DATA COMPRESSION FOR EMITTER LOCATION

Mark L. Fowler

Department of Electrical Engineering
State University of New York at Binghamton
Binghamton, NY 13902
mfowler@binghamton.edu

ABSTRACT

Locating emitters by cross-correlating received signals to compute their time-difference-of-arrival (TDOA) and the frequency-difference-of-arrival (FDOA) requires that signal data received at one platform be transferred to the other platform. Often the data link used has insufficient bandwidth to accomplish the transfer within the time requirement, and therefore use of data compression is needed. This paper outlines a useful progression in compression techniques from those that consider only mean square error to those that consider the true impact on the estimated TDOA/FDOA accuracy. Within this context, specific results are presented for two compression approaches for TDOA/FDOA systems. The first is the application of block-adaptive quantizers (BAQ) to the real and imaginary parts of the complex baseband signal to be transferred. The second uses a wavelet transform together with an adaptively allocated set of quantizers; also, certain wavelet coefficients can be eliminated with lower impact on the TDOA/FDOA accuracy than expected from a mean-square quantization point of view.

1. INTRODUCTION

An effective way to locate electromagnetic emitters is to measure the time-difference-of-arrival (TDOA) and the frequency-difference-of-arrival (FDOA) between pairs of signals received at geographically separated sites [1]-[3]. The measurement of TDOA/FDOA between these signals is done by coherently cross-correlating the signal pairs [2], [3], and requires that the signal samples of the two signals are available at a common site, which is generally accomplished by transferring the signal samples over a data link from one site to the other site. An important aspect of this that is not widely addressed in the literature is that often the available data link rate is insufficient to accomplish the transfer within the time requirement unless some form of lossy data compression is employed. For the case of Gaussian signals and noises, the results in [4] establish bounds on the rate-distortion performance for the TDOA/FDOA problem and compare them to the performance achievable using scalar quantizers, where distortion is measured in terms of lost SNR due to compression. However, these results are not applicable when locating radar and communication emitters because the signals encountered are not Gaussian.

The two noisy signals received are complex-valued baseband signals given by

$$\begin{aligned}\hat{s}(k) &= s(k) + n(k) \\ &= [s_r(k) + js_i(k)] + [n_r(k) + jn_i(k)] \\ \hat{d}(k) &= d(k) + v(k) \\ &= [d_r(k) + jd_i(k)] + [v_r(k) + jv_i(k)]\end{aligned}$$

where $s(k)$ and $d(k)$ are the complex baseband signals of interest and $n(k)$ and $v(k)$ are complex white Gaussian noises, each with real and imaginary parts notated as indicated. The signal $d(k)$ is a delayed and doppler shifted version of $s(k)$. The signal-to-noise ratios (SNR) for these two signals are denoted *SNR* and *DNR*, respectively[1]. To cross correlate these two signals one of them (assumed to be $\hat{s}(k)$ here) is compressed, transferred to the other site, and then decompressed before cross-correlation. After lossy compression, signal $\hat{s}(k)$ has SNR of SNR_q, and the SNR after cross-correlation is given by

$$SNR_{cc} = \frac{WT}{\dfrac{1}{SNR_q} + \dfrac{1}{DNR} + \dfrac{1}{SNR_q DNR}}$$

where *WT* is the time-bandwidth product (or coherent processing gain), with *W* being the noise bandwidth of the receiver and *T* being the duration of the received signal [3]. Note that the post-correlation SNR is dominated by the smaller of SNR_q and *DNR*. The accuracies of the TDOA/FDOA estimates are

$$\begin{aligned}s_t &\geq \frac{1}{2p\; B_{rms}\sqrt{SNR_{cc}}} \\ s_n &\geq \frac{1}{2p\; D_{rms}\sqrt{SNR_{cc}}},\end{aligned} \qquad (1)$$

where B_{rms} is the signal's rms (or Gabor) bandwidth in Hz, D_{rms} is the signal's rms (or Gabor) duration in seconds [3]; we

[1] SNR (non-italic) represents an acronym for signal-to-noise ratio; *SNR* (italic) represents the SNR for $\hat{s}(k)$.

will refer collectively to these two signal parameters as the signal's "rms widths".

In this paper we present results for two different compression approaches for TDOA/FDOA estimation. The first, called R/I quantization, applies block-adaptive quantizers (BAQ) to the real and imaginary parts of the signal [7], and the second uses a wavelet transform together with an adaptively allocated set of quantizers [8].

A BAQ technique has been applied to the compression of synthetic aperture array (SAR) signals [5], where block-adaptive quantizers designed for Gaussian signals are applied to the real and imaginary components. However, for the TDOA/FDOA case considered here, the signal's pdf is not known nor is it expected to be Gaussian. Thus, first, we derive bounds on the SNR performance of the BAQ technique when applied to radar and communication signals and investigate the impact on TDOA/FDOA estimation. Simulation results are compared to these bounds to show that they are indeed viable. Second, we develop a wavelet-based algorithm and demonstrate through simulations that it outperforms the R/I quantization approach for signals of interest. Finally, the wavelet-based method is briefly discussed in a setting that considers a trade-off between quantization and decimation for TDOA/FDOA emitter location. This idea is generalized to the following: certain wavelet coefficients can be eliminated with less impact on the TDOA/FDOA accuracy than expected from a mean-square quantization error point of view.

2. R/I QUANTIZATION METHOD

The R/I quantization methods decrease the number of bits used to represent signal samples by converting each sample to a form that is more coarsely quantized; this is done on a block-by-block basis by scaling the real and imaginary samples in a block by scale factors g_r and g_i, respectively, and then rounding to the desired number of bits. For analysis purposes we consider that the compressed signal is unscaled before cross-correlation. We consider two ways of scaling: (i) Multiply-Scaling: multiplying the samples in a block by an appropriate positive factor chosen to set to 1's all the non-sign bits of the largest-magnitude sample in the block, and (ii) Shift-Scaling: left-shifting the bits of the samples in a block by an amount chosen to shift the most-significant 1 of the largest-magnitude sample in the block into the MSB position. Thus, shift-scaling is multiply-scaling with a scaling factor that is a power of two.

The quantized signal can be considered to have a $B = b + 1$ bit representation of the form $S_\bullet B_1 B_2 B_3 \cdots B_b$, where S denotes the sign bit, $\bullet$ denotes the binary point, and the B_i's denote the b bits used to represent the magnitude. Then $s_q^2 = 2^{-2b}/12$ is the variance of the quantization noise [6]. After scaling and quantizing the real part of $\hat{s}(k)$, we get

$$\tilde{s}_r(k) = s_r(k) + n_r(k) + e_r(k)/g_r$$

where $e_r(k)$ is the quantization noise; a similar expression for the imaginary part can be written. Using this and defining the signal peak-to-rms parameters

$$a_r = \frac{\max\{|s_r(k)|\}}{\sqrt{P_{s_r}}} \qquad a_i = \frac{\max\{|s_i(k)|\}}{\sqrt{P_{s_i}}}$$

it is possible to show [7] that $a_r \approx a_i \triangleq a$ for the signals of interest and that the SNR of the complex signal after compression and decompression is bounded by

$$SNR_q < \frac{SNR}{1 + a^2 SNR\left(\frac{2^{-2B}}{3}\right)} \tag{2}$$

From this we see that the degradation due to quantization is more pronounced when SNR is large or when α is large (or both); of course, when SNR is large a larger SNR loss might be more easily tolerated by the system. Note that larger values of α reduce the bound on the post-quantization SNR; large α corresponds to signals that are "peaky"– such as speech, while small α corresponds to signals that are not – such as FM signals.

For multiply-scaling this upper bound can be very nearly achieved. However, when using shift-scaling there is no guarantee that this bound will be even close to being achieved. Shift-scaling can only assure that the scaled version of the largest signal sample lies in $[1/2, 1)$, which leads to the lower bound given in

$$\frac{SNR}{1 + 4a^2 SNR\left(\frac{2^{-2B}}{3}\right)} \le SNR_q < \frac{SNR}{1 + a^2 SNR\left(\frac{2^{-2B}}{3}\right)}$$

These bounds can be used to establish bounds on the cross-correlator output SNR with quantization that, together with simulations, show that it is possible to reduce the quantization to $B = 4$ bits per real part sample and $B = 4$ bits per imaginary part sample and suffer negligible reduction in output SNR (see Figure 1) and in TDOA/FDOA accuracy (see Figure 3 discussed later). Fewer than $B = 4$ bits has been found to be unsatisfactory due to the excessive nonlinearity of the quantization. Figure 2 shows the theoretical and simulated results for both the quantized and the nonquantized cases for a simulated single-sideband signal with $B = 4$. For the case shown it is clear that there is little impact on the output SNR; this is true in general when $SQR \triangleq P_s / s_q^2$ is kept higher than either *SNR* or *DNR*. This is easily seen by recognizing that

$$\frac{1}{SNR_q} = \frac{1}{SNR} + \frac{1}{SQR}$$

so that the output SNR is

$$SNR_{cc} = \frac{BT}{\frac{1}{SNR} + \frac{1}{SQR} + \frac{1}{DNR} + \frac{1}{SNR\ DNR} + \frac{1}{SQR\ DNR}}.$$

Thus, SNR_{cc} is dominated by the lowest of *SNR*, *DNR*, and SQR. Thus, when either (or both) of the signals has low SNR, the additional impact of moderate quantization is negligible. If both signals have high SNR then even moderate quantization has a significant impact, but that is a case where perhaps the degradation can be tolerated.

3. WAVELET TRANSFORM METHOD

The wavelet transform compression algorithm [8] consists of breaking the signal into blocks of $N = 2^p$ samples, applying an *L*-level wavelet transform to each block for $L<p$ (i.e., stopping the cascade of wavelet transform filter bank stages at the level where the filter outputs have $N_B = N/2^L$ elements), grouping the resulting *N* wavelet coefficients into $K = 2^L$ subblocks of $N_B = 2^{p-L}$ samples each, and adaptively quantizing each of these subblocks. For the complex baseband signals used here, this procedure is applied independently to the real and the imaginary components.

The subblocks of the wavelet coefficients are formed within wavelet scale levels as follows: the *N*/2 wavelet transform coefficients from the first filter bank stage are grouped into 2^{L-1} subblocks of 2^{p-L} coefficients each, the *N*/4 wavelet transform coefficients from the second filter bank stage are grouped into 2^{L-2} subblocks of 2^{p-L} coefficients each, . . ., and finally the 2^{p-L} wavelet transform coefficients from the last filter bank stage form a single subblock, and the 2^{p-L} scaling coefficients from the last stage also form a single subblock.

Each one of these subblocks is quantized with a quantizer designed to achieve the desired level of quantization noise. The choice of these quantizers is made easy by the fact that the wavelet transform preserves energy; this property can be used to show that the proper choice of the quantizer cell width is given by

$$\Delta = \sqrt{\frac{12\ P_x}{SQR}}\ ,$$

where *SQR* is the desired signal-to-quantization noise ratio and P_x is the power of the input signal $x(n)$ (in this case, either $\hat{s}_r(k)$ or $\hat{s}_i(k)$). Thus, to obtain a desired SQR, the quantizers $\{Q_1, Q_2, \ldots Q_K\}$ should each have a quantization step size given by Δ. Then the number of bits B_k used by the k^{th} quantizer is chosen to assure that the resulting quantizer covers the range of the k^{th} subblock. This leads to the rule

$$B_k = \lceil_0 \left(\log_2\left[\max\{|W_x^k|\}\right] - \log_2 \Delta + 1\right),$$

where the maximum is taken over the wavelet coefficients in the k^{th} block and the symbol $\lceil_0\ a$ means "the smallest integer not less than 0 that is larger than a;" this means that when the expression in parentheses in the equation for B_k is negative we set $B_k = 0$.

In addition to sending the quantized wavelet coefficients, this scheme requires sending side information to the receiver about the number of bits used for each quantizer as well as the step size used. If the maximum number of bits used by any of the subblocks is $B_{\max}$, then the allowable quantizers are those that use between 0 and $B_{\max}$ bits, for a total of $B_{\max} + 1$ different quantizers; the number of bits required to specify which of these is used for a specific subblock is $\log_2(B_{\max} + 1)$ bits. Since this must be done for each of the *K* subblocks, we require $K \log_2(B_{\max} + 1)$ bits of side information; side information on the quantizer step size also must be sent, which will be no more than the number of bits to which the original signal is quantized (we have assumed 8 bits here). So the total amount of side information is

$$R_{\text{side}} = K \times \log_2(B_{\max} + 1) + 8 \text{ (bits)} .$$

Simulations have shown that it is possible to limit $B_{\max}$ to 7 bits.

In this approach, the wavelet transform is used together with bit allocation to provide a means of reducing the number of bits per (real or imaginary) sample with negligible degradation of the TDOA/FDOA accuracy. This scheme accepts a specific desired signal-to-quantization ratio (SQR) and attempts to minimize the number of bits needed to achieve that SQR value. In practice, the desired SQR can be set either (i) to be roughly equal to the estimated SNR of the signal to ensure that the impact of the compression on the TDOA/FDOA accuracy is negligible, or (ii) to some fixed *a priori* value.

An algorithm parameter that can be adjusted is called B_{min}; it is possible to set all values of B_k, as determined above, that are below some specified value B_{min} to zero. This helps to eliminate wavelet coefficients that contain only noise, and thus helps to reduce the amount of information that must be transmitted. Increasing B_{min} causes a larger number of coefficients to be set to zero and can therefore increase the compression ratio with only a small impact on accuracy.

The R/I quantization and wavelet compression methods described above have focused on minimizing the mean-square error (MSE) due to compression. However, because the goal is to estimate TDOA/FDOA, the minimum MSE criterion is not the most appropriate one because it fails to fully exploit how the signal's structure impacts the parameter estimates. Because TDOA/FDOA accuracy depends not only on SNR but also on the signal's rms bandwidth and rms duration (see (1)), compression approaches that can reduce the amount of data while negligibly impacting the signal's rms widths are desired. Accordingly, one intent of increasing B_{min} in the wavelet method is to remove small wavelet coefficients that may contribute insignificantly to the signal's rms widths. The wavelet transform approach is a natural tool to enable removing time-frequency components of the signal that contribute very little to the signal's rms widths. These ideas are natural generalizations of results we have obtained to determine the correct balance between quantization and decimation for data compression for TDOA systems, which we present here to illustrate the potential for a more general wavelet-based approach.

We now investigate this trade-off between decimation and quantization. Assume that we quantize the complex signal samples using $2B$ bits (B for the real part, B for the imaginary part) and that we sample it at F_s complex samples/second. If the signals are collected for T seconds, then the total number of bits collected is $2BTF_s$. System requirements often specify a fixed length of time, T_l, for the data transmission over the link at a rate R_l bits/second. Then the total number of bits collected must be able to be sent over the link at rate R_l in no more than link time T_l; thus, the constraint to satisfy is $2BTF_s \le R_l T_l$. Equivalently, if we define $R = R_l T_l / T$ as an effective rate and assume equality in the constraint we get

$$R = 2BF_s . \tag{3}$$

Now consider that the received signals are filtered and decimated to a bandwidth of W_f (now $F_s = W_f$ for critical sampling of the complex signals after decimation) and assume that the signals= spectra are flat so that the two SNRs don't depend on W_f. After filtering and decimation, the signal to be transmitted is quantized using $2B$ bits per complex sample (B bits for the real part and B bits for the imaginary part). The result is that the decimated and quantized signal has, from (2), approximate SNR given by

$$SNR_q(B) = \frac{SNR}{1 + a^2 SNR\left(\frac{2^{-2B}}{3}\right)} \tag{4}$$

and the output SNR then depends on the filtered bandwidth and the quantization level according to

$$SNR_o\left(W_f, B\right) = \frac{W_f T}{\frac{1}{SNR_q(B)} + \frac{1}{DNR} + \frac{1}{SNR_q(B)\ DNR}}$$

$$= W_f\, T\, SNR_{eff}(B) , \tag{5}$$

where DNR is the SNR of $\hat{d}(k)$, the signal that is not quantized. Using Equation (5) in Equation (1) gives a bound on TDOA accuracy that depends on the bandwidth after decimation (W_f Hz) and the quantization level (B bits), that is

$$s_t \ge \frac{1}{1.8\, W_f^{3/2} \sqrt{T\ \ SNR_{eff}(B)}} .$$

Using $W_f = R/2B$ from the Arate-bandwidth-bits@ constraint in (3) with $F_s = W_f$ gives

$$s_t(B) \ge \frac{2^{3/2}}{1.8\ R^{3/2}\sqrt{T}} \left[b^{3/2}\sqrt{SNR_{eff}(B)}\right], \tag{6}$$

where it is really the bracketed term that is of interest here, since it shows the tradeoff between decimation and quantization. It is important to remember that (6) includes the rate-bandwidth-bits constraint, so for a fixed R, increasing B necessarily decreases W_f, and vice versa. The nonbracketed term in (6) just scales the result up or down depending on the values of the system parameters effective rate R and collection time T. However, one important insight does come from the first term: the bound on s_t varies as the -3/2 power of the rate R; thus, if you double the allowable rate you get almost three times better accuracy, and if you quadruple the allowable rate you get eight times better accuracy. The reason that increasing the data rate improves the accuracy is because we have constrained the time available to transmit the data, so increasing the data rate allows an increase in the amount of information about the signal that can be transmitted. This is an important insight into the system design issues.

To compute the bracketed term in (6) we first compute $SNR_q(B)$ using (4). Then it is used in (6) to compute the bracketed term for a particular set of values for the parameters á, *B*, *SNR*, and *DNR*. Plots of the bracketed term in (6) versus *B*, parameterized by á, *SNR*, and *DNR* reveal the proper way to choose the optimal value of *B*; that is, how to tradeoff decimation and quantization. Since the value of *R* does not affect these curves, the optimal level of quantization is *not* set by the allowable data rate. Instead, the optimal degree of quantization is set by the interplay between the SNRs of the two signals and the peak factor of the signal to be quantized. Once this level of quantization is determined, the appropriate amount of decimation is defined by determining the allowable amount of bandwidth by solving (3) for the sampling rate F_s given the allowable data rate and the number of bits determined above. Figure 2 shows a plot of the bracketed term in (6) for the case of effective rate R=10 kbps, signal BW=100 MHz, original *B*=8 bits, α=5, *SNR*=20, *DNR*=20. The desired operating point is where the curve is at a minimum; however, note that for this case the minimum requires use of a very small number of bits, where the theory of quantization used here breaks down. Thus, in this case we would choose *B*=4 bits as the best operating point and the signal would be decimated to 1250 Hz. For that choice the resulting s_t is 1.7 times lower than that for decimation only; note that quantization alone can not meet the link requirements for this case. If instead the signals had $a = \sqrt{2}$ (e.g., and FM signal), *SNR*=20, and *DNR*=20 the resulting s_t would be nearly 3 times lower for combined quantization and decimation compared to decimation alone.

This investigation shows that compression algorithms that balance quantization noise and the reduction of rms widths can be more effective than just quantization. In other words, to understand the performance of the wavelet technique we must really consider how the quantization's zeroing of the coefficients impacts the signals rms widths, and to improve its performance we should seek to zero-out wavelet coefficients that contribute insignificantly to the signal's rms widths.

4. SIMULATION RESULTS

Simulations are used to demonstrate the performance of the wavelet transform using adaptive quantization method. These simulations also made use of the compression-correction method proposed in [9], in which prior to sending the compressed signal it is cross-correlated with its original version and the location of the peak of this correlation surface is then sent to the other platform where it is subtracted from the peak locations of the surface computed there. Such an approach is very effective at removing bias imparted by the compression method.

The results presented here are for the case of a radar pulse train whose samples between pulses have been removed by a pre-compression detection procedure; timing pointers are also sent to allow reassembling the pulses into their original timing relationships. The pulse trains are complex baseband linear FM signals having a pulse width of 4 µs and a frequency deviation of ±0.7 MHz, and consisted of 4096 samples generated at 4 MSPS using 8 bits/sample for the real samples and 8 bits/sample for the imaginary samples. The signal that was not compressed had an SNR of *DNR* = 40 dB; the signal that was compressed had SNRs prior to compression in the range $SNR \in [10,40]$ dB.

The R/I quantization scheme used multiply scaling on blocks of 128 samples and quantized the samples to 4 bits/sample for the real samples and 4 bits/sample for the imaginary samples.

The wavelet transform method used a transform having size of N = 2048 and L = 8 levels. Thus, the number of sub-blocks per transform was 256, each having 8 samples per sub-block. The values SQR = 10 dB and B_{min} = 2 were used.

Figure 3 shows three plots. Each plot shows three curves: no compression, wavelet transform (WT) compression, and R/I quantization. The first two plots show the achieved TDOA and FDOA accuracies, respectively, as a function of the compressed signal's *SNR* for the R/I and wavelet transform (WT) methods. The third is a plot of the achieved compression ratios vs. the compressed signal's *SNR*. The impact of R/I quantization on the TDOA/FDOA accuracies can be negligible when a 2:1 compression ratio (4 bit quantization) is used. The wavelet method, however, can achieve a much larger compression ratio but at the expense of about 25% larger TDOA error on but virtually no degradation in the FDOA error.

5. REFERENCES

[1] P. C. Chestnut, "Emitter location accuracy using TDOA and differential doppler ," *IEEE Trans. Aero. and Electronic Systems*, vol. AES-18, pp. 214-218, March 1982.

[2] S. Stein, "Differential delay/doppler ML estimation with unknown signals," *IEEE Trans. Sig. Proc.*, vol. 41, pp. 2717 - 2719, August 1993.

[3] S. Stein, "Algorithms for ambiguity function processing," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. ASSP-29, pp. 588 - 599, June 1981.

[4] D. J. Matthiesen and G. D. Miller, "Data transfer minimization for coherent passive location systems," Report No. ESD-TR-81-129, Air Force Project No. 4110, June 1981.

[5] U. Benz, K. Strodl, and A. Moreira, "A comparison of several algorithms for SAR raw data compression," *IEEE Trans. Geosci. Remote Sensing*, vol. 33, pp. 1266 - 1276, Sept. 1995.

[6] Y. Neuvo, "Digital filter implementation considerations," in S. Mitra and J. Kaiser (Eds.), *Handbook for Digital Signal Processing*, New York: John Wiley & Sons, 1993.

[7] M. L. Fowler, "Coarse quantization for data compression in coherent location systems," under revision for *IEEE Transactions on Aerospace and Electronic Systems*.

[8] M. L. Fowler, "Data compression for TDOA/DD-based location system," US Patent #5,991,454 issued Nov. 23, 1999, Lockheed Martin Federal Systems.

[9] G. Desjardins, "TDOA/FDOA technique for locating a transmitter," US Patent #5,570,099 issued Oct. 29, 1996, Lockheed Martin Federal Systems.

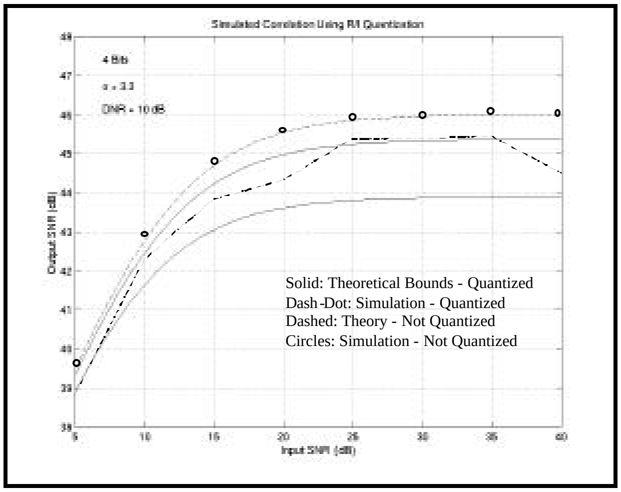

Figure 1: Bounds and Simulation Results for Output SNR using R/I Approach

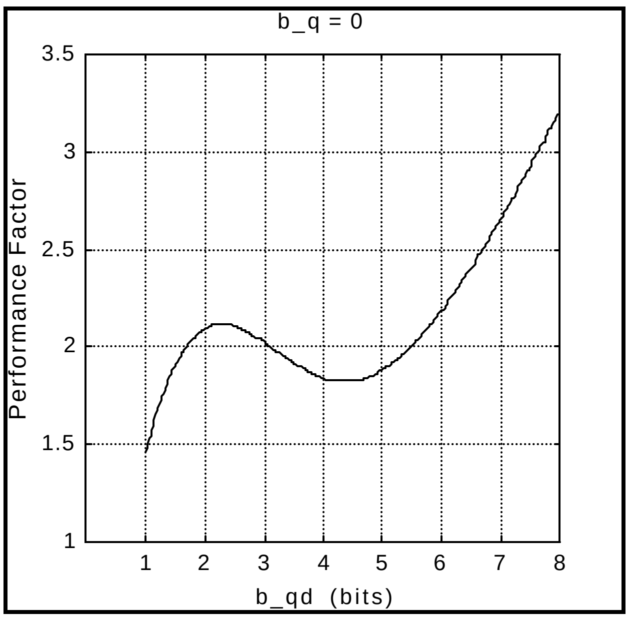

Figure 2: "Performance Factor" for Decimation vs. Quantization Tradeoff

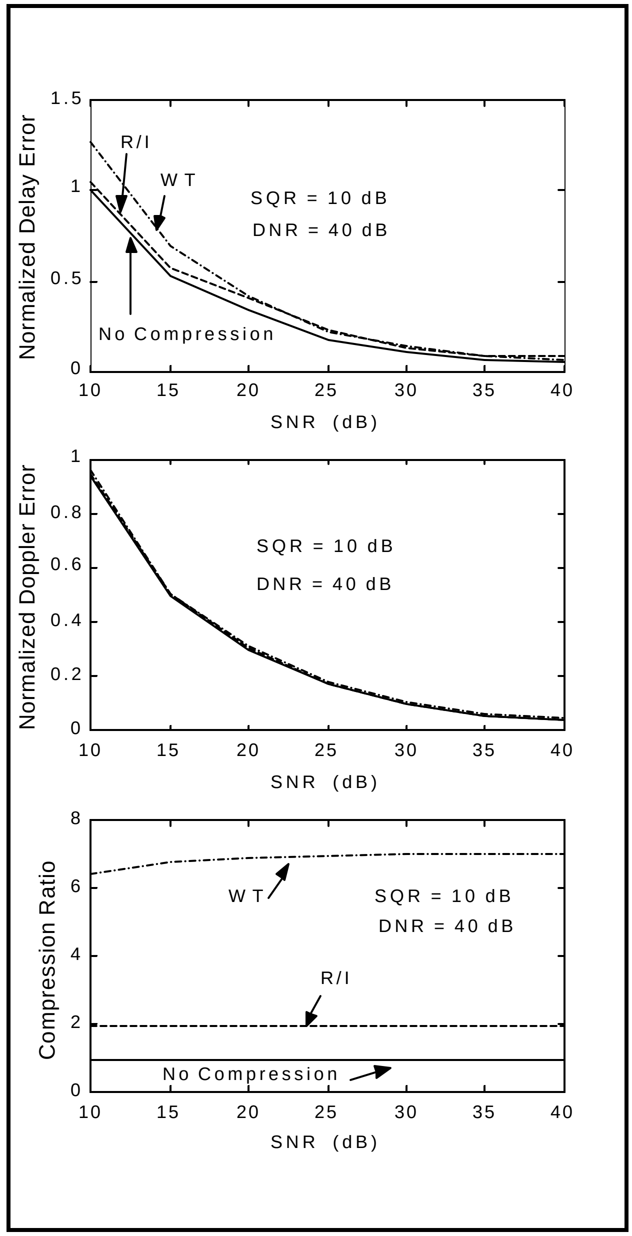

Figure 3:Simualtions for TDOA/FDOA Accuracy using R/I Quantization and Wavelet Method