Section 2 Sensory Communication

Chapter 1. Sensory Communication

Academic and Research Staff

Professor Louis D. Braida, Professor Richard M. Held, Nathaniel I. Durlach, Dr. Anuradha M. Annaswamy, Dr. David L. Brock, Dr. Donald K. Eddington, Dr. Susan L. Goldman, Dr. Jeng-Feng Lee, Dr. William M. Rabinowitz, Dr. Christine M. Rankovic, Dr. Charlotte M. Reed, Dr. J. Kenneth Salisbury, Dr. Mandayam A. Srinivasan, Dr. Annie H. Takeuchi, Dr. Rosalie M. Uchanski, Dr. Thomas E. Wiegand, Dr. David Zeltzer, Dr. Patrick M. Zurek, Lorraine A. Delhorne, Seth M. Hall

Visiting Scientists and Research Affiliates

Dr. Kenneth W. Grant, Dr. Kaoru Sekiyama, Gerald L. Beauregard, Geoffrey L. Plant

Graduate Students

Walter A. Aviles, Jyh-Shing Chen, Kiran B. Dandekar, Joseph G. Desloge, Paul Duchnowski, Eric M. Foxlin, Joseph A. Frisbie, Svetlana Gaffron, Julie E. Greenberg, Rogeve J. Gulati, Rakesh Gupta, Louise Jandura, Steingrimur P. Karason, Gregory R. Martin, Thomas H. Massie, Joseph A. Maxwell, Michael P. O'Connell, Matthew H. Power, Barbara G. Shinn-Cunningham, Nick Swarup, Hong Z. Tan, Daniel P. Welker, Craig B. Zilles

Undergraduate Students

Walter E. Babiec, Julien B. Beasley, Bridget L. Coffman, Gail Denesvich, Ashanthi Gajaweera, Andrew L. Howe, Megan C. Jasek, Michael T. Keagy, Kevin C. Knoedler, Irene Kotok, Danielle G. LeMay, David S. Lum, Jason Mueller, Philip P. Nadeau, Erik Nygren, Charlie Pan, Nicholas Pioch, Diane E. Ronan, Andrew Ugarov, Evan F. Wies, Jennifer A. Wozniak

Technical and Support Staff

Ann K. Dix, Eleanora M. Luongo, Michael T. Tuyo

1.1 Introduction

The Sensory Communication Group is conducting research on (1) the auditory and tactual senses, (2) auditory, visual, and tactual aids for individuals who are hearing-impaired or deaf, and (3) human-machine interfaces for teleoperator and virtual-environment systems (involving the visual as well as the auditory and tactual senses). Within the domain of hearing aids, research is being conducted on systems that bypass the outer and middle ear and directly stimulate the auditory nerve electrically (cochlear prostheses), as well as on systems that stimulate the ears acoustically. The research on taction is focused not only on speech reception for the totally deaf, but also on the ability of the human hand to sense and manipulate the environment. Within the domain of human-machine interfaces, topics of special interest concern (1) development of principles for mapping the human sensorimotor system into nonanthropomorphic slave mechanisms (or the equivalent in virtual space) and (2) ability of the human sensorimotor system to adapt to alterations of normal sensorimotor loops caused by the presence of the interface.

1.2 Hearing Aid Research

Sponsor

National Institutes of Health
Grant 5 R01 DC00117

Project Staff

Professor Louis D. Braida, Lorraine A. Delhorne, Dr. Kenneth W. Grant, Dr. Christine M. Rankovic, Dr. Charlotte M. Reed, Dr. Kaoru Sekiyama, Dr. Rosalie M. Uchanski, Dr. Patrick M. Zurek, Bridget L. Coffman, Paul Duchnowski, Joseph A. Frisbie, Andrew L. Howe, Michael T. Keagy, Danielle G. LeMay, David S. Lum, Gregory R. Martin, Joseph A. Maxwell, Jason Mueller, Philip P. Nadeau, Matthew H. Power, Diane E. Ronan, Jennifer A. Wozniak

During the past year, our research on hearing aids has focused on (1) simulation of sensorineural hearing impairment, (2) basic studies of speech intelligibility, (3) speech production factors, (4) control of acoustic feedback, and (5) aids to speechreading.

The following is a description of our research in these areas.

1.2.1 Simulation of Sensorineural Hearing Impairment

Our previous work employed noise masking to simulate the effects of hearing loss. Zurek and Delhorne,[1] as well as Humes et al.[2] and Dubno and Schaefer,[3] found good agreement between masked-normal and hearing-impaired listeners' reception of consonants. Together with other findings,[4] these results support the view that the main difficulty in speech reception comes from audibility restrictions, as opposed to suprathreshold disorders as suggested by others,[5] at least for listeners with mild and moderate hearing losses.

While noise masking provides functionally accurate simulations of hearing impairment, it is phenomenologically unrealistic[6] and can be used only to simulate losses of less than roughly 70 dB because of the intense noise levels needed. We have developed and evaluated an alternate technique based on multiband expansion amplification[7] that simulates both threshold shift and recruitment without the use of masking noise. This simulation, a digitally-implemented elaboration of Villchur's scheme,[8] is completely specified by the audiogram of a particular hearing-impaired listener and empirical relations specifying the steepness of the recruitment function for a given hearing loss.

In initial evaluations, the performance of two normal-hearing subjects listening to the processed output of the simulator matched fairly well the consonant-reception scores of both the hearing-impaired and noise-masked normal listeners who had been previously tested by Zurek and Delhorne.[9] In more extensive evaluations, the performance of each of three severely hearing-impaired subjects on consonant and sentence intelligibility tests was compared to that of three normal-hearing subjects listening through the simulator. Generally, the simulation reproduced overall intelligibility scores and error patterns for individual impaired listeners very well when the frequency response was flat. When high-frequency emphasis placed more speech energy above threshold at high frequencies, the simulation gave better intelligibility than observed for two of the three impaired listeners. Overall, this study verified the equivalence of the AGC and noise simulations, explored the applicability of the AGC simulation for severe impairments, and identified frequency-specific deficits not accounted for by audibility and recruitment in some of the listeners.

---

1 P.M. Zurek and L.A. Delhorne, "Consonant Reception in Noise by Listeners with Mild and Moderate Hearing Impairment," *J. Acoust. Soc. Am.* 82: 1548-1559 (1987).

2 L.E. Humes, "An Evaluation of Several Rationales for Selecting Hearing Aid Gain," *J. Speech Hear. Res.* 51: 272-281 (1986).

3 J.R. Dubno and A.B. Schaefer, "Comparison of Frequency Selectivity and Consonant Recognition Among Hearing-impaired and Masked Normal-hearing Listeners," *J. Acoust. Soc. Am.* 91: 2110-2121 (1992).

4 J.R. Dubno and D.D. Dirks, "Factors Affecting Performance on Psychoacoustic and Speech-recognition Tasks in the Presence of Hearing Loss," in *Acoustical Factors Affecting Hearing Aid Performance II,* eds. G.A. Studebaker and I. Hochberg (Boston: Allyn and Bacon, 1993).

5 R. Plomp, "Auditory Handicap of Hearing Impairment and the Limited Benefit of Hearing Aids," *J. Acoust. Soc. Am.* 63: 533-549 (1978); B.R. Glasberg and B.C.J. Moore, "Psychoacoustic Abilities of Subjects with Unilateral and Bilateral Cochlear Hearing Impairments and Their Relationship to the Ability to Understand Speech," *Scand. Audiol. Suppl.* 32: 1-25 (1989).

6 D.P. Philips, "Stimulus Intensity and Loudness Recruitment: Neural Correlates," *J. Acoust. Soc. Am.* 82: 1-12 (1987).

7 P.M. Duchnowski, *Simulation of Sensorineural Hearing Impairment.* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1989; P.M. Duchnowski and P.M. Zurek, "Villchur Revisited: Another Look at AGC Simulation of Recruiting Hearing Loss," submitted to *J. Acoust. Soc. Am.*

8 E. Villchur, "Electronic Models to Simulate the Effect of Sensory Distortions on Speech Perception by the Deaf," *J. Acoust. Soc. Am.* 62: 665-674 (1977).

9 P.M. Zurek and L.A. Delhorne, "Consonant Reception in Noise by Listeners with Mild and Moderate Hearing Impairment," *J. Acoust. Soc. Am.* 82: 1548-1559 (1987).

1.2.2 Basic Studies of Speech Intelligibility

Although many current theories of speech perception assume that the perception of speech elements involves accurate cross-spectral comparisons, studies of spectral shape perception have revealed limits on our ability to make such comparisons. To gain insight into the role of cross-spectral comparisons in speech perception, we are studying the ability of normal-hearing listeners to combine cues across frequency bands when identifying speech segments. This research involves both speech identification experiments and analysis of the results of these experiments using the pre-labeling and post-labeling models[10] that were initially applied to studies of audiovisual integration.

When bands arising from lowpass and highpass filtering of consonants are considered,[11] the situation is similar to the audiovisual case: the pre-labeling models gives a good account of the combined band (i.e., wideband) score, while the post-labeling model underestimates it. To evaluate the models' abilities to predict cross frequency-band integration, we undertook a series of consonant-identification experiments in which the stimuli were filtered into disjoint bands, which were presented both individually and in various combinations. In the first experiment, the stimuli were CV syllables, the bands were (A) 0-0.7, (B) 0.7-1.4, and (C) 1.4-2.8 kHz, and the combined bands were A+B and A+C. Bands were combined both diotically and dichotically to evaluate peripheral interference effects. In the second experiment, the stimuli were CVC syllables, the bands were (A) 0.1-1.4, (B) 1.4-2.1, (C) 2.1-2.8, and (D) 2.8-5.5 kHz, and the combined bands (diotic presentation only) were all 15 combinations of the individual bands.

In each experiment a group of five young adult listeners with normal hearing was tested. We derived predictions for the identification score in the combined-band conditions based on performance in the single-band conditions using the following approaches: (1) the ANSI 20-band AI calculation procedure;[12] (2) the ANSI one-third octave AI calculation procedure; (3) Fletcher's rule in which the error probability for combined bands is the product of the error probabilities for the individual bands; (4) Fletcher's rule applied to individual consonants rather than the overall score; (5) the fuzzy logical model of perception;[13] (6) the pre-labeling model; and (7) the post-labeling model.

Results from the CV experiment indicate that the score in the combined band condition was generally only slightly greater than the higher score obtained for the two constituent bands. This result is inconsistent with predictions based on the articulation index. The predictions of the post-labeling model are closest to observed scores, and are statistically indistinguishable from the data in roughly half the cases. The other models generally overestimate performance in the combined-band conditions. Scores for one listener were lower in the two diotic combined-band conditions than for the better constituent band, but were greater in the dichotic combined-band conditions, suggesting that peripheral interference may have occurred in the diotic condition.

Overall scores in the CVC experiment are fairly well predicted by the Articulation Index calculation, and all of the rules considered except the post-labeling model, which tended to underestimate scores in the combined-band conditions. The other models predicted overall scores satisfactorily in five to seven of ten possible cases: (five subjects, initial and final consonants). When identification scores for individual consonants are considered, the Fletcher Rule applied to individual consonants and the pre-labeling model gave the closest predictions.

Although the results of these two experiments suggest that different integration mechanisms may apply to low- and high-frequency bands of speech, they are consistent with respect to the (lack of) need for explicit cross-spectral comparisons in consonant identification. Identification scores for combinations of bands generally can be adequately accounted for without requiring such comparison mechanisms.

We have recently developed an automatic speech recognition system for use in studies of across-

---

10 L.D. Braida, "Crossmodal Integration in the Identification of Consonant Segments," *Quart. J. Expt. Psych.* 43A(3): 647-677 (1991).

11 G.A. Miller and P. Nicely, "An Analysis of Perceptual Confusions Among Some English Consonants," *J. Acoust. Soc. Am.* 27: 328-352 (1955).

12 *American National Standard Methods for the Calculation of the Articulation Index,* ANSI S3.5-1969 (New York: American National Standards Institute, 1969).

13 D.W. Massaro, *Speech Perception by Ear and Eye: A Paradigm for Psychological Inquiry,* (Hinsdale, N.J.: Lawrence Earlbaum Assoc., 1987.

band integration phenomena.[14] The system identifies speech sounds by combining estimates made in each of four non-overlapping frequency bands. Each of the four sub-recognizers used in this scheme is a conventional Hidden Markov Model system that operates upon a simpler signal than wideband speech. The processing required for recognition is thus potentially less complicated than in conventional recognizers because only a relatively narrow spectral region need be represented for each recognizer, and the training required may be substantially reduced. The bands used in the initial implementation were 0.0-0.7, 0.7-1.4, 1.4-2.8, and 2.8-5.0 kHz. The bandpass signal characterizations studied included cepstral, LPC, and autocorrelation coefficients, with both static parameter values and time-differences used to represent the signals. Integration across frequency bands was accomplished by estimating the joint distribution of band estimates for each speech sound. Five methods were evaluated: statistical independence, pairwise independence, log-linear modeling, pseudo-bayseian modeling, and occurrence-weighted interpolation. The highest recognition accuracy (58.5 percent correct recognition of 39 speech sounds in continuous speech spoken by a large number of talkers in the TIMIT database) was achieved using the Cepstral representation and the Occurrence-Weighted Interpolation combination method. Comparable results obtained by more conventional recognizers that operate on wideband speech range from 55 percent[15] to 64 percent.[16] While the results obtained thus far using the four-band system are only comparable to those for conventional systems, they are nevertheless highly encouraging, given the relatively early state of development of the approach. In addition, they are consistent with the results of our perceptual studies. The speech sounds of English (including both consonants and vowels) can be recognized with high accuracy in running speech without relying heavily on cross-spectral comparisons.

We have recently completed our preliminary investigation[17] of a new computational approach for determining the intelligibility of speech subjected to waveform degradations or signal-processing transformations. The approach incorporates a model of the human auditory system that derives a sequence of discrete symbols from the speech waveform. Each symbol represents a prototypical vector of parameter values characteristic of the short-term spectrum measured in a single frame (10 ms) of speech. The perceptual effect of the degradation is estimated by assessing the consistency between the symbol sequence derived from an untransformed (input) speech signal and that derived from a transformed (output) signal. This is implemented via calculation of percent transmitted information (PTI), hypothesized to be monotonically related to intelligibility for a given auditory model, listener, and speech test. Two computational models of auditory system function were studied: one (CBL) based on critical-band levels[18] and the other (EIH) based on the ensemble interval histogram representation[19] of auditory nerve firing patterns. Degradations studied thus far include linear filtering and additive noise, whose effects on speech intelligibility are relatively well understood. For both the CBL and EIH models, PTI decreases monotonically as the severity of filtering or level of additive noise is increased. Since intelligibility is assumed to be monotonically related to PTI, the crossover frequency (equal PTI for highpass and lowpass filtering) and the value of S/N which yields the PTI at the crossover frequency, characterize the performance of the auditory models. For the CBL model, a PTI of 55 percent is achieved at a crossover fre-

---

14 P.M. Duchnowski, *A Novel Structure for Speech Segment Recognizers,* Ph.D. diss. Dept. of Electr. Eng. and Comput. Sci., MIT, 1993.

15 V. Zue, "Recent Progress on the SUMMIT System," *Proceedings of the Third DARPA Workshop on Speech and Natural Language,* 1990, pp. 380-384.

16 K.-F. Lee, and H.-W. Hon, "Speaker Independent Phone Recognition Using Hidden Markov Models," *IEEE Trans. Acoust. Speech Sig. Process.* 37(11): 1641-1648 (1989).

17 M.H. Power, *A Physical Measure of Consistency Among Speech Parameter Vectors: Application to Speech Intelligibility Determination,* Ph.D. diss., Dept. Electr. Eng. and Comput. Sci., MIT, 1993; M.H. Power and L.D. Braida, "A Physical Measure of Consistency Among Speech Parameter Vectors: Application to Speech Intelligibility Determination," *J. Acoust. Soc. Am.* 90: S2327 (1991).

18 R. Plomp, "Timbre as a Multidimensional Attribute of Complex Tones," in *Frequency Analysis and Periodicity Detection in Hearing,* Sijthoff, Leiden, The Netherlands, 1970; M. Florentine and S. Buus, "An Excitation-Pattern Model for Intensity Discrimination," *J. Acoust. Soc. Am.* 70: 1646-1654 (1981); C.L. Farrar, C.M. Reed, N.I. Durlach, L.A. Delhorne, P.M. Zurek, Y. Ito, and L.D. Braida, "Spectral-shape Discrimination. I. Results from Normal-hearing Listeners for Stationary Broadband Noise," *J. Acoust. Soc. Am.* 81: 1085-1092 (1987).

19 O. Ghitza, "Temporal Non-Place Information in the Auditory-Nerve Firing Patterns as a Front-End for Speech Recognition in a Noisy Environment," *J. Phonetics* 16: 109-123 (1988).

quency of 1800 Hz and also for S/N = + 23 dB. For the EIH model, a PTI of 40 percent is achieved at a crossover frequency of 1000 Hz and also for a S/N = +10 dB. Thus, neither the CBL nor the EIH model provides completely satisfactory accounts of the dependence of intelligibility on the degradations considered. The CBL model predictions were much more sensitive to the effects of noise, and EIH predictions to those of highpass filtering.

1.2.3 Clear Speech

Sentences spoken clearly are more intelligible than those spoken conversationally for (1) hearing impaired listeners in a quiet background,[20] (2) normal-hearing listeners in a noisy background,[21] and (3) both normal-hearing and hearing-impaired listeners in reverberant and noisy-reverberant backgrounds.[22] In a recent study of normal-hearing listeners we tested the interaction between three distortions multiplicative noise;[23] 500 Hz lowpass filtering; and reverberation with speaking rate on intelligibility. We used two sets of materials Clarke sentences[24] and Harvard sentences[25] which were recorded by two teachers of the deaf as part of our study of manual cued speech.[26] Each type of material was spoken both while the teachers produced manual cues (speaking rate of 100 wpm) and while not producing clues (150 wpm). This difference in speaking rates is only about half that between clear and conversational speech, and little exaggeration of articulation is evident in the speech produced while cueing. Roughly equal intelligibility test scores were obtained for the two speaking rates for all conditions with the exception that the IEEE sentences were less affected by reverberation at the slower speaking rate than at the faster one. These results suggest that there may be an interaction between speaking style and type of distortion on the intelligibility advantage associated with speaking clearly. Such an interaction should aid greatly in identifying the factors responsible for the difference in intelligibility between various speaking styles.

1.2.4 Control of Acoustic Feedback

Acoustic feedback imposes a limit on hearing aid gain that can result in low-level signals not being amplified to reach audibility. Although several techniques for reducing feedback have been proposed recently, evaluations of these techniques have been limited to measurements of the increase in stable gain that can be achieved; signal distortions introduced by the processing have not been considered, and measurement conditions have differed from study to study. We have implemented the most promising signal processing techniques and compared them not only for stable gain but also for their effects on sound quality and annoyance. In the process, we have developed a novel technique for reducing the effects of feedback.[27]

The five feedback-reduction algorithms examined were an adaptive notch filter and four variants of adaptive cancellation. The adaptive notch method has been proposed previously for use in public-address systems but has not been evaluated in the hearing aid application, though it has been used in an algorithm for detecting oscillation.[28] The adaptive cancellation methods have in common the goal of estimating the feedback path in order to cancel it

20 M.A. Picheny, N.I. Durlach, and L.D. Braida, "Speaking Clearly for the Hard of Hearing I: Intelligibility Differences between Clear and Conversational Speech," *J. Speech Hear. Res.* 28: 96-103 (1985).

21 R.M. Uchanski, *Spectral and Temporal Contributions to Speech Clarity for Hearing Impaired Listeners,* Ph.D. diss., Dept. of Electr. Eng. and Comput. Sci., MIT, 1988.

22 K.L. Payton, R.M. Uchanski, and L.D. Braida, "Intelligibility of Conversational and Clear Speech in Noise and Reverberation for Listeners with Normal and Impaired Hearing," *J. Acoust. Soc. Am.* 95: 1581-1592 (1994).

23 M.R. Schroeder, "Reference Signal for Signal Quality Studies," *J. Acoust. Soc. Am.* 44: 1735-1736 (1968).

24 M.E. Magner, *A Speech Intelligibility Test for Deaf Children,* (Northampton, Mass.: Clarke School for the Deaf, 1972).

25 Institute of Electrical and Electronics Engineers. *IEEE Recommended Practice for Speech Quality Measurements.* Publication No. 297 (New York: IEEE, 1969).

26 R.M. Uchanski, L.A. Delhorne, A.K. Dix, L.D. Braida, C.M. Reed, and N.I. Durlach, "Automatic Speech Recognition to Aid the Hearing Impaired. Prospects for the Automatic Generation of Cued Speech," *J. Rehab. Res. and Dev.* Spring 1994, forthcoming.

27 J.A. Maxwell, *Acoustic Feedback in Hearing Aids,* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1993; J.A. Maxwell and P.M. Zurek, "Reducing Acoustic Feedback in Hearing Aids," in preparation.

28 J.M. Kates, "Feedback Cancellation in Hearing Aids: Results from a Computer Simulation," *IEEE Trans. Sig. Proc.* 39: 553-562 (1991).

with another out-of-phase feedback path. In the systems described by Bustamante et al.[29] and Engebretsen et al.,[30] the adaptation takes place continuously. Kates' approach differs in that the system monitors the signal for oscillation and, if detected, interrupts the main signal path and inserts a probe noise for feedback-path estimation.[28] The fifth system[31] also interrupts the signal, but it does so not only when oscillation is detected, but also during selected time intervals when the input signal is estimated to be low. This approach of quiet-interval adaptation has four clear advantages: (1) it provides better cancellation because the misadjustment caused by the input signal is minimized; (2) it provides a quasi-continuous, but stable, tracking of changes in the feedback path; (3) it achieves cancellation with minimal disturbance to the signal; and (4) minimal levels of probe noise are required.

All five methods were implemented in real time with a Motorola DSP96002 DSP and evaluated for increased stable gain (measured physically) and speech quality (measured by subjective ratings). Acoustic feedback paths (two static and one time-varying) were simulated electrically. Results showed the following: First, a single adaptive notch can be effective, but only if the feedback is confined to a single frequency region that is narrower than the notch width. Second, the continuously-adapting systems of Bustamante et al.[32] and Engebretson et al.[33] and, as far as we can determine, the system described by Dyrlund and Bisgaard,[34] which is now available in commercial hearing aids, are inherently unstable. This instability manifests as bursts and warbling distortions that can be minimized only at the expense of degree of cancellation. Third, the noncontinuously adapting systems and the quiet-interval system[35] perform equivalantly only under static, quiet conditions. When the feedback path changes and/or when adaptation occurs in the presence of speech, the advantages of the quiet-interval method are apparent. Overall, the quiet-interval system, by providing maximal feedback cancellation with minimal disturbance to the user, represents a substantial advance over the other techniques.

1.2.5 Aids to Speechreading

During the past grant year, our research on aids to speechreading focused the development of low bandwidth acoustic signals that enhance speechreading for listeners with severe hearing impairments. One class of such signals consists of tones that are amplitude modulated by the envelopes of filtered bands of speech. These signals have been shown to be effective supplements to speechreading by normal-hearing listeners.[36] However, when the center frequency of the carrier bands is lowered to match the residual hearing of listeners with severe high-frequency loss, the signals no longer successfully supplement speechreading.[37] A

[29] D.K. Bustamante, T.L. Worrall, and M.J. Williamson, "Measurement of Adaptive Suppression of Acoustic Feedback in Hearing Aids," *Proceedings of ICASSP-89,* 1989, pp. 2017-2020.

[30] A.M. Engebretson, M.P. O'Connell, and F. Gong, "An Adaptive Feedback Equalization Algorithm for the CID Digital Hearing Aid," *Proceedings of the 12th Annual International Conference on IEEE Engineering in Medicine and Biology Society,* 1990, pp. 2286-2287.

[31] J.A. Maxwell, *Acoustic Feedback in Hearing Aids,* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1993.

[32] D.K. Bustamante, T.L. Worrall, and M.J. Williamson, "Measurement of Adaptive Suppression of Acoustic Feedback in Hearing Aids," *Proceedings of ICASSP-89,* 1989, pp. 2017-2020.

[33] A.M. Engebretson, M.P. O'Connell, and F. Gong, "An Adaptive Feedback Equalization Algorithm for the CID Digital Hearing Aid," *Proceedings of the 12th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,* 1990, pp. 2286-2287.

[34] O. Dyrlund and N. Bisgaard, "Acoustic Feedback Margin Improvements in Hearing Instruments Using a Prototype DFS (Digital Feedback Suppression) System," *Scand. Audiol.* 20: 49-53 (1991).

[35] J.M. Kates, "Feedback Cancellation in Hearing Aids: Results from a Computer Simulation," *IEEE Trans. Sig. Proc.* 39: 553-562 (1991).

[36] M. Breeuwer and R. Plomp, "Speechreading Supplemented with Frequency-selective Sound-pressure Information," *J. Acoust. Soc. Am.* 76: 686-691 (1984); M. Breeuwer and R. Plomp, "Speechreading Supplemented with Auditorily Presented Speech Parameters," *J. Acoust. Soc. Am.* 79: 481-499 (1986); K.W. Grant, L.D. Braida, and R.J. Renn, "Single Band Amplitude Envelope Cues as an Aid to Speechreading," *Quart. J. Exp. Psych.* 43A(3): 621-645 (1991).

[37] K.W. Grant, L.D. Braida, and R.J. Renn, "Single Band Amplitude Envelope Cues as an Aid to Speechreading," *Quart. J. Exp. Psych.* 43A(3): 621-645 (1991); K.W. Grant, L.D. Braida, and R.J. Renn, "Auditory Supplements to Speechreading: Combining Amplitude Envelope Cues from Different Spectral Regions of Speech," *J. Acoust. Soc. Am.* 95: 1065-1073 (1994).

second class of low-bandwidth supplements consists of tones that are modulated in frequency by estimates of the time-varying fundamental frequency of the speech signal and modulated in amplitude by the broadband speech envelope.

Takeuchi and Braida[38] used a match-to-sample task to study the effect of a spectrally remote amplitude-modulated signal on the ability of listeners to compare two patterns of amplitude modulation in the presence of a distractor. Discrimination performance was measured as a function of the correlation between the envelopes of the comparison and distractor signals and difference between the carrier frequencies used to convey these envelopes. Although there are some differences across listeners, the addition of a distractor, even if it was spectrally remote from the comparison signal, degraded discrimination. The degree of interference tended to increase as the degree of correlation between comparison and distractor envelopes decreased. Performance on conditions with small carrier-frequency shifts between the comparison and distractor signals was generally worse than on conditions with wide spacing. When the comparison was higher in frequency than the distractor, there was a larger effect of frequency spacing than when the comparison was lower than the distractor. These results suggest that there may be some impairment to the perception of the envelopes of individual speech bands when an additional band is presented simultaneously. Since this impairment is somewhat greater when the bands are closely spaced in frequency, hearing-impaired listeners with reduced frequency selectivity may be more affected by the simultaneous presentation of multiple bands than listeners with normal hearing. It is noteworthy, however, that even on the conditions that produced the most impairment, performance generally did not fall below 70 percent of the level with no added signal.

An alternative acoustic signal that has been proposed as a speechreading supplement is based on conveying the voice fundamental frequency (F0) and the amplitude envelope of the broadband speech waveform. For research purposes, this signal can be based on accelerometer measurements of glottal vibrations and pressure-microphone measurements of the acoustic speech signal. To compare the effectiveness of such supplements with those based on the envelopes of bands of speech, we have developed a system capable of synthesizing a sinusoidal waveform whose amplitude is controlled by the envelope of the speech waveform and whose frequency is controlled by the variation in glottal-pulse periods.[39] Since the system is implemented in software on a real-time signal processor, both the mapping from speech envelope amplitude to sinusoidal amplitude and the mapping from glottal pulse amplitude to tone frequency can be specified flexibly to accommodate the hearing loss and reduced dynamic range of impaired listeners.[40] Using a system-produced supplement that incorporates only frequency modulation with a linear mapping of F0 to tone frequency, five relatively untrained listeners with normal hearing showed improved scores for word recognition in sentence context (Clarke Sentences)[41] from 20 to 39 percent, roughly comparable to the benefit expected from a fixed-frequency tone amplitude modulated by the envelope of an octave band of speech.

1.2.6 Publications

Dubno, J.R., and D.D. Dirks. "Factors Affecting Performance on Psychoacoustic and Speech-recognition Tasks in the Presence of Hearing Loss." In *Acoustical Factors Affecting Hearing Aid Performance II.* Eds. G.A. Studebaker and I. Hochberg. Boston: Allyn and Bacon, 1993.

Duchnowski, P.M. *A Novel Structure for Speech Segment Recognizers.* Ph.D. diss. Dept. of Electr. Eng. and Comput. Sci., MIT, 1993.

Duchnowski, P.M., and P.M. Zurek. "Villchur Revisited: Another Look at AGC Simulation of Recruiting Hearing Loss." Submitted to *J. Acoust. Soc. Am.* 1993.

Grant, K.W., L.D. Braida, and R.J. Renn. "Auditory Supplements to Speechreading: Combining Amplitude Envelope Cues from Different Spectral Regions of Speech." *J. Acoust. Soc. Am.* 95: 1065-1073 (1994).

---

38 A. Takeuchi and L.D. Braida, "Recognition of Envelope Patterns in the Presence of a Distractor," submitted to *J. Acoust. Soc. Am.*

39 M.T. Keagy, *A Voice Fundamental Frequency Extractor to Aid Speechreading,* S.B. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1993.

40 K.W. Grant, "Identification of Intonation Contours by Normally Hearing and Profoundly Hearing-impaired Listeners," *J. Acoust. Soc. Am.* 82: 1172-1178 (1987).

41 M.E. Magner, *A Speech Intelligibility Test for Deaf Children* (Northampton, Mass.: Clarke School for the Deaf, 1972).

Keagy, M.T. *A Voice Fundamental Frequency Extractor to Aid Speechreading.* S.B. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1993.

Maxwell, J.A. *Acoustic Feedback in Hearing Aids.* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1993.

Payton, K.L., R.M. Uchanski, and L.D. Braida. "Intelligibility of Conversational and Clear Speech in Noise and Reverberation for Listeners with Normal and Impaired Hearing." *J. Acoust. Soc. Am.* 95: 1581-1592 (1994).

Power, M.H. *A Physical Measure of Consistency Among Speech Parameter Vectors: Application to Speech Intelligibility Determination.* Ph.D. diss., Dept. Electr. Eng. and Comput. Sci., MIT, 1993.

Takeuchi, A., and L.D. Braida. "Recognition of Envelope Patterns in the Presence of a Distractor." Submitted to *J. Acoust. Soc. Am.*

Uchanski, R.M., L.A. Delhorne, A.K. Dix, L.D. Braida, C.M. Reed, and N.I. Durlach. "Automatic Speech Recognition to Aid the Hearing Impaired. Prospects for the Automatic Generation of Cued Speech." *J. Rehab. Res. and Dev.* Spring 1994. Forthcoming.

1.3 Cochlear Prostheses

Sponsor

National Institutes of Health
Contract 2 P01 DC00361[42]

Project Staff

Professor Louis D. Braida, Lorraine A. Delhorne, Dr. Donald K. Eddington, Dr. William M. Rabinowitz

The development of cochlear prostheses is directed at aiding individuals with profound sensorineural deafness who are unable to derive useful benefits from acoustic input to the ear. The prosthesis bypasses the impaired acoustic tranduction mechanism and directly stimulates surviving auditory nerve fibers with currents delivered via an electrode array implanted within the cochlea. The overall goals of our research are to understand the mechanisms responsible for the improved hearing provided by these prostheses and to exploit this understanding for the development of improved systems. Our program makes use of a dedicated group of postlingually deafened adults implanted with a multichannel cochlear prosthesis (Ineraid, Richards Medical Corporation), who participate in intensive multifactorial studies. Our research capitalizes on having direct accessibility to the implanted electrode array through a percutaneous connector.

During the past year, work has involved (1) analysis of cue integration in audiovisual speech reception, (2) acoustic simulations of cochlear-implant speech reception, and (3) alternative speech processing for improved implant performance. The work in (3) is performed with Joseph Tierney and Marc Zissman; progress is described in part 5, section 3, chapter 1, section 1.7.3.

The work on audiovisual integration assesses the ability of an implantee to combine cues that are available from separately using vision and audition. Because most implantees require audiovisual input for reliable communication, analysis of integration is of particular significance.

A series of experiments on audiovisual integration were continued. Closed-set identification for a set of ten vowels, a set of the 12 most frequently occurring consonants, and the full set of 24 consonants is being tested. The vowels are tested in a medial position, within nonsense /b/-V-/d/ syllables; the consonants are tested in initial position, within nonsense C1-/a/-C2 syllables. Stimuli are from a single female talker, with multiple tokens for each stimulus, and are presented through a computer-controlled videodisc. Results are being obtained using vision alone, audio (the implant) alone, and audiovisually. Most subjects have completed the ten-vowel and 12-consonant tests. The 24-consonant tests require considerably more trials for reliable estimates of confusion patterns for each of the three test modes; these tests are in varying states of completion for the subjects. Analyses to determine the quantitative efficiency of audiovisual integration are underway. These analyses exploit an analytical framework[43] that quantifies how performance in a combined mode (e.g., audiovisual stimulation) is influenced by integration or interference of cues available in the isolated component modes (e.g., audio and visual stimulation alone).

[42] Subcontract from Massachusetts Eye and Ear Infirmary. Dr. Joseph B. Nadol, M.D., Principal Investigator.

[43] L.D. Braida, "Crossmodal Integration in the Identification of Consonant Segments," *Quart. J. Exp. Psychol.* 43A: 647-677 (1990).

The work in acoustic simulations is designed to test our understanding of the cues that implantees receive when using the prosthesis alone. Speech is processed to remove specific details and the resulting waveform serves as a new sound input for speech testing. Insofar as an implantee's speech reception is unchanged, the details that were removed must have been either unimportant for speech reception or were not being received by the subject.

Several lines of evidence suggest that Ineraid users perceive information about the fine structure of the speech waveform only up to a limited frequency, roughly 300 to 1000 Hz depending upon the subject. Information about higher frequency waveform components is mediated principally by the envelopes of the bandpass signals that are delivered to the electrodes. To test this hypothesis directly, experiments have begun in which the signals normally delivered to the three upper channels of the implant are replaced by their envelopes modulating tones at the respective center frequencies of the bands. The low frequency portion of the spectrum (0 to 500 Hz) was unmodified. The modified system has been tested with a single subject who exhibits exceptionally good performance with his normal system (87 percent-correct on 12-consonant identification and 49 percent-correct on whole-word identification). His performance with the modified system was 72 and 26 percent-correct on the consonant and word tests, respectively. These scores are below his normal performance, but they are remarkably good, exceeding that of the average implantee using the normal system (about 55 percent on consonants and 15 percent on words).[44] Performance that was much worse (14 percent on consonants) was obtained with this same subject when the three upper-band signals were removed, and only the lowpass speech (0 to 500 Hz) was presented. These results indicate that the three upper-band envelope signals can provide spectral-shape information that is very useful and that continued study with this processing model would be fruitful.

1.4 Tactile Communication of Speech

Sponsor

National Institutes of Health
Grant 5 R01 DC00126

Project Staff

Dr. Charlotte M. Reed, Gerald L. Beauregard, Lorraine A. Delhorne, Gail Denesvich, Nathaniel I. Durlach, Ashanthi Gajaweera, Irene Kotok, Charlie Pan, Geoffrey L. Plant, Dr. William M. Rabinowitz, Dr. Mandayam A. Srinivasan, Hong Z. Tan, Andrew Ugarov

The ultimate goal of this research is to develop tactual aids for the deaf and deaf-blind that can serve as substitutes for hearing in speech communication. The objectives and specific aims of our research are as follows:

1. Basic study of encoding and display schemes to develop methods of displaying acoustic signals to the tactual sense for optimal information transfer.
2. Research on tactual systems designed specifically to aid speechreading, including systems based on both acoustic and articulatory-based signal processing.
3. Evaluations of experienced deaf users of portable, wearable tactual aids to determine improvements to speech reception provided by these aids and to compare this performance to that of users of other types of auditory prostheses.

1.4.1 Basic Study of Encoding and Display Schemes

During the current year, work has been conducted on five projects concerned with increased understanding of the proprioceptive/kinesthetic as well as the cutaneous sensory system of the hand.

[44] W.M. Rabinowitz, D.K. Eddington, L.A. Delhorne, and P.A. Cuneo, "Relations Among Different Measures of Speech Reception in Subjects Using a Cochlear Implant," *J. Acoust. Soc. Am.* 92: 1869-1881 (1992).

Amplitude and Frequency Resolution for Motional Stimulation

This study is concerned with characterizing psychophysical performance for low-frequency (1-32 Hz), high-amplitude (up to 2 cm) signals delivered to the fingerpad using a position-controlled servomotor. Measurements of amplitude discrimination have been obtained on three subjects as a function of frequency for several values of reference displacement and stimulus duration. The amplitude jnd (which is roughly independent of duration in the range of 250-1000 msec) appears to decrease slightly (from roughly 1.1 to 0.6 dB) as frequency increases from 2 to 32 Hz. The value obtained at 32 Hz is similar to results reported in the literature for vibrational frequencies. Measurements of frequency discrimination are complicated at very low frequencies by the fact that durational differences, as well as frequency differences, can be used as a cue for discriminating F and F+ ΔF. To limit performance solely to frequency cues, the overall duration of the reference and comparison stimuli are being randomized within a run of trials. Preliminary results for frequency discrimination indicate that, for d'=1, $\Delta f/f$ is roughly 0.1 across the frequency range 4-16 Hz. The difference limen reported in the literature for frequencies in the range 30-200 Hz increases with frequency from roughly 0.3 to 0.5. Thus, frequency discrimination at very low frequencies appears to be somewhat better than at the higher frequencies.

Tactual Reception of Morse Code Sequences

This study, which employs the same motion-stimulation device as described above, is concerned with the tactual reception of Morse Code in subjects who are either highly skilled or naive in the sending/receiving of Morse Code. Computer-generated sequences of Morse Code are delivered to the fingertip using high-amplitude square-wave stimulation. Reception ability is being studied as a function of the rate of presentation (12, 16, and 24 words/minute) and the length of the stimulus stream (isolated letters, random three-letter strings, common words, and conversational sentences). Results for two subjects who are already highly experienced in the sending and in the auditory reception of Morse Code indicate faster learning curves and higher levels of tactual performance than those obtained for naive subjects. For example, the experienced subjects achieved criterion performance of 95 percent on the single-letter recogni1ion task at all three presentation rates within a total of 10,000 trials. For the naive subjects, learning is proceeding at a slower pace, and differences between the skilled and naive subjects appear to increase with rate. While the skilled subjects were able to receive CUNY sentences nearly perfectly at a rate of 16 words/minute, the unskilled subjects have not advanced beyond the single-letter task. The auditory reception of sentences in Morse Code by the skilled subjects, however, is superior to their tactual reception since they can obtain near-perfect scores for auditory reception of the code at rates of 40-50 words/minute.

Discrimination of Thickness

Study of the discrimination of thickness has continued over the current year with the development of a device and software for experimentation and with collecting preliminary data on several subjects. The ability to discriminate thickness is being studied for three materials with different bending moduli (steel, brass, and plastic, in decreasing order of stiffness). Various steps have been taken to eliminate potential cues other than thickness in performing the discrimination task, including the use of different tokens (i.e., individual plates) for each material and thickness, limiting the area of the plate in which the thumb and forefinger may make contact, and placing the apparatus in a temperature-controlled box to minimize the use of thermal cues. On a given trial, two plates of different thickness (but of the same material) are presented sequentially, and the subject's task is to judge the order of presentation (thick-thin or thin-thick). Training and testing has been conducted thus far with 8-mil reference plates and with comparison values in the range of 8-20 mils. Preliminary results indicate a possible correlation between stiffness and jnd for thickness in that the jnd for steel (the stiffest of the three materials) is higher than for brass and plastic (presumably due to additional cues supplied by compression of the fingerpads as the material bends).

Design of Multifinger Tactual Stimulator

Work has been conducted on the design of a new multifinger device for use in studies of the information capabilities of the tactual (cutaneous plus kinesthetic) sense. The specifications for this device include simultaneous stimulation of three to five fingers with one degree of freedom in movement for each finger. The frequency range of stimulation is specified as 0-300 Hz, with maximum motion of one inch peak-to-peak at very low frequencies decreasing to roughly 100 microns for frequencies above 100 Hz. The geometric constraints on the display are dictated by the shape of the hand.

Study of Manual Compliance

New experiments have been completed[45] on compliance discrimination in which work cues were eliminated and force cues were minimized. When compared to data obtained previously in which both work and terminal force cues, as well as compliance cues, were available to the subject, the new data indicate (a) that manual resolution of compliance deteriorates when force and/or work cues are reduced or eliminated and (b) that subjects appear to rely strongly on terminal force cues (when they are available) in making compliance comparisons.

1.4.2 Tactual Supplements to Speechreading

Work during the current year has been concerned with developing a more complete understanding of the differential results obtained for auditory and tactile presentation of acoustic-based low-bandwidth supplements to speechreading. Our previous work has been extended in two areas described below.

Comparison of Auditory and Tactile Presentation of a Single-band Envelope Cue as a Supplement to Speechreading

Further experiments have been conducted using a low bandwidth supplemental speechreading cue. The envelope of an octave band of speech centered at 500 Hz is used to modulate the amplitude of a 200-Hz tone. The usefulness of this cue for speechreading, when presented either auditorily through headphones or tactilely through a single-channel vibrator, was evaluated for a set of 24 initial consonants in CVC syllables and a set of 16 vowels and diphthongs in CVC syllables.[46] Consonant results, averaged across three subjects, showed improvement over lipreading alone (53 percent-correct performance) of 10 percentage points for the tactile supplement and 22 percentage points for the auditory supplement. The improvement observed with the supplements appears to be due primarily to improved reception of the voicing feature for the tactile condition and of the voicing and nasality features for the auditory condition. Vowel results indicated near-ceiling performance on all three conditions. The segmental results will be related to those obtained previously for sentence materials.

Tactile and Auditory Measures of Amplitude-modulation Discrimination

One possible explanation for the better results obtained for auditory compared to tactile presentation of the same supplemental cue to speechreading may be differences in basic psychophysical properties of hearing and touch. In particular, the ability to discriminate differences in amplitude modulation may be highly relevant to the supplemental cue we are studying. Experiments are being conducted to assess AM discrimination as a function of modulation depth ($m=0$ and 1) and modulation frequency ($f_m=5$ or 50 Hz) for a 200-Hz carrier. Results indicate that, on average, auditory thresholds are roughly 6 dB more sensitive than tactile thresholds for a given set of stimulus parameters, with the exception of the condition $m=0$ and $f_m=50$, where auditory and tactile performance are nearly the same. Our results extend those previously available in the literature by including data for nonzero values of reference modulation, which may be particularly relevant to the information that is conveyed in the single-band speech envelope.

1.4.3 Evaluation of Practical Aids

In addition to continuing periodic laboratory evaluations of eight subjects who have been fitted with tactile devices and are wearing them in the field,[47] two other studies have been initiated during the current year.

---

45 H.Z. Tan and N.I. Durlach, "Manual Discrimination Using Active Finger Motion: Compliance, Force, and Work," submitted to *Perception and Psychophysics*; H.Z. Tan, N.I. Durlach, Y. Shao, and M. Wei, "Manual Resolution of Compliance When Work and Force Cues are Minimized," *Proceedings of the American Society of Mechanical Engineers WAM'93* (Book H00851), 99-104, 1993.

46 M.S. Bratakos, *Supplements to Speechreading: A Comparison of Auditory and Tactile Presentation of a Single-Band Envelope Cue*, S.B. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1993.

47 C.M. Reed, L.A. Delhorne, and N.I. Durlach, "Results Obtained with Tactaid II and Tactaid VII," in *Proceedings of Second International Conference on Tactile Aids, Hearing Aids, and Cochlear Implants*, eds. A. Risberg, S. Felicetti, G. Plant, and K.-E. Spens (Stockholm, Sweden: Department of Speech Communication and Music Acoustics, KTH, 1992), pp. 149-155.

The Effects of Training on Adult Performance Using the Tactaid VII

Two additional adults have been fitted with the Tactaid VII and are participating in weekly training sessions in the use of the device as a supplement to speechreading. Pre-training measurements were obtained on the subjects' speechreading ability in two tasks (sentence reception and tracking) both with and without the tactile aid. A detailed training program using both analytic and synthetic training materials has been prepared for use in the program, emphasizing the use of the tactile aid as a supplement to lipreading but also including exercises for the tactile reception of speech alone. After a six-month training period, the speechreading tests will be administered again to determine the effects of training. Performance will then be compared with the group of adults fitted with the Tactaid VII but provided with only basic aid orientation and limited training.

The Effects of Training on the Use of Tactile Aids by Congenitally Deaf Children

This study involves long-term training of a group of congenitally deaf children (age range five to ten years) attending the Boston School for the Deaf. These children, who derive limited benefit from amplification (hearing aids and FM trainers), are provided with weekly therapy sessions using the Tactaid VII. The training sessions emphasize the use of the tactile aid for speech development, auditory/tactile speech reception, and the use of the tactile aid as a lipreading supplement. Each child's speech production and speech reception skills were assessed at the beginning of the program and will be retested at three month intervals. An additional child, who has normal hearing up to 1.5 kHz but no measureable hearing above 2 kHz, is receiving training on a modified tactile aid which provides information on high-frequency consonants. Speech-production tests consisting primarily of materials that include sounds with high-frequency energy were administered prior to the start of a training program with the aid, and will be re-administered at three-month intervals.

1.4.4 Significance

The research conducted here leads to increased understanding of tactual speech communication and contributes to the development of improved tactual aids for individuals who are deaf and deaf-blind. This research plays a significant role in helping such individuals to achieve improved speech reception, speech production, and language competence. At the more basic scientific level, this research contributes to increased knowledge about speech communication, tactual perception, manual sensing, display design, and sensory substitution.

1.4.5 Publications

Reed, C.M., L.A. Delhorne, and N.I. Durlach. "Results Obtained with Tactaid II and Tactaid VII." *Proceedings of the Second International Conference on Tactile Aids, Hearing Aids, and Cochlear Implants*, pp. 149-155, Eds. A. Risberg, S. Felicetti, G. Plant, and K.-E. Spens. Stockholm, Sweden: Department of Speech Communication and Music Acoustics, KTH, 1992.

Tan, H.Z., and N.I. Durlach. "Manual Discrimination Using Active Finger Motion: Compliance, Force, and Work." Submitted to *Percept. Psychophys.*

Tan, H.Z., N.I. Durlach, Y. Shao, and M. Wei. "Manual Resolution of Compliance when Work and Force Cues are Minimized." *Proceedings of the American Society of Mechanical Engineers WAM'93*, Book H00851: 99-104, New Orleans, Louisiana, November 28-December 3, 1993.

Thesis

Bratakos, M.S. *Supplements to Speechreading: A Comparison of Auditory and Tactile Presentation of a Single-Band Envelope Cue.* S.B. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1993.

1.5 Multimicrophone Hearing Aids

Sponsor

National Institutes of Health
Grant R01-DC00270

Project Staff

Joseph G. Desloge, Nathaniel I. Durlach, Julie E. Greenberg, Michael P. O'Connell, Dr. William M. Rabinowitz, Daniel P. Welker, Dr. Patrick M. Zurek

The long-term goal of this research is the development of sensory aids that improve the ability of hearing-impaired listeners to function in complex acoustic environments through the use of microphone arrays. Since the reception of speech is the most important problem for the hearing impaired, the target signal of primary interest is speech.

We envision a microphone array system that resolves incoming signals into simultaneous directional channels followed by a coding operation that transforms these resolved signals so that resolution is preserved at the perceptual level after the signals are summed for presentation either to one or two ears. Such a system would permit even a monaural listener to monitor all directions simultaneously, detect and localize in the same operation, and focus on a single direction.

Our current work on microphone arrays is directed toward the extraction of a single target signal (assumed to arise from a target source straight ahead of the listener) and the reduction of interference from sources directionally distinct from the target source. Approaches based both on fixed and adaptive arrays, with processing for monaural or binaural outputs, are being studied.

Microphone arrays with fixed (time-invariant) weights are directed at enhancing a desired signal from one direction (straight ahead) while attenuating spatially distributed interference and reverberation. Past work[48] applied the theory of sensitivity-constrained optimal beamforming[49] to free-field head-sized arrays. This analysis showed that broadband intelligibility-weighted directivities[50] of 7-8 dB can be obtained either with a broadside array of cardioid microphones or with an endfire array with weights based on analog gains and pure time delays. Current work is directed at implementing portable prototype multimicrophone arrays to be worn and evaluated by hearing-aid users.

Work on adaptive array hearing aids has recently been focused on algorithmic analysis and design. Issues of adaptation control to minimize target cancellation and misadjustment, and robustness in reverberation, issues that were outlined previously,[51] are being thoroughly analyzed. The result of this analysis will be a specification for a relatively simple and robust five-microphone broadside adaptive array.

1.6 Superauditory Localization for Improved Human-Machine Interfaces

Sponsor

U.S. Air Force - Office of Scientific Research
Contract AFOSR-90-0200

Project Staff

Nathaniel I. Durlach, Professor Richard M. Held, Dr. William M. Rabinowitz, Barbara G. Shinn-Cunningham, Eric M. Foxlin

The motivation and background for this project were described in pages 312-313 of the *RLE Progress Report Number 134* and pages 306-307 of the *RLE Progress Report Number 135.*

In the past year, work on facilities development included the development of specifications for an improved auditory spatial synthesis device and further development of a new inertial position tracker with enhanced resolution and speed.

In examining ways in which supernormal localization cues could be produced, the idea of synthesizing HRTFs from larger-than-normal heads by frequency-scaling normal HRTFs was considered. The theoretical effects of frequency-scaling HRTFs to approximate a larger than normal head were investigated and reported in Rabinowitz, Maxwell, Shao, and Wei (1993). In this work, it was shown that frequency-scaling normal HRTFs will produce results very similar to HRTFs from larger than normal heads, provided the sources to be simulated are relatively far from the listener.

Earlier experiments showed that some adaptation could occur when active sensorimotor tasks were employed to train subjects. These basic results were investigated in more detail in the past year by varying the (1) type of training given to subjects, (2) number of positions employed, (3) complexity of the acoustic environment, and (3) presence of visual cues.

In all experiments performed to date, introduction of the transformation produced the anticipated changes in resolution and increased resolution in

48 R.W. Stadler and W.M. Rabinowitz, "On the Potential of Fixed Arrays for Hearing Aids," *J. Acoust. Soc. Am.* 94: 1332-1342 (1992).

49 H. Cox, R.M. Zeskind, and T. Kooij, "Practical Supergain," *IEEE Trans. Acoust. Speech Sig. Proc.* ASSP-34: 393-398 (1986).

50 J.E. Greenberg, P.M. Peterson, and P.M. Zurek, "Inelligibility-Weighted Measures of Speech-to-Interference Ratio and Speech System Gain," *J. Acoust. Soc. Am.* 94: 3009-3010 (1993).

51 J.E. Greenberg and P.M. Zurek, "Evaluation of an Adaptive Beamforming Method for Hearing Aids," *J. Acoust. Soc. Am.* 91: 1662-1676 (1992).

the center of the field. Furthermore, most of the experiments showed clear evidence of adaptation. Not only did the subjects in these experiments show a reduction in bias (and localization error) with exposure to the altered cues, but they also showed an increase in bias (and localization error) in the opposite direction when tested with normal cues following altered-cue exposure (a negative after-effect). Of particular interest in the context of classical adaptation work, adaptation was found to occur, with essentially comparable strength, without involvement of the sensorimotor system in the adaptation process: results were nearly identical when feedback was purely cognitive. This result contrasts with previous work (e.g., by Held, Mikaelian et al., Freedman et al., and Kalil) which focused strongly on the importance of sensorimotor involvement.

Independent of the issues of sensorimotor involvement, two results complicate the simplified picture described above. First, no adaptation occurred in the experiment in which the subjects were blindfolded. Second, in over half the experiments there appeared to be a tendency for enhanced resolution to decrease with increased exposure to the altered cues.

In the future, we will focus on analyzing the apparent decrease in resolution with exposure, examining the number of necessary and sufficient visual cues to achieve adaptation, investigating how the strength of the remapping function affects the speed and extent of adaptation, and developing a statistical model that describes experimental results.

Additional work connected with this grant has involved participation in meetings at government agencies (e.g., NASA and ONR) and the Acoustical Society of America.

An invention disclosure has been submitted to MIT's Office of Technology Licensing for the work on an inertial tracking system. A patent may ensue for this tracker, which was supported both by this project and NASA contract NCC 2-771.

1.6.1 Publications

Durlach, N.I., B.G. Shinn-Cunningham, and R.M. Held. "Supernormal Auditory Localization. I. General Background." *Presence* 2(2): 1-103 (1993).

Rabinowitz, W.M., J.A. Maxwell, Y. Shao, and M. Wei. "Sound Localization Cues for a Magnified Head: Implications from Sound Diffraction about a Rigid Sphere." *Presence* 2(2): 125-129 (1993).

Talks

Shinn-Cunningham, B.G. "Auditory Virtual Environments." Presented at the MIT Workshop on Space Life Sciences and Virtual Reality, Endicott House, Dedham, Masschusetts, January 6, 1993.

Shinn-Cunningham, B.G., N.I. Durlach, and R. Held. "Super Auditory Localization for Improved Human-machine Interface." Paper presented at the Air Force Office of Scientific Research Review of Research in Hearing, Fairborn, Ohio, June 1993.

Shinn-Cunningham, B.G., N.I. Durlach, and R. Held. "Auditory Display and Localization." Paper presented at the Conference on Binaural and Spatial Hearing, sponsored by the Air Force Office of Scientific Research and Armstrong Laboratory, WPAFB, Ohio, September 9-12, 1993.

1.7 Mechanistic Modeling of the Primate Fingerpad

Sponsor

National Institutes of Health
Grant R29-DC00625

Project Staff

Dr. Mandayam A. Srinivasan, Kiran B. Dandekar, Walter E. Babiec

When we touch an object, the source of all tactile information is the spatio-temporal distribution of mechanical loads on the skin at the contact interface. The relationship between these loads and the resulting stresses and strains at the nerve terminals within the skin plays a fundamental role in the neural coding of tactile information. Although empirical determination of the stress or strain state of a mechanoreceptor is not possible at present, mechanistic models of the skin and subcutaneous tissues enable us to generate testable hypotheses on skin deformations and associated peripheral neural responses. Verification of the hypotheses can then be accomplished by comparing the calculated results from the models with biomechanical data on the deformation of both the skin and subcutaneous tissues and with neurophysiological data from recordings of the responses of single neural fibers. The research under this grant is directed towards applying analytical and computational mechanics to analyze the biomechanical aspects of touch—the mechanics of contact, the transmission of the mechanical signals through the skin, and their transduction into neural impulses by the mechanoreceptors.

1.7.1 Determination of Geometric and Material Properties of the Primate Fingertip

The first step in performing mechanistic analyses of the primate fingertip (distal phalanx) is to determine its geometric and material properties. Three-dimensional (3D) external geometry of the primate fingertip was determined from accurate epoxy replicas of human and monkey fingertips. Using a videomicroscopy setup, we obtained images of orthographic projections of the epoxy replicas at various known orientations. The images were then digitized and processed to determine the boundary of the finger at each orientation. By combining the boundary data for all the different orientations, we were able to reconstruct the 3D external geometry of the fingertip.[52] We have reconstructed several human and monkey fingertips using this method.

For mechanical modeling of the human fingerpad, the Poisson's ratio, which is a measure of its compressibility, is required as an input to the mathematical models. The Poisson's ratio for the human fingerpad in vivo is unknown at present. In previous noninvasive experiments on human subjects, we have measured the change in volume of the fingerpad under static indentations with different indentors.[53] Our results show that the compressibility of the fingertip increases with increases in both the depth of indentation and the contact area with the indentor. The highest change in fingertip volume was about 5 percent. We have now developed an experimental setup involving a computer controlled linear actuator for fingertip volume change measurements under dynamic conditions.

1.7.2 Fingertip Models and Finite Element Analyses

We have performed linear and nonlinear finite element analyses of a series of mechanistic models of the fingerpad under a variety of mechanical stimuli.[54] The models range from a semi-infinite medium to a 3D model based on the actual finger geometry and composed of either a homogeneous elastic material or a thick elastic shell containing a fluid. Simulations of the mechanistic aspects of neurophysiological experiments involving mapping of receptive fields with single point loads, determination of spatial resolution of two-point stimuli, and indentations by single bars as well as periodic and aperiodic gratings have been carried out for the 2D models. We have also solved the nonlinear contact problem of indentations by cylindrical objects. The large number of numerical calculations needed even for the linear 2D models necessitated the use of the MIT supercomputer.

The results show that the (1) model geometry has a significant influence on the spatial distribution of the mechanical signals and (2) elastic medium acts like a low-pass filter in causing blurring of the mechanical signals imposed at the surface. None of the other models predicted the experimentally observed surface deflection profiles under line loads as closely as a simple heterogeneous model that treated the fingerpad as a membrane enclosing an incompressible fluid.[55] The results from the analyses of these models were then related to recordings of the responses of slowly and rapidly adapting mechanoreceptors (SA and RA, respectively). Under indentations by rectangular gratings, the maximum compressive strain and strain energy density at the receptor locations emerged as the two strain measures that are directly related to the response rate of SAs. The same result was true for the contact problem of indentations by cylinders. However, strain energy density is a better candidate since it is a scalar that is invariant with respect to receptor orientation and a direct measure of the distortion of the receptor caused by loads imposed on the skin. For the model calculations to match both biomechanical and neurophysiological data more

---

[52] T.R.R. Perez and M.A. Srinivasan, *Videomicroscopic Reconstruction of the Human Finger,* Project report to the MIT Minority Summer Science Research Program, 1992.

[53] M.A. Srinivasan, R.J. Gulati, and K. Dandekar, "In vivo Compressibility of the Human Fingertip," *ASME Adv. Bioeng. 22: 573-576 (1992).*

[54] M.A. Srinivasan and K. Dandekar. "Role of Fingertip Geometry in the Transmission of Tactile Mechanical Signals," *ASME Adv. Bioeng.* 22: 569-572 (1992); M.A. Srinivasan and K. Dandekar, "An Investigation of the Mechanics of Tactile Sense Using Two-dimensional Models of the Primate Fingertip," submitted to *J. Biomech. Eng.*

[55] M.A. Srinivasan, "Surface Deflection of Primate Fingertip Under Line Load," *J. Biomech.* 22(4): 343-349 (1989).

accurately, it is likely that the models will have to take into account the inhomogeneity of the primate fingerpad.

1.8 Peripheral Neural Mechanisms of Haptic Touch

Sponsor

U.S. Navy - Office of Naval Research
Grant N00014-88-K-0604

Project Staff

Dr. Mandayam A. Srinivasan, Dr. Anuradha M. Annaswamy, Dr. Robert H. LaMotte,[56] Steingrimur P. Karason

We have been collaborating with Dr. LaMotte of Yale University School of Medicine in conducting psychophysical and neurophysiological studies on the tactile perception of the microtexture, shape and softness of objects. We have shown that humans can detect extremly fine textures composed of 50 nanometers high parallel bars on plane glass plates.[57] Our neurophysiological recordings indicate that when such fine textures are stroked on the fingerpad skin, the fingerprint ridges vibrate and cause Pacinian Corpuscles to respond, thus enabling the detection of the microtexure.[58] In studies of the tactile perception of shape, a series of two- and three-dimensional objects (e.g., cylinders, spheres, ellipsoids and wavy surfaces) were pressed or stroked across the fingerpads of anesthetized monkeys and the evoked responses in cutaneous mechanoreceptive primary afferent nerve fibers were recorded.[59] Major geometrical properties of the shapes were well represented in the spatio-temporal responses of SA and RA afferent fiber populations, particularly those of the SAs. The results show that the following hypothesis explains qualitatively all the data we have obtained: the depth of indentation and the change in curvature of the skin surface are encoded by the discharge rates of SAs. In addition, the velocity and the rate of change in skin surface curvature are encoded by the discharge rates of both SAs and RAs.

Thus, intensive parameters of shapes, such as the magnitude of change in skin curvature produced by contact with the object surface, were encoded in the discharge rates of SAs and RAs, but this neural code was also influenced by changes in stroke velocity. Spatial parameters of shapes such as the curvature width and the changes in contour that characterize a shape as belonging to a particular category (such as a sphere as opposed to a cylinder) are encoded in the spatially distributed discharge rates of the SA population. This spatial response profile provides a neural code that is probably invariant with moderate changes in the parameters that govern contact conditions between the object and the skin, such as the contact force or orientation and velocity of its trajectory. Therefore, among the different possible geometric representations of the shape of objects, the intrinsic description, i.e., the surface curvature as a function of the distance along the surface, seems to be relevant for tactile sensing of shape.

Based on a theoretical analysis of the mechanics of contact, we have proposed a mechanism by which shapes of objects within contact regions are perceived through the tactile sense. The curvature of the skin surface under an object, which we know from differential geometry is approximated by the second spatial derivative of surface deflection, is coded not by differentiating (which is a noise enhancing process), but by exploiting its relation to surface pressure. Pressure peaks occur where the depths of indentation and/or changes in the skin surface curvature are high. The skin effectively acts as a low-pass filter in transmitting the mechanical signals, and the mechanoreceptors respond to the blurred versions of the surface pressure dis-

---

56 Yale University School of Medicine, New Haven, Connecticut.

57 R.H. LaMotte and M.A. Srinivasan, "Surface Microgeometry: Neural Encoding and Perception," in *Information Processing in the Somatosensory System,* eds. O. Franzen and J. Westman, Wenner-Gren International Symposium Series (New York: Macmillan Press, 1991).

58 M.A. Srinivasan, J.M. Whitehouse, and R.H. LaMotte, "Tactile Detection of Slip: Surface Microgeometry and Peripheral Neural Codes," *J. Neurophys.* 63(6): 1323-1332 (1990).

59 M.A. Srinivasan and R.H. LaMotte, "Encoding of Shape in the Responses of Cutaneous Mechanoreceptors," in *Information Processing in the Somatosensory System,* eds. O. Franzen and J. Westman, Wenner-Gren International Symposium Series (New York: Macmillan Press, 1991); R.H. LaMotte and M.A. Srinivasan, "Responses of Cutaneous Mechanoreceptors to the Shape of Objects Applied to the Primate Fingerpad," *Acta Psychol.* 84: 41-51 (1993); R.H. LaMotte, M.A. Srinivasan, C. Lu, and A. Klusch-Petersen, "Responses of Cutaneous Mechanoreceptors to Two- and Three-dimensional Shapes Stroked Across the Monkey Fingerpad. *Soc. Neurosci. Abstr.* 19: 105 (1993); LaMotte, R.H., M.A. Srinivasan, C. Lu, and A. Klusch-Petersen. "Cutaneous Neural Codes for Shape." *Can. J. Physiol. Pharm.* Forthcoming.

tribution, thus encoding the shape of the object in terms of its surface curvatures.[60]

We have also shown that the human discriminability of softness or compliance of objects depends on whether the object has a deformable or rigid surface.[61] When the surface is deformable, the spatial pressure distribution within the contact region is dependent on object compliance, and hence information from cutaneous mechanoreceptors is sufficient for discrimination of subtle differences in compliance. When the surface is rigid, kinesthetic information is necessary for discrimination, and the discriminability is much poorer than that for objects with deformable surfaces.

1.8.1 Development of a Computational Theory of Haptics

Our research on computational theory of haptics is focused on developing a theoretical framework for studying the information processing and control strategies common to both humans and robots performing haptic tasks. For example, although the "hardware" of the tactile apparatus in humans and robots are different, they have the common feature of mechanosensors embedded in a deformable medium. Thus the computational problem of coding (predicting sensor response for a given mechanical stimulus at the surface) and decoding (inferring the mechanical stimulus at the surface by suitably processing the sensor response) need similar mechanistic analyses for their solution. We have developed such a "computational theory" for an idealized medium subjected to arbitrary pressure or displacement loading conditions, and give explicit formulae for the coding and decoding problems.[62]

In our work during the past year, we were successful in answering basic identification and control questions that arise during manipulation of compliant objects using compliant fingerpads.[63] To understand the fundamental aspects of these tasks, we have analyzed the problem of identification of compliant objects with a single finger contact, as well as under a two-finger grasp. Using lumped parameter models, we have carried out the identification of human and object parameters, using either force or displacement inputs to the rigid backing of the end-effector. Based on identified parameters, control strategies are developed to achieve a desired manipulation of the object in the workspace.

We have also modeled the dynamic interactions that occur between compliant end-effectors and deformable objects by a class of nonlinear systems. It was shown that standard geometric techniques for exact feedback linearization were inadequate. New algorithms were developed by using adaptive feedback techniques which judiciously employed the stability characteristics of the underlying nonlinear dynamics. Both theoretical and simulation results show that these adaptive control algorithms led to successful manipulation. The theoretical results can be used to generate testable hypotheses for experiments on human or robot haptics.

1.9 Biomechanics of Skin-Object Contact

Sponsor

U.S. Navy - Office of Naval Research
Grant N00014-91-J-1454

Project Staff

Dr. Mandayam A. Srinivasan, Jyh-Shing Chen, Rogeve J. Gulati

Although physical contact is ubiquitous in our interactions with objects in the environment, we do not yet understand the mechanistic phenomena occurring at the skin-object interface. As mentioned before, the spatio-temporal distribution of mechanical loads on the skin at the contact interface is the source of all tactile information. These loads, specified as pressure, displacements, etc., depend on the geometrical and material properties of both the contacting entities, as well as the overall forces of interaction. The goal of this project is to determine the growth and motion of contact regions and vari-

60 M.A. Srinivasan and R.H. LaMotte, "Encoding of Shape in the Responses of Cutaneous Mechanoreceptors," in *Information Processing in the Somatosensory System,* eds. O. Franzen and J. Westman, Wenner-Gren International Symposium Series (New York: Macmillan Press, 1991).

61 M.A. Srinivasan and R.H. LaMotte, "Tactual Discrimination of Softness," submitted to *J. Neurophys.*

62 M.A. Srinivasan, *Tactile Sensing in Humans and Robots: Computational Theory and Algorithms,* Newman Laboratory Technical Report, Dept. of Mech. Eng., MIT, 1988.

63 A.M. Annaswamy and M.A. Srinivasan, "The Role of Compliant Fingerpads in Grasping and Manipulation: Identification and Control," Institute of Mathematics (New York: Springer Verlag, 1993).

ations in the associated force over time between the human fingerpad and carefully chosen transparent test objects whose microtexture, shape or softness is varied in a controlled manner. These results are being used to gain a deeper understanding of the data we have already obtained for the same test objects, namely, electrophysiologically recorded responses of cutaneous mechanoreceptive afferent fibers, and psychophysically determined human discriminabilities.

To measure the *in vivo* surface deformations of the fingerpad under various tactile stimuli, we have designed a videomicroscopy system together with high precision force sensors. The videomicroscopy system consists of a set of video zoom lenses attached to a high-resolution CCD camera, whose output can either be digitized directly at about 5 frames/s, or stored on a laserdisk at real-time frame rates (30 frames/s) for off-line digitization. The zoom lenses enable continuous variation of magnification, with the field of view covering the entire fingerpad, or just a few fingerprint ridges. High contrast images are achieved with coaxial and other fiberoptic lighting. In collaboration with our colleagues at the MIT Artificial Intelligence Laboratory, we designed and built two 6-axis force sensors that are customized to our application. These sensors have much higher resolutions (10 bit) than commercial sensors operating in a comparable range of forces (5 Newtons). Transparent test objects can be attached to these sensors for both biomechanical and psychophysical experiments.

Using the test facility described above, we have performed a set of experiments with human subjects to investigate the relationship between the contact force, contact area and compliance of the object. The experiments involved active indentation of transparent compliant rubber specimens and a glass plate with the subjects' fingerpads. Static video images of the contact regions were captured at various force levels and magnifications. In order to minimize the effects of nonuniform illumination, we implemented homomorphic image processing algorithms with or without image decimation. The processed images showed that contact regions consisted of discontinuous "islands" along each finger ridge, with clear distinction between contact and noncontact regions over the entire field of view.

Results show that for objects whose compliances are discriminable, even when the overall contact areas under a given contact force are the same, the actual contact areas can differ by a factor of two or more. The actual pressure distribution, which acts only within the discontinuous contact islands on the skin, will therefore be radically different for the objects. Consequently, a spatio-temporal neural code for object compliance emerges with far higher resolution than an intensive code such as the average pressure over the overall contact area. These results are in agreement with our hypothesis that the neural coding of objects with deformable surfaces (such as rubber) is based on the spatio-temporal pressure distribution on the skin. This was one of the conclusions from our psychophysical, biomechanical and neurophysiological experiments in a companion project conducted in collaboration with Dr. LaMotte of Yale University School of Medicine.

In another set of experiments, the images of a human fingertip before and after indentation with rectangular and cylindrical bars were obtained. By superposing the images, we can measure the surface profile as well as the displacements of the points on the skin surface marked with ink. Such measurements for several bars at different orientations serve as biomechanical data to verify and guide the finite element models of the fingertips described above.

1.10 Human and Robot Hands: Mechanics, Sensorimotor Functions and Cognition

Sponsor

U.S. Navy - Office of Naval Research
Grant N00014-92-J-1814

Project Staff

Dr. Mandayam A. Srinivasan, Dr. J. Kenneth Salisbury, Nathaniel I. Durlach, Dr. Robert D. Howe,[64] Dr. Robert H. LaMotte, Kiran B. Dandekar, Louise Jandura, Kevin C. Knoedler

The premise of this University Research Initiative project is that the integrated study of human and robot haptics can provide complementary knowledge of the processes of prehension and manipulation. From the human research side, we wish to understand the basic mechanical, perceptual, and strategic capabilities that lead to the dexterity and deftness we observe in human task performance. By studying the underlying competences that

[64] Harvard University, Cambridge, Massachusetts.

humans bring to bear on task performance, we seek guidelines on how to better build robots. From the robotic research side, we wish to understand how mechanism and sensor design choices can best be made to maximize grasping and manipulative competences in robot task performance. By understanding better the mechanical demands of task performance, we seek to understand the performance demands which underlie skillful human manipulation. To bridge these two areas of study, we have assembled a unique team of researchers with diverse backgrounds in robotic and human studies at MIT, Harvard and Yale Universities.

The main components of the research conducted under this project during the past year are (1) development of new hardware for robotic and human studies, (2) processing of robot sensor signals and task-level control of the devices, and (3) experiments on human perception and control using some of the devices. The subsections to follow provide description of the results obtained in each of the three topics.

1.10.1 Developments of New Hardware

Hardware developments by the Dextrous Manipulation Group at the Artificial Intelligence (AI) Laboratory at MIT include two versions of glass-smooth actuator test-beds, the first version of the Precision Manipulation Instrument, and PVDF-2 contact sensors. Collaboration between the AI Lab Group and the Human Haptics Group at MIT has resulted in the design and fabrication of miniature, high-resolution six-axes force sensors, and an Instrumented ScrewDriver for experiments on human force control. The Harvard Group and the MIT Human Haptics Group have collaborated in building a reconfigured robot finger to be used as a tactile stimulator in experiments on human subjects. The Harvard Group has fabricated a prototype of a flexible tactile array pressure distribution sensor and an integrated sensor system. They have also begun development of a multi-element stress rate sensor for improved slip and contact detection along with new signal processing electronics for improved rejection of noise originating in the manipulator structure. The Yale group has designed and fabricated an instrumented probe that is capable of measuring dynamic grasp forces. The probe can be mounted on a torque motor for recording and playing back the transient force signals of impact with compliant objects in psychophysical experiments under both active and passive touch.

1.10.2 Robot Signal Processing and Task-Level Control

Interpretation of Contact Force Signals from a Robot Fingertip

Through theoretical advances in processing of temporal signals from a robot fingertip force sensor, the MIT AI Lab group has successfully applied the generalized sequential likelihood ratio test to temporal contact force signals and used it to segment them into statistically equivalent pieces indicating significant contact events. This permits detection and labeling of impacts, slip, changes in texture, and contact condition. While of immediate value in detecting and differentiating contact events, the results of the segmentation are designed to be used in a higher level procedure which will interpret the results within a manipulation context. All the algorithms are based on forms of the the generalized sequential likelihood ratio test. This test is optimal under mild conditions. Our algorithms yield significantly better performance on standard tasks, like sensing contact, than current thresholding techniques. In addition, our new algorithms allow continuous self-calibrating (adaptive) sensing and the detection/segmentation of more complex phenomena.

Task Level Control System

A task level control system has been developed for the Salisbury Hand and applied to the problem of grasping. The system is designed to grasp objects in a semi-structured environment (static objects which may be displaced randomly). The hand is outfitted with force sensing fingertips and a simple contact sensing palm. The system uses sparse contact information (spatially and in time) obtained from these sensors and coarse geometric models of objects to recursively refine its "knowledge" about object position and grasp state. Each contact event provides partial information about the task state and is used to increase the knowledge about it. Aging or inconsistent data is rejected, thus permitting a degree of robustness to error and unexpected disturbances. Local motions aimed at improving the task knowledge (and ultimately, grasp quality) are planned "on the fly" based on the current knowledge state. The system thus incorporates the responsiveness of sensor-based reactive systems and the correctness of planning systems. Proofs of stability and convergence have been obtained and the system has been demonstrated in the grasping of cylinders and blocks.

Regrasping Gaits

An investigation of regrasping "gaits" has been conducted. This work is aimed at both developing methods for robot hands to continuously reorient objects by repositioning fingers and at revealing the underlying performance of this element of skillful manipulation. A systematic approach for enumerating acceptable regrasping sequences has been developed, and this permits us to identify those patterns of finger movement that allow continual reorientation of objects acknowledging friction limits, object weight, and object orientation.

1.10.3 Experiments on Human Perception and Control

Human Performance in Controlling Normal Forces of Contact

The Human Haptics Group at MIT has investigated the ability of humans in actively controlling contact forces with normal and locally anesthetized fingertips.[65] The effect of force magnitude, time-profiles of forces, and visual feedback has been studied. The experiments consisted of tracking constant, linear-ramp and sinusoidal time-profiles of forces applied on a custom-built six-axis force sensor. In most of the experiments, visual feedback was provided by displaying on a computer monitor the time-profiles of both the desired target forces and actual forces exerted by the subjects. Some experiments on maintaining constant force were conducted without visual feedback.

The results show that for a constant target force, the mean error in tracking remained approximately constant with respect to force magnitude when visual feedback was present, but increased with target force magnitude in the absence of vision. The lack of tactile feedback due to local anesthesia caused the error to increase by 25 to 80 percent. Higher force rates also generally increased the percentage of error. Among the three target profiles, the constant force had the least error. For a given force rate, the errors for the sinusoidal and ramp targets were about the same. A design specification that emerges from these results is that haptic interfaces for virtual environments and teleoperation must have a force resolution of about 0.01 Newton at low force rates to make full use of human haptic capabilities.

Human Performance in Controling Torque

The human ability to control torque has been investigated in preliminary experiments with the Instrumented ScrewDriver (ISD). The ISD, comprised of a single shaft which is supported by low friction bearings, is connected to a reaction torque sensor and a magnetic particle brake. Angular position of the shaft is measured by an incremental optical encoder. In this set of experiments, the subjects were asked to grasp the handle between the thumb and index finger and turn the shaft clockwise at a constant velocity. At a certain point during this maneuver, the resistive torque acting on the shaft was increased in the form of a sharp step of variable magnitude. When the step change in resistive torque occurred, the initial motion ceased while the subject increased the applied torque. After this time delay, motion of the shaft resumed and continued until the experiment was over. The results show that the time delay increased approximately linearly with resistive torque, indicating that the subjects increased the applied torque at an almost constant rate until shaft motion resumed. Increasing standard deviations about the mean applied torque after the delay indicated that increases in the resistive torque level cause decreases in the human ability to control torque.

1.10.4 Processing of Tactile Information Obtained by Touching an Object with a Tool

The Yale group has recorded and analyzed tactile signals delivered to a single fingerpad in contact with a tool (a stylus) that strikes objects of differing compliance. One end of the stylus, attached to a servocontrolled torque motor, rested against the restrained fingerpad of the subject and exerted a base force of about 0.2 Newton. The other end of the stylus was struck with a rubber object whose compliance was varied. The force exerted by the stylus against the skin upon impact with the object was recorded. In a series of experiments, this signal was edited slightly and "played back" under servocontrol of force by the motor through the stylus. The subject judged the magnitude of subjective softness by assigning a number (the method of Magnitude Estimation). On certain trials, the back end of the stylus was struck with a rubber object by the experimenter while on other trials the

---

65 M.A. Srinivasan and J.S. Chen, "Human Performance in Controlling Normal Forces of Contact with Rigid Objects," *Advances in Robotics, Mechatronics, and Haptic Interfaces*, DSC-Vol. 49, ASME, Winter Annual Meeting, New Orleans, 1993.

motor played back a signal that had been previously recorded. In addition to judging the magnitude of softness, the subjects made a two-alternative forced choice as to whether the tactile signal was "live" or "played back." Results demonstrated that the subjects were unable to make this discrimination above the level of chance performance. This meant that the tactile signals played back by the motor felt natural and represented those generated when a stylus is struck by a compliant object.

A series of experiments in which the recorded traces were filtered or otherwise edited demonstrated that the critical signal associated with the compliance of an object was the slope of the rising phase of the force trace. This feature was more important than others such as peak amplitude or higher frequency oscillations in the signal. In the future, these signals will then be delivered to the cutaneous receptive fields of mechanoreceptors to learn how tool-generated tactile information is transduced and how associated neuronal events in sensory pathways are processed by the brain during haptic performance.

1.10.5 Publications

Annaswamy, A.M., and D. Seto. "Object Manipulation Using Compliant Fingerpads: Modeling and Control." *ASME J. Dynamic Syst. Measure. Contr.* 115: 638-648 (1993).

Eberman, B., and J.K. Salisbury. *Application of Change Detection to Dynamic Contact Sensing.* MIT Artificial Intelligence Laboratory Memo 1421, MIT, 1993.

Eberman, B., and J.K. Salisbury. "Segmentation and Interpretation of Temporal Contact Signals," Third ISER Conference, Kyoto, Japan, October 1993.

Eberman, B., and J.K. Salisbury. "Application of Change Detection to Dynamic Contact Sensing." Submitted to *Int. J. Robot. Res.*

LaMotte, R.H., and M.A. Srinivasan. "Responses of Cutaneous Mechanoreceptors to the Shape of Objects Applied to the Primate Fingerpad." *Acta Psychol.* 84: 41-51 (1993).

LaMotte, R.H., M.A. Srinivasan, C. Lu, and A. Klusch-Petersen. "Responses of Cutaneous Mechanoreceptors to Two- and Three-dimensional Shapes Stroked Across the Monkey Fingerpad. *Soc. Neurosci. Abstr.* 19: 105 (1993).

Pollard, N. "Planning Grasps for a Robot Hand in the Presence of Obstacles." *Proceedings of the IEEE International Conference on Robotics and Automation,* Atlanta, Georgia, May 1993, vol. 3, pp. 723-728.

Srinivasan, M.A., and K. Dandekar. "An Investigation of the Mechanics of Tactile Sense Using Two-dimensional Models of the Primate Fingertip." Submitted to *J. Biomech. Eng.*

Srinivasan, M.A., and R.H. LaMotte. "Tactual Discrimination of Softness." Submitted to *J. Neurophys.*

Srinivasan, M.A., and J.S. Chen. "Human Performance in Controlling Normal Forces of Contact with Rigid Objects, *Advances in Robotics, Mechatronics, and Haptic Interfaces,* DSC-Vol. 49, ASME Annual Winter Meeting, New Orleans, 1993.

Srinivasan, M.A. "Virtual Haptic Environments: Issues in Force and Tactile Feedback." Abstract in *Proceedings of Virtual Reality Systems '93.* New York: Sig-advanced Applications, 1993.

Papers Accepted for Publication

Annaswamy, A.M., and M.A. Srinivasan. "The Role of Compliant Fingerpads in Grasping and Manipulation: Identification and Control." In *Mathematics and Its Applications.* Institute of Mathematics series. New York: Springer Verlag, 1994.

Howe, R.D. "Tactile Sensing and Control of Robotic Manipulation." *J. Adv. Robot.* Forthcoming.

LaMotte, R.H., M.A. Srinivasan, C. Lu, and A. Klusch-Petersen. "Cutaneous Neural Codes for Shape." *Can. J. Physiol. Pharm.* Forthcoming.

Seto, D., A.M. Annaswamy, and J. Baillieul. "Adaptive Control of a Class of Nonlinear Systems with a Triangular Structure." *IEEE Trans. Auto. Control.* Forthcoming.

Son, J., A. Monteverde, and R.D. Howe. "A Tactile Sensor for Localizing Transient Events in Manipulation." Submitted to the IEEE International Conference of Robotics and Automation. Forthcoming.

Theses

Brock, D.L. *A Sensor Based Strategy for Automatic Robotic Grasping.* Ph.D. diss., Dept. of Mech. Eng., MIT, January 1993.

Massie, T.H. *Design of a Three Degree of Freedom Force Reflecting Haptic Interface.* S.B. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1993.

1.11 Virtual Environment Technology for Training (VETT)

Sponsor

U.S. Navy - Naval Training Systems Center
Contract N61339-93-M-1213
Contract N61339-93-C-0055
Contract N61339-93-C-0083

The VETT program consists of five related research projects that include a virtual environment (VE) system that serves as a research testbed and four satellite projects. The testbed and two of the satellites, sensorimotor adaptation and improved haptics interfaces, are underway at RLE. Research on the whole body motion satellite is being conducted at the Graybiel Laboratory of Brandeis University, and the training satellite research is being conducted at Bolt, Beranek and Newman, Inc., of Cambridge. This report will cover the VETT testbed, the sensorimotor satellite, and the haptics interface satellite.

1.11.1 VETT Testbed

Project Staff

Dr. David Zeltzer, Dr. Jeng-Feng Lee, Walter A. Aviles, Rakesh Gupta, Erik Nygren, Nicholas Pioch

The first two phases of the first year of the VETT Testbed effort have been completed as scheduled. All major computational and human/machine interface systems have been procured, installed, and tested. Computational equipment includes an SGI 2-processor ONYX system with VTX graphics subsystem, two SGI Indigo2 Extremes (one procured under another grant), a 486 PC, and a Macintosh Centris 660AV, all installed in a networked environment. Interface devices include a color video projection system capable of field-interlaced stereo display, a Virtual Research EYEGEN3 color helmet mounted display, a three-sensor Polhemus FASTRAK position/orientation sensing system, a VPL Dataglove, a PHANToM 3-axis desktop force-reflecting haptic interface device, and a BEACHTRON sound spatialization system from Crystal River Engineering.

A demonstration virtual environment—a three-dimensional model of the VETT Testbed laboratory—has been implemented and tested with all interface devices. Frame rates for the demonstration VE are approximately 40Hz for bi-ocular viewing (same image presented to each eye) and 20Hz for stereoscopic viewing.

Plans for the remainder of the first year include: (1) characterization of testbed performance (update rate and lag); (2) implementation of a prototype version of the currently envisioned first testbed-based training task (vehicle dynamics training); (3) configuration of the testbed (i.e., architecture, layout, and instrumentation) to support precisely controlled psychophysical experiments and device evaluations; (4) addition of speech-based control; and (5) incorporation of additional visual, auditory, and haptic interface devices.

1.11.2 Sensorimotor Satellite

Project Staff

Nathaniel I. Durlach, Professor Richard M. Held, Dr. Thomas E. Wiegand, Julien B. Beasley, Evan F. Wies

Since the start of the VETT project, we have assembled all the equipment required for the planned experiments. This equipment consists of a planar four-degree-of-freedom (DOF) manipulandum and a visual display mounted in a viewing box. The display is mounted in such a way that the visual image representing the state of the manipulandum (i.e., the subject's hand that is grasping the manipulandum) appears in the plane of the manipulandum. The relation between the state of the manipulandum and the visual image representing the manipulandum are under the control of a Macintosh computer and can be programmed to simulate various types of distortions, delays, and stochastic disturbances (noises) that are likely to arise in virtual-environment training systems.

We have built two separate jointed-arm manipulanda for use in the proposed sensorimotor experiments. Each device has four DOFs (x, y, twist, and squeeze) in a planar operating range of about 10 by 14 inches. Manipulandum number 1 also includes a brake on the squeeze dimension to simulate the presence of an object between the grip pads. Manipulandum number 2 was constructed of thin-walled aluminum tubing (without the brake) to minimize the mass of the moving elements. Which manipulandum will be used will depend on the detailed requirements of the specific experiment under consideration.

During the last half of the first year, we expect to complete a set of preliminary experiments in which we introduce spatial, temporal, and stochastic alterations corresponding to the limitations of available interface devices. These data will enable us to make informed choices of technical tradeoffs and will also allow us to begin developing a model for predicting the effects of such alterations.

A further component of the second half of the year's research effort concerns the investigation of sensorimotor performance involving spatial auditory targets. This work will include installation of enhanced spatial tracking hardware and planning of experiments to be conducted with the new auditory display equipment being developed at Crystal River Engineering (CRE) for delivery during the second year.

1.11.3 Haptics Satellite

Project Staff

Dr. J. Kenneth Salisbury, Dr. Mandayam A. Srinivasan, Dr. David L. Brock, Gerald L. Beauregard, Nick Swarup, Craig B. Zilles, Thomas H. Massie

During the first half of the first year, we have developed device hardware and interaction software and performed psychophysical experiments pertaining to haptic interactions with virtual environments. The Haptics laboratory space in RLE has been reorganized and outfitted with new test and computer equipment as originally proposed. Two major devices for performing psychophysical experiments, the Linear and Planar Graspers, have been extensively refurbished with improved mechanics and sensors. A new force-reflecting gripper to be mounted on the planar grasper has been constructed and is currently being interfaced. Another haptic display device developed previously, the PHANToM, has been used to prototype a wide range of force-based haptic display primitives. It permits precisely controllable forces to be exerted on a user's fingertip through a thimble connection and allows simple tool interactions through a stylus-like grip. Copies of this device, along with demonstration software, were delivered to Naval Air Warfare Center Training Systems Division (NAWCTSD) at the end of December 1993 and to the VETT Testbed in January 1994.

In preparation for the design of a tool-handle haptic interface suitable for a large workspace, a full-scale mockup has been fabricated to determine workspace constraints. An analysis of force/inertia/friction requirements of the device aimed at specifying kinematic, structural and actuator requirements has been initiated. In addition, an operational prototype of a powered screwdriver tool-handle has been built and interfaced to provide an early reality check on the tool-handle concept and feedback useful for design of the full scale version. At the MIT Newman Laboratory, final commissioning and testing of the instrumentation and control package was done for MANUS, a 5-DOF haptic interface device with workspace comparable to the human arm, and a data communication link was established between MANUS and MANTA, a 2-DOF haptic interface device.

We have developed software for the demonstration of haptic displays with the PHANToM, such as interactions with virtual pushbuttons, sliding friction and surface texture displays, virtual screw turning, and virtual grasping, as well as 2-person haptic interactions with two PHANToMs. Software has also been developed for performing new psychophysical experiments with the Linear Grasper. Following analysis of the data obtained in previous experiments on the compliance just noticeable difference (JND) and its relationship to the discrimination of mechanical work and terminal force, a new series of experiments for measuring human discriminability of compliance and work done have been initiated.

During the remaining portion of the first year, we plan to further characterize human performance in the discrimination of fundamental mechanical properties of objects such as compliance, viscosity and inertia. This requires additional hardware improvements to the linear grasper, such as the mounting of tachometers and accelerometers. Other hardware developments consist of the addition of the force reflecting gripper to the Planar Grasper to permit pick-and-place tasks, and the design and construction of tool-handle prototypes 1 and 2. We will provide support for the interfacing of the PHANToM into the Testbed and the performance of psychophysical experiments under multimodal conditions. Our haptic display software efforts will focus on two areas, the packaging of display algorithms for transfer to Testbed use, and the exploration of new display techniques for evoking more subtle contact events and states, such as surface microtexture, stiction, rigidity, and edge sharpness.

1.12 The Virtual Sailor: Synthetic Humans for Virtual Environments

Sponsor

U.S. Navy - Office of Naval Research
Grant N00014-92-J-4005

Project Staff

Dr. David Zeltzer, Svetlana Gaffron

1.12.1 Project Overview

Many virtual environment (VE) applications, such as a system for training team members to perform vehicle maintenance, call for the representation and display of interactive, realtime synthetic humans, which we call virtual actors. To do this, we need to develop adaptive and efficient computational models of human motor behavior.

In the past year, we have undertaken a research program designed to identify the basic issues involved in developing a system, which we call SkillBuilder, for computer-aided design of motor skills for virtual actors. The skill builder will allow us to implement a generic human model and train it to adaptively execute new skills, such as handling and operating a particular tool. As part of this effort, we have implemented a model of visually-guided reaching and grasping and controllers for human head and eye movement based on psychophysical studies. We have demonstrated that these models can support realistic, realtime interaction with human VE participants.[66]

Real-time, interactive behavior models would be of little use without appropriate human interface software that supports task level interaction using speech and gestures. In addition to our work on SkillBuilder, therefore, we have also developed an architecture that we believe will support task level interaction using a restricted natural language interface. This work is based on our previous work on reactive planning systems,[67] as well as the natural language research of Schank and colleagues in the early 1970s,[68] efforts to develop a language for specifying human behavior in a clinical setting,[69] and the notion of affordances first described by J.J. Gibson.[70]

Finally, in addition to modeling human motor behavior for virtual actors, it is important to correctly capture human appearance, especially when close-up interactions are required. Therefore, we have also continued our efforts to develop accurate, realtime finite element models (FEM) of human facial tissue. This year we have developed a set of motor programs that coordinate the actions of our finite element model of the human face, such that we can display a range of facial expressions. Since we cannot update the FEM in realtime, we have also developed a method for pre-computing a set of key expressions and later playing them back in realtime.

1.12.2 Real-Time, Multimodal Interaction With Virtual Actors

The long-range goal of this work is to implement a virtual environment that includes autonomous and interactive synthetic humans. These virtual actors should be capable of interacting with human participants in the VE, as well as with simulated objects and events in the VE."

On July 30, 1993, we were able to demonstrate a prototypical virtual actor capable of responding, in realtime, to physical interaction with human participants using position-sensing devices. In addition, this virtual actor also performed interactive, realtime, visually-guided grasping and reaching behaviors. Finally, the actor was capable of responsive head and eye movements. Not only could it

66 S. Gaffron, *SkillBuilder: A Motor Program Design Tool for Virtual Actors,* S.M. thesis, Dept. of Mech. Eng., MIT, 1994.

67 D. Zeltzer and M. Johnson, "Motor Planning: Specifying and Controlling the Behavior of Autonomous Animated Agents," *J. Vis. Comput. Animat.* 2(2): 74-80 (1991).

68 R.C. Schank, "Conceptual Information Processing," (New York: American Elsevier Publishing Company, 1975); R.C. Schank, and R.P. Abelson, "Scripts, Plans, Goals and Understanding," (Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977).

69 E.S. Reed, M. Montgomery, M. Schwartz, C. Palmer, and J.B. Pittenger, "Visually Based Descriptions of Everyday Action," *Ecological Psychology* 4(3): 129-152 (YEAR??); M.F. Schwartz, E.S. Reed, M. Montgomery, C. Palmer, and N.H. Mayer, "The Quantitative Description of Action Disorganization after Brain Damage: A Case Study," *Cog. Neuropsychol.* 8(5): 381-414 (1991).

70 J.J. Gibson, "The Theory of Affordances," in *Perceiving, Acting, and Knowing: Toward an Ecological Psychology,* eds. R. Shaw and J. Bransford (Hillsdale NJ: Lawrence Erlbaum Associates, 1977, pp. 67-82); W.M. Mace, "James J. Gibsons Strategy for Perceiving: Ask Not Whats Inside Your Head, but What Your Heads Inside Of," in *Perceiving, Acting, and Knowing: Toward an Ecological Psychology,* eds. R. Shaw and J. Bransford (Hillsdale NJ: Lawrence Erlbaum Associates, 1977, pp. 43-65).

track target objects with realistic head and eye movements, but if a human were to wear a head-mounted position-sensor, the actor would maintain eye contact with the human as he or she moved![71]

While there are a small number of human-modeling efforts underway elsewhere in the U.S. and Europe, our work is unique because of our focus on realtime, adaptive, multimodal interaction, rather than offline scripting of behaviors.

1.12.3 Task Level Interaction With Virtual Actors

In earlier work we developed and implemented a reactive planner, which we call the skill network, for adaptively selecting and executing behavior modules, called motor skills, in order to satisfy some motor goal, say, "Go to the door and open it." To do this, however, we need to identify the motor skills explicitly and implicitly required to accomplish the given motor act. We call this process motor goal parsing (MGP), and this year we have designed an architecture to support this process.

While most humans are capable of seemingly effortless manipulation of ordinary objects in everyday life, computational models of such behavior must account for affordances of objects and motor equivalence. That is, given a command that names one or more objects in the virtual world, how can a virtual actor compute at runtime whether and how the named objects can be manipulated? Next, once the virtual actor has determined that an object can be manipulated, the appropriate effector system must be selected. For example, can the object be grasped with one hand or must both arms be employed?

In the past year, we have described the representations we feel are necessary to transform descriptions of everyday activities into a series of motor goals that can be executed by the skill network. To account for motor equivalence and the affordances of objects in the environment, we have specified an effector-free set of task primitives to be used as an intermediate representation in the parsing process, and we have outlined the process of motor goal parsing.[72] It is our intention to implement this architecture in the coming year.

1.13 Research on Improved Sonar Displays: A Human/Machine Processing System

Sponsor

U.S. Navy - Office of Naval Research
Grant N00014-93-1-1198

Project Staff

Dr. David Zeltzer, Nathaniel I. Durlach, Dr. Jeng-Feng Lee, Megan C. Jasek

A decrease in the perceived threat from former Soviet-bloc countries and an increasing threat from quiet, comparatively low-cost, diesel/electric submarines and swimmer delivery vehicles have resulted in strong emphasis on littoral region surveillance and anti-submarine warfare by the U.S. Navy. The inshore, shallow-water environment, however, is extremely challenging for passive and active acoustic sensor systems. Many of the simplifications used in modeling and interpreting deep-water acoustic propagation do not hold in the highly reverberant, range dependent and noisy littoral environment. Due to these factors, not only must new sensors and sensor processing methods be developed, but also new human-machine processing systems (HMPS) which best present and exploit existing sensor systems. Emerging remote presence and virtual environment technologies and techniques now provide sufficient framework to support the type of multi-modal, interactive human-machine processing systems that are envisioned.

Ideally, a general-purpose HMPS would allow viewing and interaction with received sonar signals in light of sonar system configuration and performance information, a reasonable acoustic propagation model, current information about the undersea environment, and information about the overall tactical situation. The HMPS would receive two classes of input signals. The first consists of the outputs of underwater acoustic sensing systems (both active and passive). The second consists of the outputs of auxiliary systems designed to produce continuously updated information on the state of the surface environment (e.g., ship posi-

71 D.T. Chen, S.D. Pieper, S.K. Singh, J.M. Rosen, and D. Zeltzer, "The Virtual Sailor: An Implementation of Interactive Human Body Modeling," *Proceedings of the Virtual Reality International Symposium,* Seattle Washington, September 18-22, 1993; S.K. Singh and S.D. Pieper, "Control and Coordination of Head, Eyes and Facial Expressions of Virtual Actors in Virtual Environments," *Proceedings of the IEEE International Workshop on Robot and Human Communication,* Tokyo, Japan, November 3-5, 1993.

72 D. Zeltzer and M. Johnson, "Virtual Actors and Virtual Environments: Defining, Modeling and Reasoning about Motor Skills," in *Interacting with Virtual Environments,* eds. L. MacDonald and J. Vince (Chichester, England: John Wiley & Sons, 1994, pp. 229-255).

tions) and the underwater environment (e.g., sound velocity profiles).

It is assumed that the first class of signals arises from both active sources in the environment and from reflections of sonar signals off objects in the environment and that these sounds and reflections can arise in connection with hostile targets such as submarines and swimmer-delivery vehicles, as well as various types of clutter, including biologicals, decoys, and commercial traffic. (In this document, we use the word clutter to refer not only to reflections of sonar energy off "irrelevant" objects, but also to interfering active sources.) The output of the HMPS is envisioned as consisting of a probabilistic description of the current environmental situation structured and displayed spatially utilizing multiple sensory modalities in a manner that is well-matched to the decision/action space for command and control. This display would not only facilitate analysis of incoming real-time sonar data but also help sonar personnel understand the acoustical propagation characteristics and tactical import of the current undersea environment.

Within this general context, this research is concerned with (1) determining the optimum processing for the HMPS to perform on the input signals, (2) the manner in which automatic machine processing on the one hand and processing by the human sensory systems and the human brain on the other hand can best be combined, and (3) how to display signals to the human operator to optimize the human processing. Although the focus will be on the last of these three topics, all of these are sufficiently intertwined so that consideration of the first two cannot be entirely eliminated without severely limiting the usefulness of the results.

As a first step in the development of a general purpose HMPS, we will concentrate on real-time, interactive display methodologies for ship-based active sonars. Sonar signals will be synthetically generated and their propagation through the shallow-water environment will be modeled. Inputs to the general acoustic model include bottom characteristics and water column data affecting the range and depth dependent sound velocity profile. In addition, a minimum of one acoustic source and two receivers will be modeled. Major source and receiver parameters which may be varied, in addition to the number of sources and receivers, include location, frequency response, and directional gain characteristics.

Once we have an appropriate model and testbed system up and running, we will then begin to explore display configurations that present information that we believe is particularly relevant to determining the features of the acoustic signal space that are important for ASW and present this information in a manner that takes account of both the strengths and weaknesses of the human processing system and the new technology that is now being developed in the area of human- machine interfaces for virtual environment and teleoperator systems.

Using available technologies, the testbed HMPS will allow modification of the following:

- Characteristics of the displays (in various modalities),
- Relevant undersea environment characteristics,
- Number, position and operating characteristics of sonar system elements,
- Number, position, and characteristics of clutter elements,
- Number, position and operating characteristics of targets.

A variety of spatially oriented representation techniques will be developed and evaluated against existing display techniques (such as "waterfall" techniques) for detection and classification performance. Among the issues to be considered in this proposed research are the following:

1. How best to combine human and machine processing to capitalize on the humans relatively good ability to perform complex pattern recognition tasks and to rapidly adapt to changing conditions, while simultaneously compensating for the humans relatively poor ability to store large amounts of unencoded sensory data and perform large amounts of computation rapidly;
2. How best to combine visual and auditory displays to capitalize on the respective strengths of these two modalities;
3. How best to present intersensor correlation information along with single-sensor information;
4. How best to exploit the sense of immersion and the ability to manipulate the observers point of view that can be provided by virtual environment display systems;
5. How best to partition and display information to groups of humans as well as single individuals; and
6. How best to incorporate decision support aids for localization and identification of sonar signals.

During the first half of the first year, we have evaluated a number of underwater acoustic propagation models, and we have selected a simplified propagation model for use in the program. Design of the preliminary human/machine interface is currently underway.

(Center at left) Senior Research Scientist Nathaniel I. Durlach, (center at right) Professor Louis D.B. Braida, and other staff members and students of RLE's RLE Sensory Communication Group, some wearing the group's research devices. Clockwise from left, Visiting Scientist Yun Shao, Research Specialist Hong Z. Tan, student Ross A. Yu, Research Specialist Lorraine A. Delhorne, graduate student Barbara G. Shinn-Cunningham, Reseach Assistant Walter A. Aviles, Administrative Secretary Eleanora M. Luongo, Visiting Scientist Min Wei, and Administrative Secretary Ann K. Dix.