Contrast Analysis for Side-Looking Sonar

Daniel C. Brown
Applied Research Laboratory
The Pennsylvania State University
North Atherton Street
State College, PA 16801
phone: (814) 865-1193 fax: (814) 865-8069 email: dcb19@psu.edu

Daniel A. Cook
Georgia Tech Research Institute
Sensors and Electromagnetic Applications Laboratory
Smyrna, GA 30080
phone: (404) 407-8886 fax: (404) 407-7728 email: dan.cook@gtri.gatech.edu

Award Number: N00014-12-1-0045 (ARL/PSU) & N00014-12-1-0085 (GTRI)

LONG-TERM GOALS

The long-term goal of this project is to develop a high-fidelity model for side scan sonar image contrast that is well-documented and easy to use, thus encouraging broad application. This model is relevant to a wide range of Navy interests, some of which are listed below.

- ATR Enhancement: Provide contrast predictions to ATR algorithms for the purposes of improving ATR performance and estimating the expected reliability of the results. This type of collaboration will be used to help the hardware and image processing community better understand what drives ATR performance.

- Sonar Model Validation: The model provides a lower bound for shadow depth that can be used to validate modeling tools such as SWAT (Shallow Water Acoustics Toolkit).

- Adaptive Postprocessing: Tune image formation algorithms by trading image contrast against resolution in a predictable manner to achieve the best possible performance for any combination of bottom type and environmental conditions.

- Adaptive Mission Planning: Modify mission parameters (for example, track spacing and pulse duration) to optimize shadow contrast over the area to be mapped while using the minimum amount of battery power.

- Performance Comparison: The contrast model permits a direct comparison of real and synthetic aperture sonar performance. Therefore, this work could be used for developing system requirements for a variety of sonar systems.

The model will be validated using SAS data collected against a shadow-inducing target developed under this grant. Other data will be used as appropriate. The model will also be compared to predictions.
based on SWAT developed at the Naval Surface Warfare Center, Panama City Division.

**OBJECTIVES**

The contrast ratio is developed as a metric to predict a system’s capacity to form imagery with well-defined shadows. The depth of shadows in real aperture and SAS imagery is arguably as important to image quality as the focus of the backscattered returns. The maximum contrast that can be achieved is determined by a combination of hardware, software, and environmental parameters. The work performed under this grant investigates and seeks to quantify the impact of these parameters on shadow contrast.

**APPROACH**

This research is a collaborative effort between ARL/PSU and GTRI. Roughly speaking, GTRI is implementing the overall contrast model, while ARL/PSU is developing the most complicated single component of that model, which is a representation of the shallow water multipath. ARL/PSU has also designed and constructed a reusable sonar target that casts enhanced shadows. The boundaries across institutions are less distinct during the software integration and validation stages of the work.

Dan Brown is the PI at ARL, and he is working with Zack Lowe on the multipath model. Dan Cook is the PI at GTRI. He is being supported to varying degrees by Mike Davis, Brian Mulvaney, and Kristin Bing. Mike Davis is a recognized expert in the field of synthetic aperture radar performance prediction, and Brian Mulvaney is assisting with the data processing and software development. Kristin Bing is a co-author of a recent book chapter on radar ATR, and she is providing advice for linking the contrast model to the ATR community.

*Signal-to-Noise Equation*

The signal to noise ratio for a SAS image of the sea floor can be derived from the standard sonar equation for a single target with a given backscatter cross section (BCS). This result is then modified by including the effect of the coherent integration associated with SAS processing and modeling the sea floor as having a backscatter coefficient that has an effective BCS per unit area of terrain. The result is Eq. (1), where the constituent terms are defined in Table 1.

\[
\text{SNR} = \frac{G_{\text{rx}}P_{\text{tx}}}{4\pi r^2} \cdot \frac{\sigma^0}{4\pi r^2 \cos(\theta_{grz})} \cdot \frac{G_{\text{rx}}\lambda^2}{4\pi} \cdot \eta_{\text{lab}} \cdot \frac{1}{P_{\text{n,add}}} \cdot N_rN_{cr}
\]  

(1)

The term \(N_r\) in (1) represent the use of extended transmit waveforms to achieve gain via pulse compression, or matched filtering. The amount of waveform gain is given by the number of independent samples \(N_r\) used in the pulse compression. This number is equal to the length of the transmitted signal times its bandwidth, or its time-bandwidth product \(\tau_pB_p\). The number \(N_r\) is independent of the temporal sampling rate because the noise bandwidth is restricted to \(B\) by the matched filtering operation as well as any analog filtering that may occur prior to sampling. Therefore, an oversampled signal contains no additional independent noise samples.

Further gain comes from the fact that a SAS image is created by coherently integrating a sequence of \(N_{cr}\) pings as the sensor traverses the synthetic aperture, whose length is denoted as \(L_s\). Each ping has a different realization of the uncorrelated additive noise so, unlike temporal sampling, increased sampling
Table 1: Variables used to compute the sidescan imaging signal-to-noise ratio.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P_{\text{TX}}$</td>
<td>Transmitted power (W)</td>
</tr>
<tr>
<td>$G_{\text{TX}}$</td>
<td>Transmit gain (dimensionless)</td>
</tr>
<tr>
<td>$G_{\text{RX}}$</td>
<td>Receive gain (dimensionless)</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Wavelength (m)</td>
</tr>
<tr>
<td>$r$</td>
<td>Range (m)</td>
</tr>
<tr>
<td>$\theta_{\text{grz}}$</td>
<td>Grazing angle (rad or deg)</td>
</tr>
<tr>
<td>$\sigma_0$</td>
<td>Backscatter coefficient (dimensionless)</td>
</tr>
<tr>
<td>$\delta_r$</td>
<td>Cross-range resolution (m)</td>
</tr>
<tr>
<td>$\delta_{\text{cr}}$</td>
<td>Range resolution (m)</td>
</tr>
<tr>
<td>$P_{n,\text{add}}$</td>
<td>Additive noise power (W)</td>
</tr>
<tr>
<td>$N_r$</td>
<td>Time-bandwidth product (dimensionless)</td>
</tr>
<tr>
<td>$N_{\text{cr}}$</td>
<td>Space-bandwidth product (dimensionless)</td>
</tr>
</tbody>
</table>

in the along-track dimension does improve the SNR. The number of pings can be written in several ways, for example:

$$N_{\text{cr}} = \frac{L_{\text{sa}}}{\Delta_x \nu} = \frac{N_{\text{ch}} f_p L_{\text{sa}}}{L_{\text{eff}} \nu} = \frac{N_{\text{ch}} f_p \lambda r}{2 \delta_{\text{cr}} \nu}$$

(2)

where $L_{\text{eff}}$ is the effective transducer length implied by the -3 dB azimuth beamwidth (if the transmit and receive beamwidths are not the same, the narrower one determines the integration angle), $\Delta_x$ is the distance between spatial samples, $f_p$ is the PRF, $N_{\text{ch}}$ is the number of channels in the Vernier receive array, and $\nu$ is the platform speed.

Our development is expressed in terms of the peak power, which is the RMS power over the time of the transmitted signal. The peak power can be obtained directly from the more commonly specified sonar source level pressure, $p_{\text{rms}}$. According to Equation (3.1.12) of [1], this relationship is:

$$p_{\text{rms}} = \left( \frac{P_{\text{TX}} \rho c}{4 \pi r_0^2} \right)^{1/2}$$

(3)

where $r_0$ is the reference distance for the RMS pressure measurement. Lastly, we point out that the quantity analogous to (1) in the radar community is usually called the clutter-to-noise ratio (CNR) to indicate that the ground is the object of concern, as opposed to a discrete target in free space.

**Terrain-to-Shadow Contrast Ratio**

The figure of merit considered here is the contrast ratio, which is the ratio of the bottom backscatter plus noise divided by the noise alone. Carrara et al. [2] describe the distributed target contrast ratio (DTCR) to quantify the ability to visually distinguish between two adjacent types of terrain, one having high backscatter and the other having low. For sidescan sonar the low (i.e., zero) backscatter regions of interest are shadows. This leads to the quantity sought, the terrain to shadow contrast ratio, or more simply, the contrast ratio (CR) of the image:

$$\text{CR} = \frac{\sigma_0 + \sigma_{n,\text{add}} + \sigma_{n,\text{mult}}}{\sigma_{n,\text{add}} + \sigma_{n,\text{mult}}}.$$  

(4)
The value of $\sigma_0$ is understood to represent the lowest backscatter in the scene which is to be distinguished from a shadow. In the absence of other requirements, the average sea floor backscatter for a given bottom type provides a reasonable value for $\sigma_0$.

Certain values of the contrast ratio are of particular interest. First, we see that CR approaches 1 (0 dB) as the backscatter coefficient goes to zero. Depending on the bottom type, the CR can approach this case for grazing angles typical of long-range SAS imagery. A second useful case occurs when the total noise contribution is equal to the bottom scattering, resulting in CR=2 (3 dB). This is a possible starting point for determining minimum contrast requirements, although CR=10 (10 dB) is probably a better indication of what would be considered good quality imagery for visual inspection.

The following sections describe the constituent sources of additive and multiplicative noise relevant to SAS.

**Additive Noise**

Additive noise is that form of noise that would be present in the recorded signal even with the transmitter deactivated. The two most significant contributions are system electronic noise and ambient sea noise:

$$\sigma_{n,\text{add}} = \sigma_{\text{self}} + \sigma_{\text{ambient}}.$$  \hspace{1cm} (5)

At high frequencies the ambient noise is driven by thermal agitation of the water molecules, while at lower frequencies the ambient noise might be dominated by other natural or anthropogenic phenomena such as biologics, sea state, or shipping noise. For a well-designed synthetic aperture sonar the self-noise is below the noise floor established by the ambient noise.

The SNR given by (1) can be used to derive the backscatter coefficient implied by the noise floor of the sonar. This quantity is commonly used in the SAR literature, where it is called the ‘noise equivalent $\sigma_0$', or $\sigma_n$. It is also referred to as the noise equivalent sigma zero, in which case the acronym NESZ is often used. The value of $\sigma_n$ is found by setting the SNR equal to 1 and solving for the backscatter coefficient:

$$\sigma_n = \frac{(4\pi)^3 r^4 \cos(\theta_{grz}) P_{n,\text{add}}}{P_{tx} G_{tx} G_{rx} \delta_r \delta_{cr} \lambda^2 \eta_{ab} N_r N_{cr}}.$$  \hspace{1cm} (6)

This development parallels that used by the radar community. More information can be found in several sources, for example [2, 3, 4].

Other sources of additive noise, such as turbulent flow near the hydrophones, can be accounted for in the same way. Ross [5] describes these and many other sources of noise. When computing $\sigma_n$ in (6) we have discarded the absorption coefficient $\eta_{ab}$ since we are concerned with noise power measurements made at or very close to the hydrophone array. The absorption term would be retained when considering noise sources whose radiating power is known at some location far removed from the array.

**Multiplicative Noise**

Multiplicative noise is so named because it rises and falls in proportion to the strength of the backscattered signal. Its effect is independent of the transmitted signal power in the sense that doubling the transmitted power would double the multiplicative noise power resulting in no benefit to SNR. Surface reverberation is a good example of a multiplicative noise source: Increasing the transmitted power increases both the sea floor and sea surface backscatter by the same fraction. Other examples of multiplicative noise are range and along-track ambiguities, the sidelobes of the image impulse response,
quantization noise, volume reverberation, and multipath. Among these sources of noise, multipath is by far the most complicated to model accurately.

According to [2] the multiplicative noise ratio (MNR) may be approximated by combining the primary sources of multiplicative noise, which are the range and along-track ambiguity-to-signal ratio (ASR), the integrated sidelobe ratio (ISLR) of the image impulse response, and the quantization noise to signal ratio (QNSR). Application to sonar adds volume reverberation from the water column \( R_{\text{vol}} \), the sea surface \( R_{\text{surf}} \), and multipath \( R_{\text{mpath}} \). Surface reverberation and multipath must be considered for accurate estimation of the contrast for operation in littoral waters [6]. Because they are assumed to be statistically independent, the multiplicative noise terms are combined as follows:

\[
\sigma_{\text{n, mult}} = \bar{\sigma}_0 (\text{ASR} + \text{ISLR}) + \text{QNSR} + R_{\text{surf}} + R_{\text{vol}} + R_{\text{mpath}},
\]

where ASR and ISLR are modeled as being proportional to the average sea floor backscatter. If precise knowledge of the sea floor backscatter is available, this can be accounted for since the physical locations corresponding to the ambiguous regions are deterministic. If surface and volume reverberation and multipath are negligible, then QNSR is also proportional to \( \bar{\sigma}_0 \). Otherwise, it depends on the total backscattered power.

**Ambiguity-to-Signal Ratio**

The ambiguity to signal ratio (ASR) represents the superposition of along-track ambiguities (also called azimuth or Doppler ambiguities) and range ambiguities within the imaged scene. The ambiguities are assumed to be uncorrelated, allowing their total contribution to be represented by a summation. Range ambiguities are reflections from previous pings that are received at the same time as the current ping. Thus, the return at range \( r \) would be superposed with returns from ranges \( ncT/2 + r \), where \( T \) is the ping repetition period, \( c \) is the sound speed, and \( n \) is a positive integer.

Along-track ambiguities arise from the fact that the array elements are not spaced closely enough to avoid aliasing all spatial frequencies. A uniform linear array (ULA) consisting of omnidirectional elements must possess spacing finer than or equal to one-half wavelength in order to avoid spatial aliasing. SAS bypasses this constraint because the physical receivers are not omnidirectional. The transmit and receive beampatterns limit the range of spatial frequencies that can impinge upon the array. The central portion of the mainlobe acts as a bandpass filter and the remaining beampattern determines the aliasing that occurs. Viewed in this light the spatial aliasing behavior is significant since the sinc function is generally considered to be a poor bandpass filter, yet it is typical of sonar beampatterns.

The total ASR is given by the integral in Eq. (8), where \( G = G_{\text{TX}}G_{\text{RX}} \) is the composite two-way transmit receive gain for a particular spatial frequency \( k_x \), \( \sigma_0 \) is the backscatter coefficient of the terrain, \( B_p \) is the processed spatial bandwidth, \( k_{xs} \) is the spatial sampling frequency in rad/m \( (k_{xs} = 2\pi/\Delta x) \), \( \Delta x \) is the distance between along track samples (i.e., the phase center spacing), and \( f_p \) is the ping rate in Hz.

\[
\text{ASR}(\tau) = \frac{\sum_{m,n} \int_{-B_p/2}^{-B_p/2} \int_{-B_p/2}^{-B_p/2} G(k_x + mk_{xs}, \tau + n/f_p) \cdot \sigma_0(k_x + mk_{xs}, \tau + n/f_p) dk_x}{\int_{-B_p/2}^{-B_p/2} \int_{-B_p/2}^{-B_p/2} G(k_x, \tau) \cdot \sigma_0(k_x, \tau) dk_x}
\]

The ASR is expressed as a function of time \( \tau \) (or equivalently, range \( R = c\tau/2 \)). The spatial frequencies used in Eq. (8) represent a single frequency from the transmitted spectrum. For wideband sonars this
Table 2: Properties of common 2D spectral weighting functions. The mainlobe broadening factor is relative to the mainlobe width of the uniform window impulse response, which is the 2D sinc in the image domain.

<table>
<thead>
<tr>
<th>2D Window</th>
<th>ISLR (dB)</th>
<th>Mainlobe Broadening</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rectangular</td>
<td>-6.5</td>
<td>1</td>
</tr>
<tr>
<td>Taylor (5/35)</td>
<td>-24.6</td>
<td>1.32</td>
</tr>
<tr>
<td>Hamming</td>
<td>-31.3</td>
<td>1.46</td>
</tr>
<tr>
<td>Hanning</td>
<td>-29.9</td>
<td>1.56</td>
</tr>
</tbody>
</table>

The equation would need to be integrated over the transmitted acoustic bandwidth with respect to temporal wavenumber \( k \). The processed bandwidth is the extent of the spectral region of support of the wavenumber \( k_x \) used for imaging. It is bounded above by the composite transmit-receive beamwidth, and it may be intentionally reduced in order to achieve better ASR performance. The cost of reducing \( B_p \) is a coarsening of the cross-range resolution of the image since narrowing the effective beamwidth shortens the synthetic aperture length. The resulting along-track resolution is given by \( \delta_{cr} = \frac{2\pi}{B_p} \).

The radar literature frequently describes the along-track ambiguities in terms of Doppler frequency, rather than spatial frequency as is done here. For broadside imaging the two are related by \( f_D = v k_x \).

For a fixed temporal frequency, the spatial frequency is proportional to the angle of incidence across the array \( \theta \):

\[
k_x = 2k \sin \theta.
\]

**Integrated Sidelobe Ratio**

The integrated sidelobe ratio (ISLR) is the ratio of the energy in the sidelobes of the two-dimensional image impulse response to the energy in the mainlobe. Like the ASR, it is a scale factor and has no reference quantity associated with its decibel representation. Under ideal conditions, the ISLR of the image impulse response is a function of the spectral support and weighting associated with the image reconstruction. However, errors due to miscalibration or motion can corrupt the impulse response causing the ISLR to deviate from the values given in Table 2. The impact of phase errors on the impulse response is described in Chapter 5 of [2] and in [7].

**Quantization Noise to Signal Ratio**

There are a number of models that can be used to describe quantization noise introduced by the analog-to-digital converter (ADC). The QNSR of an ideal converter is \(-(6.02N + 1.76)\) dB, where \( N \) is the number of bits; but it is rare in practice to achieve this level of performance [8]. A conservative rule of thumb found in [2] states that QNSR is given by -5 dB per bit used in the ADC. Like ASR and ISLR, the QNSR is a scale factor and has no reference quantity associated with its decibel representation.

**Reverberation**

The total scattered field of sound waves propagating in the ocean environment is known as reverberation, and it is generally categorized as being associated with the sea floor, the sea surface, or a volume within the water column. In imaging applications, the primary-reflection bottom reverberation is the signal of interest, and all other reflections are sources of multiplicative noise. In this presentation we ignore volume reverberation and combine unwanted surface and bottom reflections into a single multipath model, described below, such that \( RR_{mpath} = RR_{surf} + RR_{bottom} \).

A simple model has been developed [9] to predict the influence of individual multipath rays. This model
uses an isovelocity propagation environment with planar surface and bottom boundaries. This model considers rays up to second order, with a single diffuse reflection. The reasoning is that rays containing more than 2 specular reflections, or more than one diffuse reflection, are sufficiently attenuated so as to be inconsequential compared to the direct bottom backscatter. Each of the possible ray paths, shown in Figure 1, is identified using a naming convention that indicates the order of boundary interactions as well as the type of interaction [5]. The perturbation approximation is used to model the loss associated with scattering from both the sea surface and the sea floor. Sea-floor roughness is assumed to have a power law spectrum, and sea surface roughness is based on a Pierson-Moskowitz spectrum.

WORK COMPLETED

The key achievements completed to date are as follows:

- The overall theoretical contrast model is complete
- The multipath component of the contrast model has been developed, implemented in MATLAB, and it compares favorably with measured data (see Fig. 2 discussed below)
- The overall contrast model has been implemented in MATLAB using an open architecture that allows users to easily replace components of the model. For example, the default is to compute ideal sinc-like beampatterns, but module this could be replaced with code based on measured patterns
- Two types of shadow targets have been constructed and deployed for various SAS data collections. The first casts enhanced shadows, and the second maximizes forward scatter to create a true no-return area as opposed to a shadow
- Results to date were published at the 2013 1st International Conference and Exhibition on Underwater Acoustics in Corfu, Greece [10]

The main tasks remaining to be completed are:

- Complete the documentation that will accompany the contrast prediction model
Figure 2: The reverberation and multipath model predicts the received level well until the ambient noise floor is reached at 35 meters. “Medium Sand” bottom parameters are taken from [11] and the SAX99 bottom parameters are taken from [12, 13]

- Validate the model using measured sonar data and SWAT simulations
- Continue to publish the results in appropriate conferences and journals

RESULTS

An initial model-data comparison is shown in Fig. 2 using data collected with SSAM2 at a sea test in Panama City, FL near Shell Island. The raw data segment was taken from a single receive channel in the high frequency band. One-hundred consecutive pings were incoherently averaged to reduce speckle noise. The vehicle roll and environmental data (sound speed, depth, altitude) are recorded by the sonar system, and these are used to properly configure the model for a given segment of data. Two different sets of parameters were used to describe the bottom. The first is “Medium Sand” as defined in [11], and the second set of parameters were taken from measurements of the bottom collected during the SAX99 trials [12, 13].

Beyond 35 meters, additive noise dominates the received level. Since the reverberation model does not account for additive noise the model and data diverge. However, the model and data agree well in the regions that are not limited by additive noise. The scattering strength of the sea floor has a wide range of variability within a single sediment class. For example, the SAX99 and “Medium Sand” bottom types differ by more than 5 dB. It is not surprising then that the received level and the model results have some differences in absolute level. This is likely due to a lack of knowledge regarding the specific bottom parameters necessary to accurately model the reverberation at the location where the data was collected.
IMPACT/APPLICATIONS

A model has been developed to estimate the shadow depth formed by a SAS accounting for the impact of sonar hardware design, image formation algorithm implementation and environment. This model is applicable to a range of problems including sonar system design, sonar performance prediction, and environmentally adaptive mission planning. Our current focus is on the validation of this model using data collected during sea trials.

The anticipated broad impact of this work is summarized above under “Long-Term Goals”. In this section, we identify several possible areas for applying the contrast model within the ATR community. The following ideas came from attending the 23 August session of the 2012 ONR MCM Virtual Program Review.

• In space-time adaptive processing (STAP) used for radar-based detection of ground moving targets, the Kullback-Leibler divergence equates to twice the signal to interference plus noise ratio (SINR), such that better SINR implies a greater distance between the interference plus noise distribution and the target or signal distribution. Since the goal of STAP is to maximize SINR, this divergence is a measure of the improvement in ROC characteristics due to STAP. Jose Principe’s talk, “Surprise Metric for Contact Fusion”, discussed measuring information gain between prior and posterior distributions using the K-L divergence. We wonder if there may be a relationship between this divergence and image contrast that can be exploited to predict the probability of correct classification in a given ATR scheme.

• A primary goal of our multipath model is to better characterize environmental factors for more realistic modeling and simulation of ATR inputs. We believe it could also be used to improve the selection of feature sets at given ranges and/or the selection of data being fed to the features. (See the talks “Performance Estimation of ATR Algorithms” by Cary Humber, NSWC-PCD and “Transfer Learning for ATR” by Jason Isaacs, NSWC-PCD.)

• It would be interesting to test information theoretic techniques, which are based on measures of entropy, in regions of direct path only and multipath environments. Such comparisons may be relevant to ATR algorithms that use the information content of the background to improve performance. (See the SBIR presentation “Incremental Knowledge Assimilation System” by Joseph Yadegar, Utopia Compression Corp.)

• Morphological component analysis can be used to reject unwanted clutter, but may not be performed due to computational costs. It uses a dictionary of bases, iterative matched filtering, and thresholding to estimate the components of a signal. Perhaps the contrast/multipath models could be used to identify a better set of bases, leading to fewer required iterations. (See the SBIR presentation “Underwater Target Detection and Classification with In-Situ Learning” by Neil Wachowski, CSU.)

TRANSITIONS

The bottom and surface reverberation models developed in the reverberation and multipath modeling portion of this program were applied to the development of performance models for the PMS-406 Knifefish program. These models are being integrated into the larger system performance model being developed by Dr. Terry Bazow at Metron, Inc. Discussions have also begun regarding the incorporation
of this work into SWAT and into the MSIPS (Modular SAS Image Processing Suite) and MATS (Modular Algorithm Testing Suite) testbeds developed by NSWC Panama City Division.

RELATED PROJECTS

There are no directly related projects at this time.

REFERENCES


PUBLICATIONS


• D. Cook, D. Brown, and Z. Lowe, Synthetic aperture sonar contrast, in 1st International Conference and Exhibition on Underwater Acoustics, June 2013, pp. 143–150.