.904845210000000, 0.906345210000000, 0.907845210000000, 0.909345210000000, 0.910845210000000, 0.912345210000000, 0.913845210000000, 0.915345210000000, 0.916845210000000, 0.918345210000000, 0.919845210000000, 0.921345210000000, 0.922845210000000, 0.924345210000000, 0.925845210000000, 0.927345210000000, 0.928845210000000, 0.930345210000000, 0.931845210000000, 0.933345210000000, 0.934845210000000, 0.936345210000000, 0.937845210000000, 0.939345210000000, 0.940845210000000, 0.9423452

```
sawtooth filter = [0.014285714, 0.28571429, 0.42857143, 0.57142857, 0.71428571, 0.85714286, 1.0, 0.85714286, 0.71428571, 0.57142857, 0.42857143, 0.28571429, 0.14285714, 0.]
```

```
sawtooth filter = [0.014285714, 0.28571429, 0.42857143, 0.57142857, 0.71428571, 0.85714286, 1.0, 0.85714286, 0.71428571, 0.57142857, 0.42857143, 0.28571429, 0.14285714, 0.]
```

A number of requisite statistical measurements are required, as well as a number of generic other algorithms:

Sort: a sorting algorithm, which sorts smallest to largest

Outlier removal: an algorithm for identifying sequences smaller than a specified length of a specific value which operate on integers, and replacing them with a different value

Longest Consecutive: an algorithm which finds the longest consecutive series of the value 1 in a 1d array, and returns the size of this series

Convolution: a method which convolves two arrays together using "same" padding.

STFT analysis: a method which produces a suitable STFT representation of audio

Synthesis: an overlap-add external program which maintains a buffer and produces an output of audio samples given a STFT representation

Audio-samples: a buffer which maintains an internal storage of [8192*3] samples, is persistent between successive applications of the algorithm, and which translates the internal buffer by forgetting the first samples and appending newer samples on the end, to provide the algorithm with 8192*3 or approximately 170 milliseconds of samples.

Statistic methods distinct to this algorithm:

Maximal Average: given an input, excluding all elements which have overflowed values from the calculation, find the median of the absolute form of the product of the subtraction of the median of the nonzero values from the input

Absolute True Deviation: find the square root of the mean of the square of the absolute form of the input series minus the Maximal Average of the series

Threshold: Add the median of the nonzero values to the Absolute True Deviation

Pearson Correlation Coefficient: For N equal to the input size, return a single value

Provided that all of the above is available to the program, the following temporary variables and arrays must be initialized and provided to the algorithm as well:

Unless otherwise specified, all values are to be 64 bit floats and all mathematical operations using float precision. For environments where audio samples are in a different format, they must be modified to a 64 bit version and then modified back by an external algorithm.

complex_stft = [192,257] # a complex128 buffer for stft values.

residue_ #just like complex_stft

zeros_ #just like complex_stft but it's always zeros

stft_real= #just like complex_stft but a float buffer and not complex 128 bit values

stft_real_smoothed= #just like stft_real

mask = #just like stft_real

previous_mask= #just like stft_real

entropy_unmasked = [192] #float values

entropy_smoothed = [192+12] #float values

entropy_integer = [192] #int values

threshold = #float, initialized to 0

Provided that all of the previously stated requirements are in place, the algorithm described as Intrinsic Bayesian Algorithm or IBA is provided as follows, operating on each consecutive sequence of 8192 samples until program termination:

1. Shift the audio array buffer by 8192 samples and append the input. Zero out all temporary buffers.
2. Obtain the complex STFT representation of the audio on all samples using the LOGIT window.
3. Iterating over each row representing all frequencies at a point in time:
 - 3.1. produce the absolute values for each row and store in stft_real.
 - 3.2. subtract 1 from the Pearson Correlation Coefficient of the logit array and the sorted, normalized absolute values and place in entropy_unmasked,
 - 3.3. shifted to start at the 6th element of entropy_unmasked. Replace nan values with 0.
4. convolve entropy_unmasked with an array of [1.,1.,1.,] and divide the result by 3, placing in entropy_smoothed.
 - 4.1. copy entropy_smoothed into entropy_integer starting at the 6th element, thresholding the values to either zero or one depending on if they exceed the noise constant. If they exceed, set to 1. Increment a counter by 1 for every element set to 1.
 - 4.2. If the counter remains zero at the end of 4.1, synthesize zeros[RETURN EARLY]
 - 4.3. Sum the subset of entropy_integer[32:128+32] to form a sum. find the longest consecutive run in the same subset.
 - 4.4. if the sum is less than 22 and the streak is less than 16, synthesize zeros [RETURN EARLY]
5. using an Outlier removal algorithm on entropy_integer, remove runs of less than or 6 values long which are 0, setting them to 1.
 - 5.1. using an Outlier removal algorithm on entropy_integer, remove runs of 1 less than or 2 long and set them to 0, only after removing 0s.
6. find the threshold of stft_real. find the maximum of stft_real.
7. Take the transposition of stft_real. Pad the end of each row of the transposition with 30 zeros.
 - 7.1. convolve each row frequency slice with the sawtooth filter. Divide the result by 7, remove the padding, reverse the transposition and store in stft_real_smoothed.
8. for each value equal to 1 in entropy:
 - 8.1. Take the corresponding elements from zero to the frequency bin constant[0:36] of the index corresponding row of stft_real_smoothed, truncated to the Frequency Bins.
 - 8.2. Take the value of A = add together the Absolute True Deviation and Maximal Average of said row contents.
 - 8.3. Take the value of B = absolute of ((entropy_unmasked[index]/entropy_maximum) - 1)
 - 8.4. Take the value of C = threshold*b
 - 8.5. If C = NaN, C = A, otherwise V = (C + A)/2
 - 8.6. For each element in row up to frequency bins, where stft_real_smoothed[row,element] > c:
 - 8.7. Mask[index,element] = 1
 - 8.8. stft_real[index,element] = 0

The algorithm continues with steps 9, 10, and 11, to increase the completeness of the extraction which can be skipped for very low SNR signal.

9. : Find the maximum of stft_real and divide it by the previous maximum result in step #9.
 - 9.1. if the result is greater than 1.0, set to 1.0
10. repeat step 7.
11. for each value equal to 1 in entropy:
 - 11.1. Take the corresponding elements from zero to the frequency bin constant[0:36] of the index corresponding row of stft_real_smoothed, truncated to the Frequency Bins.
 - 11.2. Take the value of A = add together the Absolute True Deviation and Maximal Average of said row contents.
 - 11.3. Take the value of B = absolute of ((entropy_unmasked[index]/entropy_maximum) - 1)
 - 11.4. Take the value of C = threshold*b
 - 11.5. If C = NaN, C = A, otherwise V = (C + A)/2
 - 11.6. For each element in row up to frequency bins, where stft_real_smoothed[row,element] > C:
 - 11.7. if mask[index,element] == 0, mask[index,element] = 1 times the result of step 9
12. Perform step 7, but on mask, and store the result in mask.
13. do this three times:
 - 13.1. produce two copies of mask, one with the columns and rows transposed.
 - 13.2. for the first copy(not transposed) append 6 zeros to the beginning and end of each row.
 - 13.3. for the second copy, append 26 zeros to beginning and end of each row.
 - 13.4. convolve each row of the first copy with [1.,1.,1.] and divide by 3.0
 - 13.5. convolve each row of the second copy with [13x1.0 elements] and divide by 13.0
 - 13.6. transpose the second copy again.
 - 13.7. add together the two arrays, with padding stripped.
 - 13.8. divide the result by 2.
 - 13.9. store the result in mask
14. normalize the mask to 0,1
15. Replace the complex STFT values with the product of analysis using the Hann window on the audio.
16. Multiply the STFT complex values by the mask
17. Extract the center-most 64 rows of the STFT array and synthesize 8192 audio samples using the overlap-add approach and a modified synthesis window to minimize error
18. [RETURN] the audio samples to the parent program

This concludes the description of the algorithm. Other representations or forms which are pleasantly refactored to encapsulate subroutines in specific compiled machine code objects provide the suitable acceleration desired and afford this algorithm to perform within a real-time environment on most modern processor architectures(Intel Core 2 onwards) with minimal memory requirements.

Individual channels are independently processed by this algorithm, but the generalization to an amplitude-modulated dual-sideband setup is possible and simply entails producing a mirrored STFT representation and utilizing consensus-based approaches to estimating the entropy and the mask for both sidebands at the same time.