Naval Submarine Medical Research Laboratory

NSMRL REPORT 1199 **11 APRIL 1996**

Effect of digital recording parameters on discrimination features of acoustic signals in noise

T. S. Santoro
J. S. Russotti
and
T. E. Hanna

Released by
R. G. Walter, CAPT, DC, USN
Commanding Officer
Naval Submarine Medical Research Laboratory

Effect of digital recording parameters on discrimination features of acoustic signals in noise.

T. S. Santoro
J. S. Russotti,
and
T. E. Hanna

Naval Submarine Medical Research Laboratory
Report 1199

Naval Medical Research and Development Command
Research Work Unit 65856N M0100.001-5051

Approved and Released by

R. G. Walter

R. G. Walter, CAPT, DC, USN
Commanding Officer
NAVSUBMEDRSCHLAB

Summary Page

Problem

Design parameters of sonar digital signal processing systems are selected for visual presentation of data and sacrifice the quality of the aural presentation to the operator. Two critical parameter choices, sample rate and quantization code, degrade aural signal discrimination in noise by the human listener and their effect on auditory perception of processed signals is not fully understood.

Findings

Critical listener perceptions for discrimination are (1) signal beat at low frequencies, (2) spectral shape at higher frequencies, and (3) individual signal temporal modulation. The importance of each perception is strongly dependent on the sample rate used in signal processing but not the quantization code.

Applications

Design of sonar signal processing equipment for optimal human auditory discrimination performance.

Administrative Information

This work was completed under Naval Medical Research Development Command Research Work Unit 65856N M0100.001-5051, Digital processing for auditory sonar. The views expressed in this report are those of the authors and do not reflect the official policy or position of the Department of the Navy, Department of Defense, or the U.S. Government. This report was approved for publication on 11 Apr 96 and designated NSMRL Report 1199.

Abstract

An experiment was performed to determine the effects of digital signal processing sample rate and quantization code on auditory perception of sonar signals. Fifteen sonar signals were sampled and played back under nine conditions of sample rate and quantization code. In each condition all pairwise combinations of these signals in noise were presented to 35 subjects in an ABX discrimination task. The resulting matrices of discrimination errors were analyzed by multidimensional scaling. The first two scaling dimensions recovered in order of statistical significance were associated with perceptions related to (1) signal beat at low frequencies and (2) signal spectral shape in the higher frequencies. Further recovered dimensions were related to particular temporal modulation of individual signals. The importance of the first three discrimination features depended on the three sample rate conditions. Each halving of the sample rate removed one of the features from any significant contribution to the discrimination task. The quantization conditions had little influence on the significance of the discrimination features except for the mid-range sample rate (6.25 kHz).

[Blank Page]

Effect of digital recording parameters on discrimination features of acoustic signals in noise.

Digital audio recording and playback equipment has introduced two new parameters into the transmission path between source and listener: sample rate and quantization code. Values for these design parameters have a major effect on both system complexity and listening quality (Blesser, 1978; Fielder, 1987). Faster sample rate provides wider recoverable bandwidth from the original signal. Hence, more of the high frequency spectrum of signal and noise sources will be preserved and presented to the listener. The number of bits used in the computer to represent each data sample (the quantization code) affects the amount of uncertainty, and thus the noise, in the representation of sample values (Hayashi & Kitawaki, 1992). We should expect that when digital audio is used by sonar operators to detect and classify complex signals in a noise background, these parameters will also be critical.

The detection of acoustic signals in background noise by a trained listener is based on the perception of both the average acoustic power difference between signal plus noise (S + N) versus noise alone (level difference) and other primitive features distinctly associated with a given S + N combination. The latter are features that arise in specific, narrow frequency bands or from various amplitude or frequency modulations over the entire signal frequency band. They account for detection performance at very low overall S/N ratios such as -10 to -15 dB and often depend on the listener's perception of what is signal and what is noise in a given stimulus. When signal S1 plus noise N and signal S2 plus noise N are carefully balanced with respect to listener detection threshold for each signal, the listener can not distinguish any level difference between the S1 + N epoch and the S2 + N epoch. Therefore, the discrimination between S1 and S2 epochs must be based on the other perceptual features alone. In this situation, a same/different discrimination task between pairs of signals using brief listening periods can define these perceptual features independent of level or cognitive effects that might arise in more complicated test procedures.

We obtained this basic characterization of signal features for a group of typical sonar signals from ship traffic under different sample rate and quantization conditions. We employed multidimensional scaling of pairwise auditory discriminations (Gray, 1977; Howard, 1977; Mackie, Wylie, Ridihalgh, Shultz, & Sletzer, 1981) to uncover the perceptual features used by the listeners. Our test stimuli were from the same categories of sonar signals as those of Mackie et.al (1981) and Howard (1977) and our scaling analysis uncovered the same types of perceptual features used by subjects in performing those discrimination tasks. In addition, this study shows how those features depend on sample rate and quantization parameters.

Method

Signals.

We selected a group of 15 signals representative of a variety of sonar sources recorded on analog tapes at very high signal strength with essentially no background noise present. The different power spectra all had one main peak at the upper frequency end (3 to 8 kHz) with varying sharpness. Starting at different places in the midband (0.25 to 2.5 kHz), some exhibited a raised flat shoulder leading to the peak. We associated these spectral characteristics with a hissing sound from the high frequency peak and a dragging sound from the mid-band shoulder. In addition, the signals had varying amounts of temporal modulation giving them certain characteristic sounds such as a laboring or galloping beat, a machinery-like hum or rumble, and gurgling or washing sounds that were quite pronounced in some cases and made them very easy to distinguish.

A spectral model was made of a sea-state 2 recording from a typical sonar system for use as background noise. The noise power spectrum was shaped from the output of a white noise generator using a series of one-third octave band filters. The resultant spectrum had a single broad peak at 8 kHz and dropped off smoothly at lower frequencies at about 5 dB per octave with no mid-band shoulder. The noise produced an unmodulated, high-pitched hissing sound.

Procedure

Each signal and the noise were digitized separately at three rates: 12.5, 6.25, and 3.125 kHz. Anti-aliasing brick-wall filters with upper cut-offs of 5.0 kHz, 2.5 kHz and 1.25 kHz, respectively, were used in both the digital recording and playback procedures. At each sample rate, the data were quantitized in three different codes; 12-, 8-, and 4-bit. Thus we had nine different combinations of sample rate and quantization code under which to measure listener performance on the task of discriminating between pairs of our stimulus group. The discrimination task was that of comparing two signals to a third standard and simply telling which of the two was the same as the standard. This is known as an ABX comparison.

An ABX trial sequence proceeded as follows. The subject wore a headset with voice only to the left ear and test stimuli to the right ear. With silence to the right ear, the voice would state the trial number during a 5 second period. The ABX sequence would then commence to the right ear. Three seconds were allotted for each of the three signals in continuous noise. Subjects knew the first signal was always the standard and one of the next two signals would be the same as the first although not the identical recording. There was no quiet time between signals so that signal differences were the only clues to determine when each of the ABX segments occurred. After the 9 second ABX exposure, both ears were left in silence for a 10 second period while the subject checked off his response on an answer sheet. Thus a complete trial took 24 seconds.

We used all four possible orderings of two stimuli in the ABX paradigm: AAB, ABA, BAB, and BBA. Thus the original 105 possible AB pairs from our 15 test signals were counterbalanced into randomized sets of 420 trials for each of the nine conditions of sample rate and quantization. We divided the sets into six groups of 70 trials each as this number required about 1/2 hour for a subject to complete. Subjects thus would require three hours to do a complete 420 trial test condition. With breaks each half hour to relieve fatigue, this would constitute a full morning or afternoon session for subjects.

At each test condition, we set S/N ratio for each signal to be 7 dB above signal detection threshold averaged over subjects from data in a previous study (Russotti & Santoro, 1992). By carefully adjusting individual signal S/N ratio relative to its threshold for each ABX trial, we did, to the extent possible, remove average signal level as a discrimination clue. Each 3 second period flowed smoothly into the next with little or no perceptible change of overall level. Hence, the discriminations for the most part were free of level threshold clues. Because we used averaged thresholds, there was of course the possibility that an individual subject with above-average sensitivity on certain signals could still detect a level difference. Subjects were queried during breaks on their general perceptions of the test signals and did report that in some trials all three ABX periods sounded the same. It was as though there were just noise alone in all three. We take these to be trials where the test conditions have obliterated all signal-specific discrimination clues and we have done a good job of balancing out S/N thresholds.

Subjects

Thirty-six naive listeners with normal hearing, none of whom were in the 1992 study, served as subjects for this study. We divided subjects into three, 12 member, groups according to the three sample rate conditions. Each group was tested on the three quantization conditions at a single given sample rate. The ordering of quantization condition for each test session was randomized so that on each day the subject would be tested at all three quantization conditions over the six sessions without repeating the same condition over any two consecutive sessions. As a control and for training purposes, all groups were presented 420 trials of 16-bit, 50-kHz sample rate signals over six sessions on the day before the start of separate group testing.

Results

Subjects entered their judgements on score sheets and the answers were converted to lower-half diagonal-absent matrices of error rates by dividing the total errors made on each stimulus pair by the number of subjects in the test group times the number of trials presented to each subject. One subject's data was discarded because of anomalies. There were thus 12, 11, and 12 subjects in the high, medium, and low sample rate groups designated Groups I, II, and III, respectively. The error rate matrices for each signal pair at the 9 test conditions are shown in Table 1a, 1b, and 1c. If subjects did pure guessing on this two-alternative forced-choice task, we could expect error rates of 50 percent. For a few test pairs, as seen in the matrices, rates did reach the chance level indicating that the two sounds in question were indeed indistinguishable for the given test conditions. Likewise, for certain conditions, there were a few pairs that all subjects could distinguish on every trial.

Average error rates over all subjects and stimulus pairs ranged from a high of 22.8% to a low of 7.48% as given in Table 2 for the nine test conditions. A mixed design 2-way analysis of variance on the data showed significant effects due to sample rate, $F(2,33)= 26.86, p < .000001$, and quantization, $F(2,66)= 59.74, p < .000001$, with an interaction statistic of $F(4,66)=4.65, p < .01$. It is clear from the table that the discrimination task was always easier for the Group I, or high sample rate, condition. Under that condition, overall error rates were always under 10%. The major overall change comes in the move to sample rate Group II or Group III from Group I. For both these groups, error rates are more than double those of Group I. Likewise, for quantization effects within each sample rate group, the major change comes from dropping to 4-bit code from 8- or 12-bit. The overall effect of going from Group II to Group III (6.25 kHz vs 3.125 kHz) or from 12-bit quantization to 8-bit is quite small.

These error rates are shown connected with solid lines in Fig. 1 superimposed on dotted lines connecting the detection threshold averages from our previous study. Both thresholds and error rates are lowest in the Group I condition. There is a major increase in discrimination error rates between the Group I to Group II condition and a slight dropback in rates between the Group II and Group III conditions. In contrast, detection thresholds smoothly increase over the three sample rate conditions.

Multi-Dimensional Scaling

In addition to the statistical tests of significance, a Multi-Dimensional Scaling (MDS) analysis was undertaken to interpret the

Table 1a

Error rate matrices group I percent error per 48 subject-trials

GROUP 1 04 BITS PERCENT ERROR RATE PER 48 SUBJECT-TRIALS
12 SUBJECTS 4 TRIALS EACH SIGNAL PAIR

8													
12	2												
4	19	12											
15	0	12	4										
21	15	17	21	4									
12	2	17	4	52	4								
23	4	21	2	29	6	12							
12	8	6	6	0	12	6	0						
15	17	10	44	8	52	4	6	4					
4	44	6	4	6	8	8	6	2	10				
10	12	25	29	4	27	8	4	0	46	15			
10	2	4	4	15	4	12	25	6	2	2	4		
2	6	2	0	2	2	0	4	2	0	2	2	4	
2	2	2	2	4	6	0	4	0	2	0	0	4	2

GROUP 1 08 BITS

4													
17	4												
6	8	15											
8	4	4	4										
6	19	15	27	8									
15	2	4	0	33	2								
12	4	21	10	23	2	10							
10	0	4	2	4	10	0	6						
6	12	4	25	2	50	4	6	10					
8	23	2	6	2	17	2	6	8	8				
12	10	15	25	0	25	2	2	0	46	8			
6	0	4	4	15	8	6	12	4	8	2	0		
6	0	0	0	0	2	0	4	4	0	2	2	4	
0	2	0	4	0	6	2	0	2	4	0	0	0	0

GROUP 1 12 BITS

6													
19	10												
4	21	4											
4	2	8	6										
10	8	10	31	12									
10	2	6	0	44	0								
8	2	33	0	17	6	23							
10	0	2	2	0	0	4	2						
12	15	8	40	0	33	0	6	6					
8	27	6	8	0	2	0	4	0	15				
12	10	10	15	0	19	2	0	10	50	4			
6	0	6	0	17	4	8	4	4	27	2	6		
2	0	0	0	2	6	0	2	0	0	2	4	0	
2	0	0	0	0	4	2	8	0	0	4	4	0	0

Table 1b

Error rate matrices group II percent error per 44 subject-trials

GROUP 2 04 BITS PERCENT ERROR RATE PER 44 SUBJECT-TRIALS
11 SUBJECTS 4 TRIALS PER SIGNAL PAIR

32													
34	45												
25	25	41											
30	16	23	34										
39	57	43	36	32									
32	34	20	27	48	25								
18	23	36	27	27	27	36							
16	25	9	14	0	20	7	9						
41	43	41	39	18	43	16	30	27					
30	30	52	30	18	43	34	25	16	32				
27	43	48	45	25	57	23	36	11	43	45			
11	20	14	11	11	11	14	23	9	23	20	36		
0	0	2	7	2	11	2	2	2	9	2	2	2	
11	7	5	9	11	11	5	23	9	20	5	16	5	9

GROUP 2 08 BITS

23													
18	11												
9	18	16											
7	27	30	23										
27	50	45	18	14									
27	18	7	16	48	27								
7	16	16	14	27	25	25							
9	11	14	2	7	18	0	7						
11	27	39	11	14	43	18	11	18					
11	43	36	18	14	34	20	18	2	36				
27	55	48	23	25	45	25	36	18	43	23			
9	14	11	16	20	14	25	9	5	11	25	11		
2	9	2	0	7	5	5	9	2	2	5	2	5	
9	18	11	9	2	9	9	16	16	16	11	2	5	5

GROUP 2 12 BITS

11													
11	39												
14	25	11											
11	14	32	20										
30	45	27	16	14									
11	25	11	9	48	18								
16	14	14	16	25	32	27							
7	20	5	14	7	34	5	11						
9	41	32	11	34	36	27	27	25					
16	36	25	18	27	34	30	27	5	30				
18	43	41	7	16	30	27	16	7	39	41			
2	11	9	11	32	20	14	9	5	11	14	11		
5	9	7	0	5	11	7	2	0	11	9	5	2	
14	16	9	2	2	14	5	18	14	11	18	7	9	0

Table 1c
Error rate matrices group III percent error per 48 subject-trials

GROUP 3 04 BITS PERCENT ERROR RATE PER 48 SUBJECT-TRIALS
12 SUBJECTS 4 TRIALS PER SIGNAL PAIR

12													
12	58												
4	27	23											
0	15	6	21										
15	58	44	12	10									
4	8	6	4	44	6								
8	29	10	10	10	12	19							
17	35	21	31	6	48	12	23						
8	46	46	10	2	52	12	17	31					
15	50	46	17	6	46	8	10	48	40				
12	44	58	10	10	54	17	10	29	54	52			
4	6	15	25	25	19	15	15	19	15	12	10		
2	6	4	2	0	8	0	10	6	0	6	10	8	
2	25	27	29	38	17	29	38	23	21	15	12	17	31

GROUP 3 08 BITS

4													
6	40												
4	12	6											
0	12	4	12										
2	35	29	12	4									
2	21	8	4	33	23								
6	10	23	8	10	6	35							
15	31	27	12	8	54	12	8						
4	48	44	17	6	46	12	12	21					
8	44	23	12	8	48	6	31	46	40				
12	58	40	17	21	29	15	8	23	48	29			
0	15	4	12	8	12	12	12	19	12	12	17		
2	6	6	2	2	2	15	17	2	2	4	8	4	
2	19	21	12	23	17	29	35	15	25	10	23	12	10

GROUP 3 12 BITS

6													
8	46												
2	12	8											
0	15	8	6										
12	40	40	4	8									
0	17	4	6	27	17								
0	4	17	10	27	17	19							
10	19	21	21	4	48	8	15						
2	50	52	8	12	29	15	12	19					
15	42	25	10	12	44	10	8	40	23				
2	60	42	8	6	33	12	12	21	50	42			
2	19	8	10	4	19	10	15	25	17	19	23		
0	8	8	6	4	8	0	12	4	4	2	2	6	
0	31	25	15	50	31	38	29	17	31	19	10	8	4

results in terms of auditory perception. The computer program INDSCAL (Carroll and Chang, 1970; Young, 1970) creates a configuration of points in an N dimensional space whose separations represent measured pairwise confusions of stimuli such as we have in the data of Table 1. Where, for example, a data matrix entry is large, signifying a high confusion between the stimuli represented by the row/column pair, the associated points would be very close to each other and vice versa. A group of M points may always be represented in M-1 dimensions regardless of the interpoint relationships. When the original point separations are highly correlated, however, only one or two independent dimensions are required to represent the resulting configuration. Dimensions that account for a small percentage of the variance in inter-point separation may be ignored in the final solution resulting in a minimal dimensional solution that facilitates interpretation. A study of the arrangement of points along each remaining dimension coupled with a knowledge of the stimuli represented by each point yields an insight to the perceptual feature represented by that dimension.

A combined scaling analysis may be done when the same stimuli are discriminated under different test conditions such as the nine conditions used in this study. A single generalized configuration of points is produced along with a set of weights for each condition. The weights represent the relative stretching or shrinking in the coordinates of each point along each dimension in going from one condition to the next. A study of the change in weights on each dimension yields an insight on the effects of test conditions on the perceptual feature the dimension represents. The dimensions are given rank labels based on the distribution of variance among them in the generalized configuration calculated from the data for all conditions. However, the relative importance of each dimension in a given

Table 2
Average error rates for nine test conditions

Quantization Code (Bits)	Group I 12.5 kHz	Group II 6.25 kHz	Group III 3.125 kHz
04	9.5 %	22.8 %	20.2 %
08	7.5 %	17.5 %	17.0 %
12	7.4 %	17.3 %	17.1 %

condition is determined by the weights for that condition.

The results of our INDSCAL analysis are generally quite similar to the earlier work of Mackie, et al. (1981) and Howard (1977) on sonar signals from the same stimulus categories. In those two related studies, scaling analysis uncovered 3 to 5 perceptual feature dimensions. One was associated with the high frequency spectral nature of the stimuli while the others involved low frequency temporal amplitude or frequency modulation that produced very strong beat clarity, tonality, and rate effects. Working for the most part with configurations scaled into six or fewer dimensions, we came to the conclusion that there are three independent perceptual features involved in discriminating our 15 signals. These are: beat presence or absence, spectral shape, and temporal modulation. We based this on the very consistent and repeatable nature of the dimension weights for each test condition over all scaling runs.

Figures 2a, 2b, and 2c show the nine sets of weight vs. dimension results organized according to the sample rate variable conditions (Groups I, II, and III) in three graphs. Each graph has three separate curves, one for

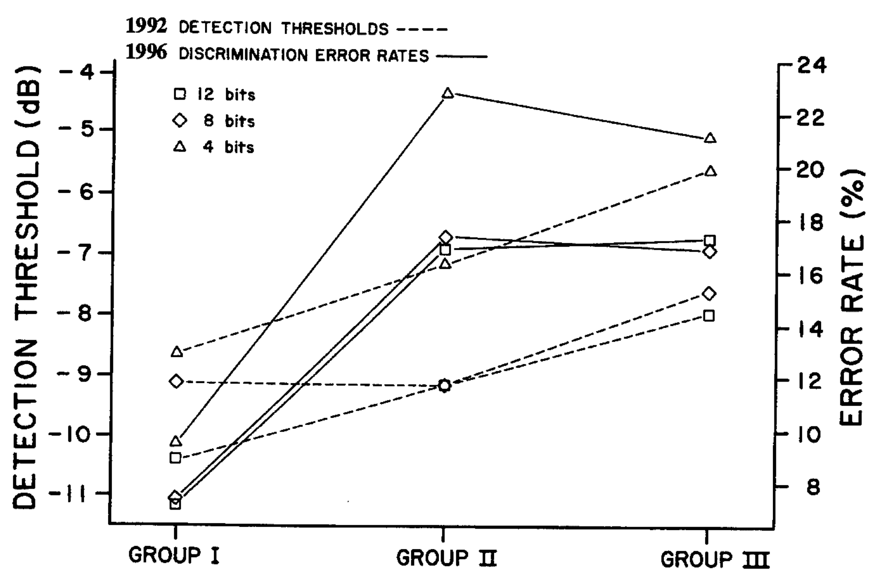

Figure 1. Average detection thresholds (Russotti & Santoro, 1992) and overall error rates for nine test conditions. Left vertical axis S/N in dB at threshold, right axis error rate in % incorrect discrimination.

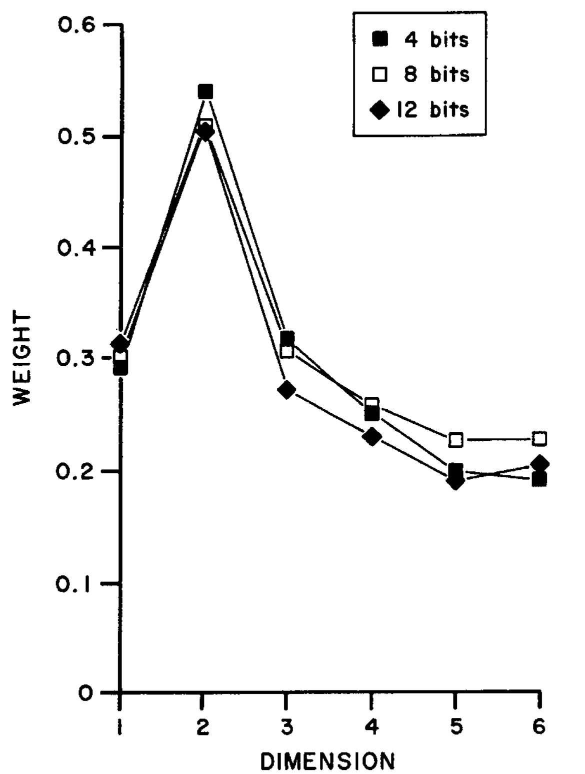

Figure 2a. Six dimensional solution weights vs. dimension - Group I condition

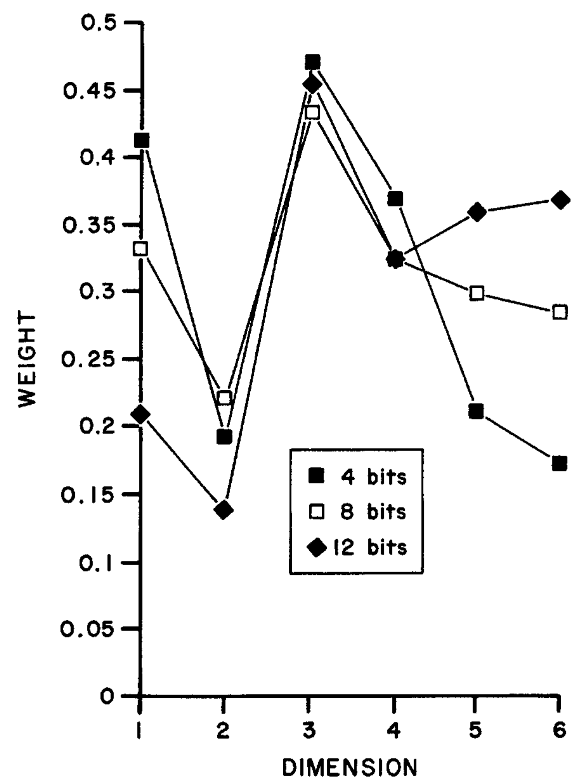

Figure 2b. Six dimensional solution weights vs. dimension - Group II condition

each of the three quantization variable conditions (12-bit, 8-bit, and 4-bit) at the same sample rate. As shown, a different scaling dimension is weighted most heavily in each of the different sample rate conditions and thus becomes the dimension accounting for the most variance in the configuration for that condition. Except for certain dimensions of the Group II sample rate, the weighting pattern is the same for the 3 quantization conditions at each sample rate condition.

Discussion

We have drawn two conclusions from the results shown on the nine plots of Fig. 2. First, we conclude that the perceptual cues are closely linked to the sample rate variable. Each sample rate condition brings with it a distinct set of dimension weights. While three to six scaling dimensions are required in

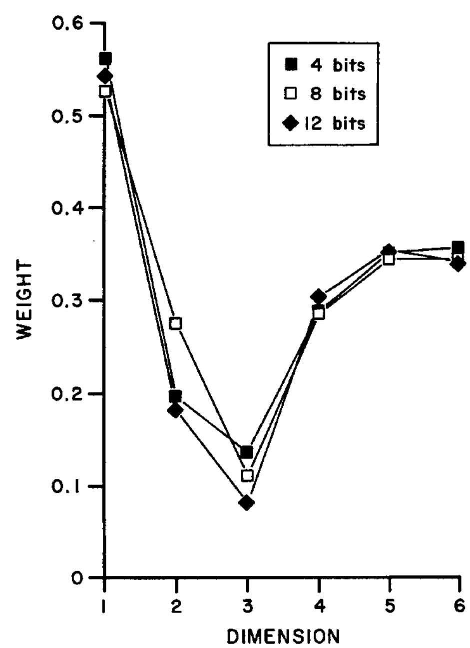

Figure 2c. Six dimensional solution weights vs. dimension - Group III condition

the complete solution for all test conditions, only one or two play the major role at each of the three sample rate conditions. Therefore, we conclude that each change in sample rate brings into play quite different perceptual features for discrimination performance. This is understandable given the range of sample rates used. As we go from the Group I rate of 12.5 kHz to the Group II rate of 6.25 kHz and then to 3.125 kHz for Group III we are removing exactly half of the power spectrum of our signals on each step. The dramatic changes observed in the discrimination data are witness to the considerable amount of information in the portion of the spectrum that was removed.

Our second conclusion is that perceptual clues are not linked to the quantization variable. While error rate is higher for the 4-bit condition as compared to the 8- and 12-bit conditions, the scaling analysis tells us this is not due to the introduction or removal of separate perceptual features. As seen in Fig. 2, the three traces for quantization code weights at each sample rate group are very close to each other. They always reach the same single peak value for the same scaling dimension. Only for Group II do these weight traces separate from each other and there it is only on the lower-weighted dimensions.

By matching up physical stimulus characteristics with corresponding MDS configuration positions, we have associated the scaling dimensions with the following three perceptual features: signal beat presence or absence, signal spectra, and temporal modulation characteristics. The features are in descending order of the amount of variance in the data accounted for by each dimension in the unweighted generalized solution over all experiment conditions. Dimension 1, associated with beat presence or absence, accounted for 17.7% of the variance; dimension 2, signal spectra, 12.8%; and dimensions 3-6, temporal modulation characteristics, together accounted for 40.1% of the total variance in the original data or about 10% for each dimension. However, which dimension accounts for the most variance, and hence the finest discrimination, for a given experimental condition depends on the weights for that condition.

For example, the weight distribution indicates that for the high sample rate condition when a broad spectrum of each signal is available, subjects use pitch variations arising from the high frequency peaks and mid-range shoulder characteristics for the major part of their discrimination decision. Howard (1978) characterized this feature by "tinniness" or the relative amount of high frequency energy. Subjects can align all 15 signals on this single perceptual dimension (dimension 2 in Fig. 2) with very good separation. We found locations on that dimension do in fact correspond to the rank ordering of the high frequency peaks of each signal. Although the error rates of Table 2 show some change with quantization at the high sample rate, the scaling solution for that rate is quite insensitive to the quantization parameter down to even the 4-bit code.

Once the signal spectrum is halved, as in Group II, and halved again as in Group III, the subjects lose this fine pitch separation capability. The scaling dimension associated with this percept is then weighted very low.

When this happens, subjects turn to other criteria that we associate with characteristic sounds due to temporal modulation, i.e.; beating, galloping, humming, gurgling, etc. (dimensions 3-6). These are related to low-frequency modulation of the signals and the dimensions labeled "beat rate" and "beat tonality" by Mackie, et al. (1981). If one of these sounds is distinct enough, it can dominate one of the higher dimensions in the general scaling solution by itself (e.g., dimensions 3-6). Once the individual temporal modulation characteristics of each signal become the perceptual dimensions of importance, sensitivity to quantization code occurs. Hence, the choice of 12-bit or 8-bit coding instead of 4-bit coding at sample rates around 6.25 kHz becomes important to discrimination performance.

The last resort for making a discrimination under the worst sample rate condition is detection of some temporal modulation, usually beats of any kind, that can be distinguished from the noise background. In the Mackie, et al. (1981) study, this dimension was called "beat clarity" and accounted for the largest percentage of variance in the data set as it also does in our overall solutions. Under adverse conditions, signals without temporal characteristics are indistinguishable from noise when the signal level cue is balanced out as in our experiments. Hence many signal pairing trials become simple comparisons of the one signal in noise whose beat can be distinguished against the other signal in noise combination that has the same perceived level but lacks a distinct beat. We contend that this is a somewhat different percept than others related to signal beat rate or beat tonality discrimination.

These discrimination results confirm and extend the general conclusions on the effects of reduced sample rate and quantization drawn from listener detection performance in our earlier study. Both detection and discrimination measures are better at 12.5 kHz (Group I) and 12 bits. When adjusted for the detection threshold differences of Fig. 1, 8-bit coded signals can be discriminated about as well as 12-bit signals at each sample rate. However, note from Fig. 1 that, while the average threshold difference is minor at 3.125 (Group III) and 6.25 kHz (Group II), it grows to about 1.5 decibels for Group I indicating clearer superiority of the 12-bit code at this sample rate. We always measure lower performance for the 4-bit code at all sample rates.

At the middle sample rate (Group II), we see changes in the scaling weights with quantization for the lower-weighted dimensions, even in going from 12- to 8-bit code. Weights, and consequently reliance, decrease for temporal modulation dimensions and increase on the beat presence or absence dimension. In this way the beat presence or absence dimension is different in kind from the other dimensions related to beat rate and tonality. At the lowest sample rate (Group III), this "beat clarity" dimension is most heavily relied upon to give listeners some indication of signal presence in the background noise.

References

Blesser, B. (1978). Digitization of audio: a comprehensive examination of theory, implementation, and current practice. *Journal of the Audio Engineering Society 25*, 739-771.

Carroll, J.D., & Chang, J.J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika, 35*, 238-319.

Fielder, L. D. (1987). Evaluation of the audible distortion and noise produced by digital audio converters. *Journal of the Audio Engineering Society 35*(7/8), 517-534.

Gray, J.M. (1977). Multidimensional perceptual scaling of musical timbres. *Journal of the Acoustic Society of America 61*(5), 1270-1277.

Hayashi, S. & Kitawaki, N. (1992). An objective quality assessment method for bit-reduction coding of wideband speech. *Journal of the Acoustic Society of America 92*(1), 106-113.

Howard, J. H. (1977). Psychophysical structure of eight complex underwater sounds. *Journal of the Acoustic Society of America 62*(1), 149-156.

Howard, J. H., & Ballas, James. A. (1978). *Feature selection in auditory perception* (Technical report ONR-78-5). Washington, DC: The Catholic University of America, Human Performance Laboratory.

Mackie, R.R., Wylie, C.D., Ridihalgh, R.R., Shultz, T.E., & Sletzer, M.L. (1981). *Some dimensions of auditory sonar signal perception and their relationships to target classification* (Technical Report ONR-2723-1). Goleta, CA: Human Factors Research Incorporated.

Russotti, J.S. & Santoro, T.P., (1992). *Effects of digital sampling rate and quantization on passive auditory sonar target detection performance* (Report No. 1184). Naval Submarine Medical Research Laboratory, Groton, CT.

Young, F. W. 1970. Nonmetric multidimensional scaling: recovery of metric information. *Psychometrika, 35*, 455-473.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 074-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF THE REPORT Approved for public release; distribution unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NSMRL Report 1199	5. MONITORING ORGANIZATION REPORT NUMBER(S) NA
6a. NAME OF PERFORMING ORGANIZATION Naval Submarine Medical Research Laboratory 6b. OFFICE SYMBOL (If Applicable)	7a. NAME OF MONITORING ORGANIZATION Naval Medical Research and Development Command
6c. ADDRESS (City, State, Zip Code) Box 900, Naval Submarine Base NLON, Groton, CT 06349-5900	7b. ADDRESS (City, State, Zip Code) 8901 Wisconsin Avenue Bethesda, MD 20889-5606
8a. NAME OF FUNDING/SPONSORING ORGANIZATION Same as 7a 8b. OFFICE SYMBOL (If Applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, Zip Code) Same as 7b	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
65856N	M0100	001	DN240501

11. TITLE (Include Security Classification)
Effect of digital recording parameters on discrimination features of acoustic signals in noise.

12. PERSONAL AUTHOR(S)
Thomas S. Santoro, Joseph S. Russotti, & Thomas E. Hanna

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
	FROM ______ TO ______	1996 Apr 11	10

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Multidimensional scaling; digital quantization and sample rate; auditory perception

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

An experiment was performed to determine the effects of digital signal processing sample rate and quantization code on auditory perception of sonar signals. Fifteen sonar signals were sampled and played back under nine conditions of sample rate and quantization code. In each condition all pairwise combinations of these signals in noise were presented to 35 subjects in an ABX discrimination task. The resulting matrices of discrimination errors were analyzed by multidimensional scaling. The first two scaling dimensions recovered in order of statistical significance were associated with perceptions related to (1) signal beat at low frequencies and (2) signal spectral shape in the higher frequencies. Further recovered dimensions were related to particular temporal modulation of individual signals. The importance of the first three discrimination features depended on the three sample rate conditions. Each halving of the sample rate removed one of the features from any significant contribution to the discrimination task. The quantization conditions had little influence on the significance of the discrimination features except for the mid-range sample rate (6.25 kHz).

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS REPORT ☐ DTIC USERS	

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL
Ellen Perkins, Publications	(203) 449-3263	

DD Form 1473, JUN 86 — Previous editions are obsolete.
S/N 0102-LF-014-6603
SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED