Section 2 Sensory Communication

Chapter 1. Sensory Communication

Academic and Research Staff

Professor Louis D. Braida, Nathaniel I. Durlach, Professor Richard M. Held, Professor Anuradha M. Annaswamy, Dr. David L. Brock, Dr. Paul Duchnowski, Dr. Donald K. Eddington, Dr. Susan L. Goldman, Dr. Kenneth W. Grant, Dr. Julie E. Greenberg, Dr. Lynette A. Jones, Dr. Jeng-Feng Lee, Dr. Karen L. Payton, Dr. Matthew H. Power, Dr. William M. Rabinowitz, Dr. Christine M. Rankovic, Dr. Charlotte M. Reed, Dr. Wendelin L. Sachtler, Dr. J. Kenneth Salisbury, Dr. Kaoru Sekiyama, Dr. Barbara G. Shinn-Cunningham, Dr. Mandayam A. Srinivasan, Dr. Anne H. Takeuchi, Dr. Thomas E.v. Wiegand, Dr. David Zeltzer, Dr. Patrick M. Zurek, Merry A. Brantley, Andrew R. Brughera, Lorraine A. Delhorne, Dorrie Hall, Seth M. Hall, Coral D. Tassa

Visiting Scientists and Research Affiliates

Dr. G. Lee Beauregard, Dr. Kalman Glantz, Geoffrey L. Plant, Dr. Kouroush Saberi

Graduate Students

Walter A. Aviles, Maroula S. Bratakos, Douglas S. Brungart, Jyh-Shing Chen, Kiran Dandekar, Suvranu De, Joseph G. Desloge, Jeffrey J. Foley, Eric M. Foxlin, Joseph A. Frisbie, Isaac Graaf, Rogeve J. Gulati, Rakesh Gupta, Chih-Hao Ho, Alexandra I. Hou, Louise Jandura, Owen D. Johnson, Steingrimur P. Karason, Jean C. Krause, Michael A. Leabman, Gregory G. Lin, Kinuku Masaki, Hugh B. Morgenbesser, Philip M. Nadeau, Michael P. O'Connell, John Park, Nicholas Pioch, Balasundara I. Raju, Christopher R. Richardson, John J. Rodkin, David W. Schloerb, Matthew G. Sexton, Nathan R. Shnidman, Jason Sroka, Hong Z. Tan, Kimberly J. Voss, Daniel P. Welker, Elron A. Yellin, John Yoon

Undergraduate Students

Stephen V. Baird, James H. Bandy, Susan E. Born, Erika N. Carmel, Frederick W. Chen, Gail Denesvich, Carmen Ho, Gabrielle Jones, Danielle G. Lemay, Rebecca F. Lippman, David C. Lossos, John J. Novak, T.H. Ogora, Jonathan Pfautz, Sudeep Rangaswamy, Frederick L. Roby, Jonathan R. Santos, Ranjini Srikantiah, Lukasz A. Weber, Evan F. Wies

Technical and Support Staff

Ann K. Dix, David S. Lum, Eleanora M. Luongo, Francis G. Taylor

1.1 Introduction

The Sensory Communication Group is interested in understanding sensorimotor and cognitive processes and exploiting this understanding to solve practical problems in a variety of application domains. Our basic research is characterized by behavioral (psychophysical) experimentation and quantitative theoretical modeling. Facilities to support this research are developed as needed. Although some research is conducted on vision, most of the group's work is focused on audition or taction. The main applications are concerned with aiding individuals who suffer form impaired hearing, developing improved human-machine interfaces for virtual environments and teleoperation (virtual reality), and the use of virtual environment technology for training. The main facilities of the group are associated with the psychoacoustics, touch (or haptics), and virtual environment laboratories.

In the following discussion, research is organized according to the sources of funding provided to our group. Thus, work on hearing and touch, and to a certain extent vision, is described in a variety of subsections in a variety of funding contexts.

1.2 Hearing Aid Research

Sponsor

National Institutes of Health
Grant RO1 DC00117

Project Staff

Professor Louis D. Braida, Dr. Paul Duchnowski, Dr. Karen L. Payton, Dr. Matthew H. Power, Dr. Christine M. Rankovic, Dr. Charlotte M. Reed, Dr. Patrick M. Zurek, Jeffrey J. Foley, Isaac Graaf, Jean C. Krause, T.H. Ogora, Jason Sroka

1.2.1 Specific Aims

Our long-term goal is to develop improved hearing aids for people suffering from sensorineural hearing impairments. Our efforts are focused on problems resulting from inadequate knowledge of the effects of various alterations of speech signals on speech reception by impaired listeners. Specifically, we seek to increase our knowledge of the fundamental limits to speech reception improvement achieved through speech processing.

Our aims are:

1. To assess the relative contribution of various functional characteristics of hearing impairments to reduced speech-reception capacity.
2. To evaluate the effect of the manner of speech articulation and variability in the production on speech reception by hearing-impaired listeners.
3. To develop and evaluate analytical models that can predict the effect of a variety of alterations of the speech signal on intelligibility.
4. To develop and evaluate signal processing techniques that might increase the effectiveness of hearing aids.

1.2.2 Characteristics of Sensorineural Hearing Impairment

Simulation of Sensorineural Hearing Loss

An accurate perceptual model of sensorineural hearing loss could assist clinicians and others who deal with hearing-impaired listeners to experience the effects of hearing impairment. It may also provide insight into the design of hearing aids intended to overcome these effects. Although there is no satisfactory model of sensorineural hearing impairment, several *functional models* that allow listeners with normal hearing to experience the perceptual distortions and speech-reception problems of hearing-impaired listeners are emerging. Our past work[1] has documented the effectiveness of masking noise that elevates the tone-detection thresholds of normal-hearing listeners to those of listeners with hearing impairments in simulating limitations on speech-reception associated with the impairment. More recently, Duchnowski and Zurek[2] have shown that a type of multiband amplitude expansion proposed by Villchur[3] is also capable of simulating these limitations. Both the noise and expansion simulations of hearing loss are addressed nominally at the minimal factors of audibility and abnormal loudness growth. These simulations differ both with respect to phenomenological realism and in the relation of the acoustic properties of processed stimuli to the normal listener's hearing. Moreover, they may differ with respect to secondary psychoacoustic variables. Our research helps determine the relative importance of these variables.

In previous work,[4] we have shown that many of the perceptual effects of sensorineural hearing loss can be simulated for listeners with normal hearing using spectrally-shaped masking noise or multiband expansion amplification. Current research is exploring the use of these simulations to interpret the results of studies of signal processing techniques under consideration for use in hearing-aids. During the past year we applied the simulations to a popular two-band wide-dynamic range compression hearing aid that has been studied

---

[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] P. Duchnowski and P.M. Zurek, "Villchur Revisited: Another Look at AGC Simulation of Recruiting Hearing Loss," *J. Acoust. Soc. Am.* 98: 3170-3181 (1995).

[3] 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).

[4] 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); P. Duchnowski and P.M. Zurek, "Villchur Revisited: Another Look at AGC Simulation of Recruiting Hearing Loss," *J. Acoust. Soc. Am.* 98: 3170-3181 (1995); S.V. DeGennaro and L.D. Braida, "Effects of Linear Amplification and Amplitude Compression on Sentence Reception by Listeners with Hearing Loss Simulated by Masking Noise," in *Modelling Sensorineural Hearing Loss,* ed. W. Jesteadt (Mahwah, New Jersey: L. Earlbaum Assoc., 1996); D.S. Lum and L.D. Braida. "DSP Implementation of a Real-Time Hearing Loss Simulator Based on Dynamic Expansion," in *Modelling Sensorineural Hearing Loss,* ed. W. Jesteadt. (Mahwah, New Jersey: L. Earlbaum Assoc., 1996).

recently.[5] In our study, sentences in background noise were processsed with either a dual-channel compression system or a linear amplification system and presented to the simulated normal listeners. The systems were implemented to employ the same cutoff frequency and gain functions as the aids used in the Moore et al. study.

Two normal hearing listeners were used to simulate the impairment of each of two hearing-impaired participants in the Moore et al. study. These participants were selected as having achieved relatively large gains from compression amplification relative to linear amplification. Each listener was tested with both the expansion and the additive noise simulations. The speech reception thresholds (SNR at which the listener scores 50 percent of the words correct) of the simulated normal listeners aided by both a compression system and a linear system were obtained. These SRTs were compared to the corresponding SRTs of the aided hearing-impaired individuals.

Although the small number of listeners tested does not allow for a statistical analysis of the results, some trends were observed in the performance of the simulated normals. First, linear amplification produced lower SRTs than amplitude compression in seven out of eight situations. This is in marked contrast with findings of Moore et al.: both hearing-impaired individuals had SRTs that were about 2 dB lower with compression than with linear processing. Second, SRTs were lower with the expansion simulation than with additive noise in seven of eight situations. A quantitative comparison of the SRTs in the two studies is not possible due to several differences in the experimental set-up, e.g., sentences presented monaurally in this study, but binaurally in the Moore et al. study.

Future work will focus on the discrepancy in the relative performance of linear amplification and compression amplification processing systems for listeners with real and simulated sensorineural hearing loss. We plan to implement the speech processing systems in real time to determine optimal signal processing parameters for both systems. If linear amplification is found consistently to out-perform amplitude compression for the simulated normals, then we will attempt to identify the source of the discrepancy and investigate other methods of simulating the hearing loss.

1.2.3 Characteristics of the Speech Signal

Models of Speech Intelligibility

Models of speech intelligibility can help quantify those aspects of the speech signal that are essential for communication. Unlike the articulation index (AI),[6] which bases predictions of speech intelligibility on the long-term power spectrum of speech, the speech transmission index (STI)[7] bases predictions on the intensity modulation spectra for individual frequency bands. This distinction is fundamental to the understanding of the dependence of intelligibility on speaking mode. Whereas clear and conversational speech have roughly the same long-term spectrum,[8] they are likely to have different modulation spectra because of differences in speaking rate and consonant-vowel energy (CV) ratios. Also the effects of reverberation and multiband amplitude compression on the speech signal are likely to be more evident in the intensity modulation spectra than in the power spectrum.

We are developing a method of calculating the Speech Transmission Index from measurements made on real speech waveforms. We have continued to develop techniques to remove noise-induced artifacts in speech envelope spectra. After several unsuccessful attempts to predict the artifacts and subsequently remove them, we chose to alter our approach. We investigated means to predict when the artifacts would dominate the envelope spectra and thus specify an upper modulation frequency limit for our spectra. We found that reliable Modulation Transfer Functions (MTFs) could be computed from speech envelope spectra if we computed the coherence function and used it to

5 B.C.J. Moore, J.S. Johnson, T.M. Clark, and V. Pluvinage, "Evaluation of a Dual-Channel Full Dynamic Range Compression System for People with Sensorineural Hearing Loss," *Ear Hear.* 13: 349-370 (1992).

6 R.L. Dugal, L.D. Braida, and N.I. Durlach, "Implications of Previous Research for the Selection of Frequency-Gain Characteristics," in *Acoustical Factors Affecting Hearing Aid Performance and Measurement,* eds. G.A. Studebaker and I. Hochberg (New York: Academic Press, 1980).

7 T. Houtgast and H.J.M. Steeneken, "A Review of the MTF Concept in Room Acoustics and Its Use for Estimating Speech Intelligibility in Auditoria," *J. Acoust. Soc. Am.* 77: 1069-1077 (1985).

8 M.A. Picheny, N.I. Durlach, and L.D. Braida, "Speaking Clearly for the Hard of Hearing II: Acoustic Characteristics of Clear and Conversational Speech," *J. Speech Hear. Res.* 29: 434-446 (1986).

indicate a modulation frequency limit.[9] A coherence criterion level of 0.8 was found to be appropriate for clear and conversational speech in noise. This same criterion level was also found to truncate reverberant speech envelope spectra in both speaking styles at appropriate modulation frequencies to remove the artifacts seen there.

We plan to compare the Speech Transmission Index (STI) computed from our truncated MTFs with those computed theoretically for the experimental conditions reported by Payton et al.[10] In addition, we plan to compute the envelope spectra of amplitude compressed speech and amplitude compressed noisy speech. STIs derived from these amplitude compressed signals will be compared with intelligibility scores obtained for the same conditions.

1.2.4 Clear Speech

Although our previous studies[11] demonstrated that clear speech is generally more intelligible than conversational speech, it is also generally enunciated more slowly. This complicates the problem of determining what is responsible for the difference in intelligibility. Moreover, the degree to which the reduction in speaking rate is essential to clarity has not been established. To improve our understanding of the role of speaking rate in determining intelligibility, Krause[12] trained talkers with significant speaking experience to produce clear and conversational speech at slow, normal, and quick rates. In addition, each talker was provided feedback on intelligibility on a sentence by sentence basis in order to elicit the clearest possible speech at each speaking rate.

Intelligibility tests used nonsense sentences lacking in semantic context and were presented in a background of speech-spectrum noise. Key word scores indicated that clear speech was more intelligible than conversational speech at speaking rates up to roughly 200 wpm, When the speaking rate exceeded 200 wpm, the intelligibility of both clear and conversational speech declined, suggesting that there may be physiological limitations on clarity at high speaking rates. These results suggest that acoustical factors other than reduced speaking rate are responsible for the high intelligibility of clear speech.

We plan to characterize the acoustical differences between clear and conversational speech produced at the same speaking rate. Our new database of clear and conversational speech will be phonetically segmented and labeled. Several short-term acoustic differences between clear and conversational speech will be investigated for their effect on intelligibility. For example, the consonant-to-vowel ratios (CVR) found in clear speech are substantially greater than those in conversational speech.[13] Recently, two studies have reported an intelligibility improvement due to artificial CVR increase for nonsense syllables presented to impaired listeners.[14]

1.2.5 Signal Processing For Hearing Aids

Frequency-lowering individuals who experience severe hearing losses at high frequencies typically experience difficulty understanding speech because they are unable to distinguish between consonants whose spectra differ primarily at these frequencies. Although there have been many attempts to use frequency lowering techniques to improve the audibility of these differences, few have produced net improvements in speech perception. Posen et

9 K.L. Payton and L.D. Braida, "Speech Modulation Transfer Functions for Different Speaking Styles," *J. Acoust. Soc. Am.* 98: 2982 (1995).

10 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).

11 M.A. Picheny, N.I. Durlach, and L.D. Braida, "Speaking Clearly for the Hard of Hearing II: Acoustic Characteristics of Clear and Conversational Speech," *J. Speech Hear. Res.* 29: 434-446 (1986); R.M. Uchanski, S. Choi, L.D. Braida, C.M. Reed, and N.I. Durlach, "Speaking Clearly for the Hard of Hearing IV: Further Studies of the Role of Speaking Rate," *J. Speech Hear. Res.*, forthcoming.

12 J.C. Krause, *The Effects of Speaking Rate and Speaking Mode on Intelligibility,* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1995.

13 M.A. Picheny, N.I. Durlach, and L.D. Braida, "Speaking Clearly for the Hard of Hearing II: Acoustic Characteristics of Clear and conversational Speech," *J. Speech Hear. Res.* 29: 434-446 (1986).

14 A.A. Montgomery and R.A. Edge, "Evaluation of Two Speech Enhancement Techniques to Improve Intelligibility for Hearing-impaired Adults," *J. Speech Hear. Res.* 29: 434-446 (1986); S. Gordon-Salant, "Effects of Acoustic Modification on Consonant Perception by Elderly Hearing-impaired Subjects," *J. Acoust. Soc. Am.* 77: S106 (1986).

al.[15] recently demonstrated that channel vocoding could be used to superimpose cues that enchanced the ability to make distinctions among such consonants at low frequencies where hearing is often less impaired. To facilitate future studies of this technique, Foley[16] implemented Posen's vocoder system digitally using a Motorola DSP96002 DSP and Ariel DSP96 signal processing board suitable in one of our laboratory PCs. This implementation allows greater flexibility and precision in the selection of vocoder filters and the determination of speech levels, and also allows for monitoring and dynamic adjustment of processing parameters.

1.2.6 Publications

DeGennaro, S.V., and L.D. Braida. "Effects of Linear Amplification and Amplitude Compression on Sentence Reception by Listeners with Hearing Loss Simulated by Masking Noise." In *Modelling Sensorineural Hearing Loss.* Ed. W. Jesteadt. Mahwah, New Jersey: L. Earlbaum Assoc., 1996.

Duchnowski, P., and P.M. Zurek. "Villchur Revisited: Another Look at AGC Simulation of Recruiting Hearing Loss." *J. Acoust. Soc. Am.* 98: 3170-3181 (1995).

Lum, D.S., and L.D. Braida. "DSP Implementation of a Real-Time Hearing Loss Simulator Based on Dynamic Expansion." In *Modelling Sensorineural Hearing Loss.* Ed. W. Jesteadt. Mahwah, New Jersey: L. Earlbaum Assoc., 1996.

Maxwell, J.A., and P.M. Zurek. "Feedback Reduction in Hearing Aids." *IEEE Trans. Speech Audio Proc.* 3: 304-313 (1995).

Payton, K.L., and L.D. Braida. "Speech Modulation Transfer Functions for Different Speaking Styles." *J. Acoust. Soc. Am.* 98: 2982 (1995).

Power, M.H., and L.D. Braida. "Consistency among Speech Parameter Vectors: Application to Predicting Speech Intelligibility." Submitted to *J. Acoust. Soc. Am.*

Rankovic, C.A. "Derivation of Frequency-gain Characteristics for Maximizing Speech Reception in Noise." *J. Speech Hear. Res.* 29: 434-446 (1995).

Uchanski, R.M., S. Choi, L.D. Braida, C.M. Reed, and N.I. Durlach. "Speaking Clearly for the Hard of Hearing IV: Further Studies of the Role of Speaking Rate," *J. Speech Hear. Res.* Forthcoming.

Theses

Foley, J.J. *Digital Implementation of a Frequency-Lowering Channel Vocoder.* M.Eng. thesis. Dept. of Electr. Eng.and and Comput. Sci., MIT, 1996.

Krause, J.C. *The Effects of Speaking Rate and Speaking Mode on Intelligibility,* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1995.

1.3 Enhanced Communication for Speechreaders

Sponsor

National Institutes of Health
Grant RO1 DC02032

Project Staff

Professor Louis D. Braida, Dr. Paul Duchnowski, Dr. Susan L. Goldman, Dr. Matthew H. Power, Dr. Anne H. Takeuchi, Lorraine A. Delhorne, Dr. Charlotte M. Reed, Dr. Kaoru Sekiyama, Dr. Kenneth W. Grant, Maroula S. Bratakos, Frederick W. Chen, Joseph A. Frisbie, Gabrielle Jones, Danielle G. Lemay, Rebecca F. Lippman, Philip M. Nadeau, Matthew G. Sexton, Ranjini Srikantiah

1.3.1 Specific Aims

Although speechreading is an essential component of communications for the hearing impaired, under nearly all conditions the ability to communicate through speechreading alone is severely constrained. This is because many acoustic distinctions important to communication are not manifest visually. Supplements derived from the

[15] M.P. Posen, C.M. Reed, and L.D. Braida, "The Intelligibility of Frequency-Lowered Speech Produced by a Channel Vocoder," *J. Rehab. Res. and Dev.* 30: 26-38 (1993).

[16] J.J. Foley, *Digital Implementation of a Frequency-Lowering Channel Vocoder,* M.Eng thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1996.

acoustic speech signal have been shown to improve speech reception markedly when the cues they provide are integrated with cues derived from the visible actions of the talker's face. Our long-term goal is to develop aids for individuals with hearing impairments so severe that their communication relies heavily on speechreading.

Our aims are:

1. To develop models of audiovisual integration to quantify how well supplementary signals are integrated with speechreading.
2. To develop and evaluate simplified acoustic signals that can be derived from acoustic speech by signal processing. Such signals would form the basis of new types of hearing aids for listeners with severe hearing impairments.
3. To develop systems for producing and displaying discrete speechreading supplements similar to the "Manual Cued Speech System" that can be derived from the acoustic signal by speech recognition technology. Such supplements would display streams of discrete symbols that would be derived automatically from acoustic speech and presented visually for integration with the speechreading signal.

1.3.2 Studies and Results

Supplements Based on Signal Processing

Acoustic supplements based on the amplitude envelopes of filtered bands of the acoustic speech signal have been shown to improve speech reception via speechreading in normal-hearing listeners.[17] These supplements may also aid a range of hearing-impaired listeners because they place minimal demands on auditory capacities. The use of band envelopes requires only intensive and temporal resolution without the need to resolve fine spectral detail. Furthermore, envelopes can be easily extracted from continuous speech and transposed in frequency to fit the listener's best region of hearing.

Field Studies of Band-Envelope Supplements

These supplements consist of one or more tones whose frequency is in the audible area of the impaired listener and whose amplitude is modulated by the amplitude envelopes of one or more bands of the speech signal. To assess the effectiveness of these supplements in everyday situations, Nadeau[18] programmed wearable DSP-based SiVo Aids[19] to produce these supplements to listeners with severe to profound impairments. Seven impaired listeners who provided audiograms were screened; two of these were selected as likely to benefit more from the amplitude-envelopes than from their own aids (on the basis of data from listeners with normal hearing) and were provided with wearable SiVo aids. Each listener wore the aid for four weeks in each processing condition. Every two weeks the listeners received speechreading tests in the laboratory and completed the PHAB questionaire to assess reaction to the aid. One listener received substantial benefit from the aid (word scores of 22 percent by speechreading alone, 38 percent for speechreading supplemented by her own aid, 51 percent for the two-tone configuration of the SiVo aid, and 55 percent for the one-tone configuration of the SiVo aid). Her reactions were sufficiently positive that she continues to use the SiVo aid in the one-tone configuration in preference to her own aid. The second listener also was able to make use of the envelope signals (word scores of 43 percent by speechreading alone, 47 percent for the two-tone configuration of the SiVo aid, and 55 percent for the one-tone configuration of the SiVo aid), but received greater benefit from her own hearing aid (a word score of 97 percent). Not surprisingly she has elected to continue using her own aid. A major limitation of this study, aside from the small number of listeners tested, is the inability of the SiVo aid to produce the two-tone signal at sufficiently high levels.

We are in the process of having the SiVo aids modified by their manufacturer, and are in the process of identifying a second set of impaired listeners. We plan to begin a second field trial during summer 1996 and a third six months later.

---

17 K.W. Grant, L.D. Braida, and R.J. Renn, "Single Band Amplitude Envelope Cues as an Aid to Speechreading," *Quart. J. Exp. Psych.* 43: 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).

18 P.M. Nadeau, *Amplitude Envelope Cues as an Aid to the Speechreading of Impaired Individuals,* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, February 1996.

19 A. Faulkner, V. Ball, S. Rosen, B.C.J. Moore, and A. Fourcin, "Speech Pattern Hearing Aids for the Profoundly Hearing Impaired: Speech Perception and Auditory Abilities," *J. Acoust. Soc. Am.* 91: 2136-2155 (1992).

Supplements Based on Automatic Speech Recognition

Visual supplements related to the manual cued speech system used in the education of the deaf have been shown to improve speechreading dramatically. These supplements can be derived using the techniques of automatic speech recognition (ASR), a technology that has advanced rapidly during the past decade. Automatically generated cues would not require a speaker to be skilled in cue production, and so could assist a wider variety of individuals in everyday situations. Two major hurdles must be overcome to develop an effective automatic cueing system: the ASR technology used must provide cues with sufficient accuracy; and the cues must be displayed in a manner that can be integrated well with speechreading.

Simulations of Automatic Cueing Systems

Although previous studies of the reception of manual cued speech and the performance of ASR systems suggests that current technology may provide an adequate basis for the design of an automatic cueing system,[20] the effects of speech recognition errors and delays associated with the recognition process on cue reception and integration are not well understood.

To estimate the effects of such imperfections we are studying of the reception of simulated cued speech. This simulation is constructed using prerecorded IEEE sentences that had been phonetically marked and labeled using the acoustic speech signal. Visual cues consisting of static images of the hand of the talker are dubbed onto these recordings. The nominal shapes and positions of the hands are those used in manual cued speech but there is no fluid articulation of movement. The determination of the cues to be superimposed is based on the rules of manual cued speech applied to (possibly) flawed phonetic labels. Three control conditions were included in the tests: SA (speechreading alone), MCS (manual cued speech), and PSC (cues derived from flawless phonetic labels). In the experimental conditions the phone sequence contained 0, 10, or 20 percent errors, with the error pattern corresponding to a Bernoulli process and with phone substitutions corresponding to our measurements of a real recognizer. We also studied the effect of delaying the superimposed cue relative to the facial image by 0-167 ms. A final test used cues derived from the phone estimates produced by a state of the art phonetic recognizer trained on roughly 1000 sentences produced by the speaker of the sentences. The dubbed materials were presented to two groups of experienced young adult cue receivers.

The first group (four cue receivers) obtained average scores of 25 percent on SA, 78 percent on MCS, and 73 percent on PSC. Scores declined to 59 percent and 49 percent when cues were produced with phone error rates of 10 and 20 percent respectively. The effect of a small delay (0.03 sec) in cue production was negligible when the phone error rate was 20 percent. However scores declined to 37 percent when cues were produced with 20 percent phone errors and 0.1 sec of delay and to 35 percent when for cues produced with 10 percent phone errors and 0.17 sec of delay.

The second group (6 cue receivers, including three members of the first group) obtained slightly higher scores in the control conditions (36 percent with SA, 89 percent with MCS, and 81 percent with PSC). When tested under conditions of random cue delay (±0.033 sec), their scores declined to 76 percent with no phone error, 69 percent with 10 percent phone errors, and 60 percent with 20 percent phone errors. This group was also tested on cues derived from the recorded acoustical speech signal by a real ASR system. Although 21 percent of the phones produced by the system were in error, word scores averaged 70 percent. Subsequent analyses indicated that the distribution of errors was different for the real recognizer than for the simulated error conditions, i.e., an excess of multiple phone errors occurring in single words.

For both groups of cue receivers differences between scores in the the MCS and PSC conditions were not statistically significant, indicating that subjects were relatively unaffected by the discrete nature of the cues, the faster speaking rate (150 wpm versus 100 wpm) and deviations between the articulated phone sequence and the phoneme sequence corresponding to the orthographic representation of the sentences. Also, while both groups were largely unaffected by superposition delays on the order of 30 ms (whether fixed or random), delays of 100 ms or more were detrimental. This indicates that practical automatic cueing systems must compensate for the delay associated with the time required for the recognition process.

Based on the results of our simulation studies, we have begun to design an real-time cueing system

[20] 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. Dev.* 31, 20-41 (1994).

based on an automatic speech recognizer. This system is intended to facilitate future studies of the application of ASR technology to the automatic cueing application. The initial system is tailored to individual speakers and consists of three parts: the acoustic processor, the phonetic recognizer, and the cue display, each of which is implemented on a separate laboratory PC. The acoustic processor samples the speech waveform (which will be sensed initially by a close-talking microphone) and computes parameters of waveform segments at a frame rate of 100/sec. The speaker-dependent phonetic recognizer, which is based on the Entropic HTK software package, derives estimates of the phone sequence from these parameters. The cue display system buffers a sequence of images of the face of the talker, allowing for delays in the recognition process, superimposing pre-recorded images of the talker's hands with shapes and at positions corresponding to the recongnized phone sequence.

We plan to continue our studies of the effects of cue imperfections on the reception of manual cued speech. We also plan to complete the initial development of our experimental automatic cueing system during the Summer of 1996 and to use it as the basis of these studies.

1.3.3 Publications

Takeuchi, A., and L.D. Braida. "Effect of Frequency Transposition on the Discrimination of Amplitude Envelope Patterns." *J. Acoust. Soc. Am.* 97: 453-460.

Takeuchi, A., and L.D. Braida. "Recognition of Envelope Patterns in the Presence of a Distractor: I. Effects of Correlation and Frequency Relation." *J. Acoust. Soc. Am.* 98: 135-141 (1995).

Takeuchi, A., and L.D. Braida. "Recognition of Envelope Patterns in the Presence of a Distractor: II. Effects of Dichotic Presentation and Unmodulated Distractors." *J. Acoust. Soc. Am.* 98: 142-147 (1995).

Theses

Bratakos, M.A. *The Effect of Imperfect Cues on the Reception of Cued Speech.* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1995.

Nadeau, P.M. *Amplitude Envelope Cues as an Aid to the Speechreading of Impaired Individuals.* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1996.

1.4 Cochlear Implants

Sponsors

National Institutes of Health
Contract P01-DC00361[21]
Contract N01-DC22402

Project Staff

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

Most people with profound hearing impairments have lost the ability to translate the acoustic energy of sound into electric signals carried to the brain by the auditory nerve. Cochlear implants bypass the impaired transduction mechanism and directly stimulate the auditory nerve via an electrode array implanted within the cochlea. Our research is directed at understanding the fundamental mechanisms of electroauditory stimulation and using that information to develop cochlear-implant systems for improved speech reception.

During the past year, work has involved initiation of studies on amplitude modulation discrimination and continuing evaluations of speech processors based on the continuous-interleaved-stimulation (CIS) strategy.[22] Work in the latter area is described below (and is performed with collaborators listed in section 8 of the following chapter on (signal transmission in the auditory system).

In the latter part of 1994, a prototype portable processor (based on a DSP-56001) became available that could realize some CIS implementations. Two subjects with nine years of previous experience using the standard INERAID sound processor were fit with CIS systems that they have subsequently worn full-time. Measures of speech reception have been obtained longitudinally; all testing is

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

[22] B.S. Wilson, C.C. Finley, D.T. Lawson, R.D. Wolford, D.K. Eddington, and W.M. Rabinowitz, "Better Speech Reception with Cochlear Implants," *Nature* 352: 236-238 (1991).

performed without lipreading. With their INERAID systems, one subject (S1) is an average performer, and the other (S2) is a high performer.

Tests of consonant and vowel identification show learning effects for both subjects. With S1, initial scores with CIS were below INERAID values; over time, however, CIS performance increased and now exceeds INERAID values. S2 showed modest gains with CIS immediately, and futher improvements over time. Consonantal feature analyses indicated that both subjects score highest for manner and voicing information and that scores with both CIS and INERAID are similar on these features. Score for place information is lower than that for manner and voicing, but increases with CIS (versus INERAID). This improved reception of place information provides important cues to enhanced perception of spectral shape, and it appears responsible for the overall increases in consonant recognition.

The ability to recognize words in sentences was tested in quiet and in noise. In quiet, S1 was tested with everyday sentences. He scored 36 percent (words correct) with his INERAID, dropped to 14 percent for his initial CIS evaluation, and increased to 70 percent over 24 weeks of CIS use. S2 was tested with more difficult, lower context sentences. He scored 62 percent with his INERAID and showed large and immediate gains with CIS scoring at, and then above, 95 percent.

Speech perception in noise was assessed by presenting sentences against a background of speech-spectrum-shaped noise at specified speech-to-noise ratios (RMS). At +8 dB S/N, S1 shows no gain with CIS (versus time or INERAID). For S2, high scores (above 80 percent) were obtained immediately with CIS; increases to near saturation (over 95 percent) occurred over several weeks of use. At a more difficult +2 dB S/N, a stronger learning effect was evident as performance increased from 35 percent at the time of fitting to about 60 percent after several weeks of use. Comparison to INERAID results indicates that S2's gains with CIS correspond to more than a 6 dB increase in noise tolerance (i.e., CIS at +2 dB exceeds INERAID at +8 dB). Furthermore, his CIS performance is only ~ 7 dB below that of normal-hearing listeners; this level of performance is remarkable, with interesting implications for understanding normal speech perception.

In summary, the data from both subjects show that substantial gains in speech reception have occurred in these CIS trials. The fact that these gains may not be immediate, as S1 clearly shows, reinforces the need for longitudinal studies with wearable systems. Finally, although we focus on objective measures of speech reception, comments from the subjects indicate that these CIS processors also provide significant enhancements in environmental sound and music perception.

1.4.1 Publications and Conference Presentations

Rabinowitz, W.M., and D.K. Eddington. "Effects of Channel-to-Electrode Mappings on Speech Reception with the Ineraid Cochlear Implant." *Ear Hear.* 16: 450-458 (1995).

Rabinowitz, W.M., D.K. Eddington, J. Tierney, L.A. Delhorne, and M.E. Whearty. "Performance with Wearable CIS Processors." Conference on Implantable Auditory Prostheses, Asilomar, California, August 19-24, 1995.

Rabinowitz, W.M., D.K. Eddington, J. Tierney, and L.A. Delhorne. "Preliminary Evaluations of Cochlear Implantees using a Wearable CIS Processor." *J. Acoust. Soc. Am.* 97: 3346(A) (1995).

1.5 Tactile Communication of Speech

Sponsor

National Institutes of Health/National Institute on Deafness and Other Communication Disorders Grant 2 R01 DC00126

Project Staff

Lorraine A. Delhorne, Gail Denesvich, Nathaniel I. Durlach, Seth M. Hall, Chih-Hao Ho, Dr. Geoffrey L. Plant, Dr. William M. Rabinowitz, Dr. Charlotte M. Reed, Dr. Mandayam A. Srinivasan, Hong Z. Tan, Jonathan R. Santos

This research is directed towards the development of effective tactual communication devices for individuals with profound deafness or deaf-blindness. Such devices would lead to improved speech reception, speech production, language competence, and awareness of environmental sounds for such individuals and would provide them with a sensory-aid option in addition to hearing aids and cochlear implants. At a more basic scientific level, this research contributes to increased understanding of speech communication, environmental-sound reception, tactual perception, manual sensing, display design, and sensory substitution.

Research in each of four project areas is described below.

1.5.1 Basic Studies of Hand Stimulation and Active Touch

Work in this area includes completion of a device for tactual stimulation, use of this display in studies of information-transfer, and experimental and theoretical work on the discrimination of thickness.

A multifinger positional display was designed to be capable of providing stimulation along a continuum from low-frequency, high-amplitude movements (kinesthetic stimulation) to high-frequency, low-amplitude vibrations (cutaneous stimulation). The device consists of three independent channels interfaced with the fingerpads of the thumb, the index finger and the middle finger. Stimuli from threshold to roughly 50 dB SL can be delivered throughout the frequency range from near DC to above 300 Hz.

With this display, absolute identification experiments were conducted using multicomponent stimuli formed by summing sinusoids from each of three frequency regions that evoked relatively distinct perceptual attributes. These signals were presented to any one or to all three of the digits. Information transfer with this set of signals at durations of 500 and 250 msec was roughly 6.6 bits, and decreased slightly (to 6.0 bits) at a duration of 125 msec.

Identification of stimuli from this set was also examined in a masking paradigm where the stimulus to be identified was preceded and followed by randomly selected signals (drawn from the same set). The optimal stimulus presentation rate was estimated to be approximately 3 items/sec regardless of stimulus duration, and the information-transfer rate was estimated to be approximately 13 bits/sec. This information-transfer rate is roughly the same as the rate achieved by Tadoma users in speech communication.

A study of thickness discrimination has been continued through experimental and theortical work. An orderly set of results has been obtained on the discrimination of thickness for plastic and steel plates, indicating that the size of the jnd in mm is constant (at roughly 0.3 mm) when reference thickness exceeds some critical value (roughly 1 mm for plastic and 0.25 mm for steel). Below the critical value of reference thickness, the size of the jnd decreases sharply. The experimental results are interpreted in terms of the contributions of kinesthetic and cutaneous cues for bendable versus non-bendable plates and have been successfully modeled using a three-dimensional finite-method analysis of plate deformation.

1.5.2 Evaluation of Wearable Tactile Aids

Research in this area includes evaluation of deaf adult users of wearable tactile aids, tactile training programs with deaf children and adolescents, and development of training materials for use with the Tactaid 7 device.

Speech-reception testing has been conducted on two adults who routinely use the Tactaid 7 device in conjunction with hearing aids. Results indicate that improvements to speechreading provided by the tactile aid or hearing aid are comparable, and that their combined use may lead to small improvements over the use of either aid alone.

Speech-production training using the Tactaid 7 device was provided twice weekly over one year to an 18-year-old male who is congenitally deaf. Tapes of the subject producing /a/-C-/a/ syllables and common words in sentences were made pre- and post-training. The intelligibility of these utterances was examined in listening tests using normal-hearing subjects and indicated significant improvements in intelligibility for the post-training tokens.

Weekly tactile training sessions with three children fitted with the Tactaid 7 device in September 1993 are continuing on a weekly basis during the school year. The training addresses improving the childrens' speech production and speech perception skills using input from the tactile aid. A training program for adults has been developed. The program consists of 22 sessions which present both analytic and synthetic training materials using the Tactaid 7 device.

1.5.3 Development of Improved Tactual Supplements to Speechreading

Work in this area includes continued study of the differences in performance for auditory versus tactual presentation of a single-band envelope cue as a supplement to speechreading. In particular, investigation has continued into the relation between benefits to speechreading with the envelope cue and the psychophysical ability to detect changes in amplitude modulation of a 200-Hz carrier tone. In addition, performance on the reception of consonant segments has been measured in adult tactile aid users to determine the contribution of various features, including voicing, manner, and place, to improvements in speechreading through tactual input.

1.5.4 Study of the Reception of Environmental Sounds through Tactual Displays

Deaf users place a high value on the assistance provided by tactile aids in receiving and interpreting non-speech sounds in the person's environment. The reception of environmental sounds provides a sense of security and connection with the outside world and is associated with a sense of psychological well-being. To date, other researchers have shown relatively little concern about the reception and evaluation of environmental sounds through tactual aids. Our research will focus on development of a test tool for environmental-sound identification as well as on aspects of signal processing and display to enhance the reception of such sounds.

1.5.5 Publications

Besing, J.M., C.M. Reed, and N.I. Durlach. "A Comparison of Auditory and Tactual Presentation of a Single-Band Envelope Cue as a Supplement to Speechreading." *Seminars in Hear.* 16: 316-327 (1995).

Plant, G. "Training Approaches with Tactile Aids." *Seminars in Hear.* 16: 394-403 (1995).

Reed, C.M. "Tadoma: An Overview of Research." In *Profound Deafness and Speech Communication,* pp. 40-55. Eds. G. Plant and K.-E. Spens. London: Whurr Publishers, 1995.

Reed, C.M., L.A. Delhorne, N.I. Durlach, and S.D. Fischer. "A Study of the Tactual Reception of Sign Language." *J. Speech Hear. Res.* 38: 477-489 (1995).

Reed, C.M., and L.A. Delhorne. "Current Results of a Field Study of Adult Users of Tactile Aids." *Seminars in Hear.* 16: 305-315 (1995).

Tan, H.Z., and N.I. Durlach. "Manual Discrimination of Compliance Using Active Pinch Grasp: The Roles of Force and Work Cues." *Percept. Psychophys.* 57: 495-510 (1995).

Tan, H.Z., N.I. Durlach, W.M. Rabinowitz, C.M. Reed, and J.R. Santos. "Reception of Morse Code through Motional, Vibrotactile, and Auditory Stimulation." *Percept. Psychophys.* Under revision.

Thesis

Denesvich, G. *Identification of Frequency and Amplitude through Cutaneous Stimulation.* S.B. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1995.

1.6 Multimicrophone Hearing Aids

Sponsor

National Institutes of Health
Grant 2 R01 DC00270

Project Staff

Andrew R. Brughera, Joseph G. Desloge, Dr. Julie E. Greenberg, Michael P. O'Connell, Dr. William M. Rabinowitz, Coral D. Tassa, Dr. Patrick M. Zurek

The goal of this research is to determine the improvements that can be made to hearing aids through the use of multiple microphones. In this work, we are developing algorithms for processing the signals from a head-worn microphone array for the primary goal of improving the intelligibility of speech (assumed to arise from a known direction) in the presence of noise and reverberation. Ideally, this intelligibility enhancement would be achieved without compromising the listener's ability to monitor and localize sound sources from all directions. In computer simulations, array processing algorithms are first implemented and evaluated in terms of signal-to-noise improvement. The most promising approaches are then implemented in real-time with wearable devices (tethered to a computer) in the laboratory for evaluation of speech reception in noise and sound localization by normal-hearing and hearing-impaired listeners.

Array signal processing is usually designed to form a single output signal. In the hearing-aid application, however, this design has to be modified. Such a system would not contain the information about the location of sound sources that is normally conveyed primarily by differences between the signals at the two ears. In addition to sound localization ability, binaural hearing provides a sensation of auditory space and improved speech reception in noise. In an attempt to provide the natural benefits of binaural hearing along with improvements from microphone array processing, we have been exploring ways of combining them.

In one project,[23] we explored designs of a fixed four-microphone array in which, by choice of a frequency-dependent parameter, directivity can be traded with fidelity of interaural time delay (the most important binaural cue). At one extreme in this trade, maximal directivity is achieved, while at the other extreme, the outermost microphone signals are simply passed to the two ears. In another project,[24] we employed a two-microphone adaptive binaural array, with the microphones worn at the ears.

The signal-processing structure that provided the best combination of intelligibility enhancement, localization performance, and implementation simplicity was one that employed a special case of the trade between directivity and interaural delay. This system results from consideration of psychoacoustic findings showing the importance of low-frequency cues for the binaural tasks of detection, localization, and speech reception. To reflect this importance the threshold tolerance on interaural delay error is set to zero below some cutoff frequency f_{cut} and to infinity above f_{cut}. The resulting system extracts the low-frequency parts of the signals below f_{cut} from the microphones nearest the ears and passes them to their respective outputs. The high-frequency parts of the outputs come from the single-channel output of the array processor.

Measurements with normal-hearing subjects using several microphone-array configurations showed the expected trade between intelligibility and sound localization as the cutoff frequency is varied. With f_{cut} near 700 Hz, both good localization and substantial intelligibility enhancement from array processing can be obtained.

These studies of binaural-output microphone arrays have been strongly encouraging. They indicate that the spatial filtering provided by array processing can be obtained without sacrificing the benefits of binaural hearing. This is especially true of fixed-processing systems, which achieve relatively little directivity at low frequencies. The success of the lowpass/highpass system suggests that a merger of directivity and binaural hearing can be accomplished in a way that is very simple, general (at least for arrays that have microphones near the ears), and flexible. By adjusting f_{cut}, a user could vary the spatial properties of the system from a focused mode to an omnidirectional mode, depending on the listening condition.

1.6.1 Publications

Desloge, J.G., W.M. Rabinowitz, and P.M. Zurek. "Microphone-Array Hearing Aids with Binaural Output. I. Fixed-processing Systems." Submitted to *IEEE Trans. Speech Audio Proc.*

Welker, D.P., J.E. Greenberg, J.G. Desloge, and P.M. Zurek. "Microphone-Array Hearing Aids with Binaural Output. II. A Two-microphone Adaptive System." Submitted to *IEEE Trans. Speech Audio Proc.*

1.7 Hearing-Aid Device Development

Sponsor

National Institutes of Health
Contract N01 DC-5-2107

Project Staff

Merry A. Brantley, Lorraine A. Delhorne, Dr. Julie E. Greenberg, Michael A. Leabman, Dr. William M. Rabinowitz, Coral Tassa, Dr. Patrick M. Zurek

The overall objective of work under this new contract is to evaluate promising signal processing algorithms for hearing aids under realistic conditions. Progress toward this goal has begun with development of a laboratory-based signal-processing and testing system. Subsequent laboratory evaluations will select signal processing algorithms for implementation in wearable devices, which are being designed and built by Sensimetrics Corporation. Later in the project, field studies of hearing-impaired persons using this device will be conducted to evaluate the effectiveness of algorithms aimed at improving speech reception in background noise, preventing loudness discomfort, and increasing maximum gain without feedback.

1.8 Binaural Hearing

Sponsor

National Institutes of Health

[23] J.G. Desloge, W.M. Rabinowitz, and P.M. Zurek, "Microphone-Array Hearing Aids with Binaural Output. I. Fixed-Processing Systems," submitted to *IEEE Trans. Speech Audio Proc.*

[24] D.P. Welker, J.E. Greenberg, J.G. Desloge, and P.M. Zurek, "Microphone-Array Hearing Aids with Binaural Output. II. A Two-microphone Adaptive System," submitted to *IEEE Trans. Speech Audio Proc.*

Grant 2 R01 DC00100[25]

Project Staff

Nathaniel I. Durlach, Dr. Patrick M. Zurek

The long-term goal of this program is (1) to develop an integrated, quantitative theory of binaural interaction that is consistent with psychophysical and physiological data on normal and impaired auditory systems, and (2) to apply our results to the diagnosis and treatment of hearing impairments. Psychoacoustic studies are currently being conducted on time-dependent mechanisms in lateralization of transient sounds.

1.9 Virtual Environment Technology for Training

Sponsors

U.S. Navy - Office of Naval Research/
Naval Air Warfare Center
Contract N61339-94-C-0087
Contract N61339-95-K-0014
Grant N00014-93-1-1399
Grant N00014-94-1-1079

Project Staff

Walter A. Aviles, James H. Bandy, G. Lee Beauregard, Dr. David L. Brock, Erika N. Carmel, Nathaniel I. Durlach, Rakesh Gupta, Dorrie Hall, Professor Richard M. Held, Alexandra I. Hou, Dr. Lynette A. Jones, Dr. Corrie Lathan, Dr. Jeng-Feng Lee, Hugh B. Morgenbesser, Jonathan Pfautz, Nicholas Pioch, Dr. Wendelin L. Sachtler, Dr. J. Kenneth Salisbury, David W. Schloerb, Dr. Barbara G. Shinn-Cunningham, Dr. Mandayam A. Srinivasan, Lukasz A. Weber, Dr. Thomas E.v. Wiegand, Evan F. Wies, Dr. David Zeltzer

The long-term goal of the virtual environment technology for training (VETT) program is to increase the cost-effectiveness of training through the development and application of virtual-environment technology. The multimodal, interactive, and immersive characteristics of VE systems, combined with their software reconfigurability and ability to generate situations (realistic or intentionally unrealistic) that are optimally matched to a variety of training tasks, strongly suggest that they have great potential for cost-effective training.

1.9.1 Research Testbed

The research testbed efforts have centered on (1) the further development of the VE training tested and (2) the use of this testbed to conduct training experiments.

Efforts in the first area include (1a) the development of improved software for the design and execution of simulations, (1b) the creation of an experimenter's interface for facilitating experimental research on the testbed, and (1c) the incorporation of the Virtual Workbench (described in *RLE Progress Report Number 137*) to serve as a front-end to the testbed for cases in which the task to be trained, or the experimental research to be performed, can be appropriately conceptualized in terms of a workbench metaphor. Efforts in the second area, include (2a) the completion of a baseline simulation for the officer of the deck (OOD) training task, incorporating models for the terrain, water, objects in the water, and the piloting team with which the OOD must interact. In addition to the submarine's visual and dynamic models, other water-bound objects include, geometric models of buoys, range markers, and adventitious marine traffic. The second work area also includes (2b) preliminary evaluations of the OOD simulation by domain experts. Finally, the primary objective of these development efforts, i.e., (2c) training experiments that make use of VE simulation, are currently in progress.

In the following paragraphs we focus on items (1b) and (2a).

Creation of an Experimenter's Interface

The experimenter's interface (EI) was designed to be a generalized top layer of software for facilitating research in a variety of experimental contexts. Development of the EI was guided by a review of the desirable characteristics of experimental systems, some of which include: providing a simple interface to a complex system, ensuring a consistent experimental procedure, enabling a consistent interface despite an evolving set of experimental software and hardware, and limiting the amount of programming overhead and redundant coding involved in data collection.

From the software developer's point of view, the ongoing creation of many different experimental systems means the development of a significant amount of redundant code. Any piece of software

[25] Subcontract from Boston University. Professor H. Steven Colburn, Principal Investigator.

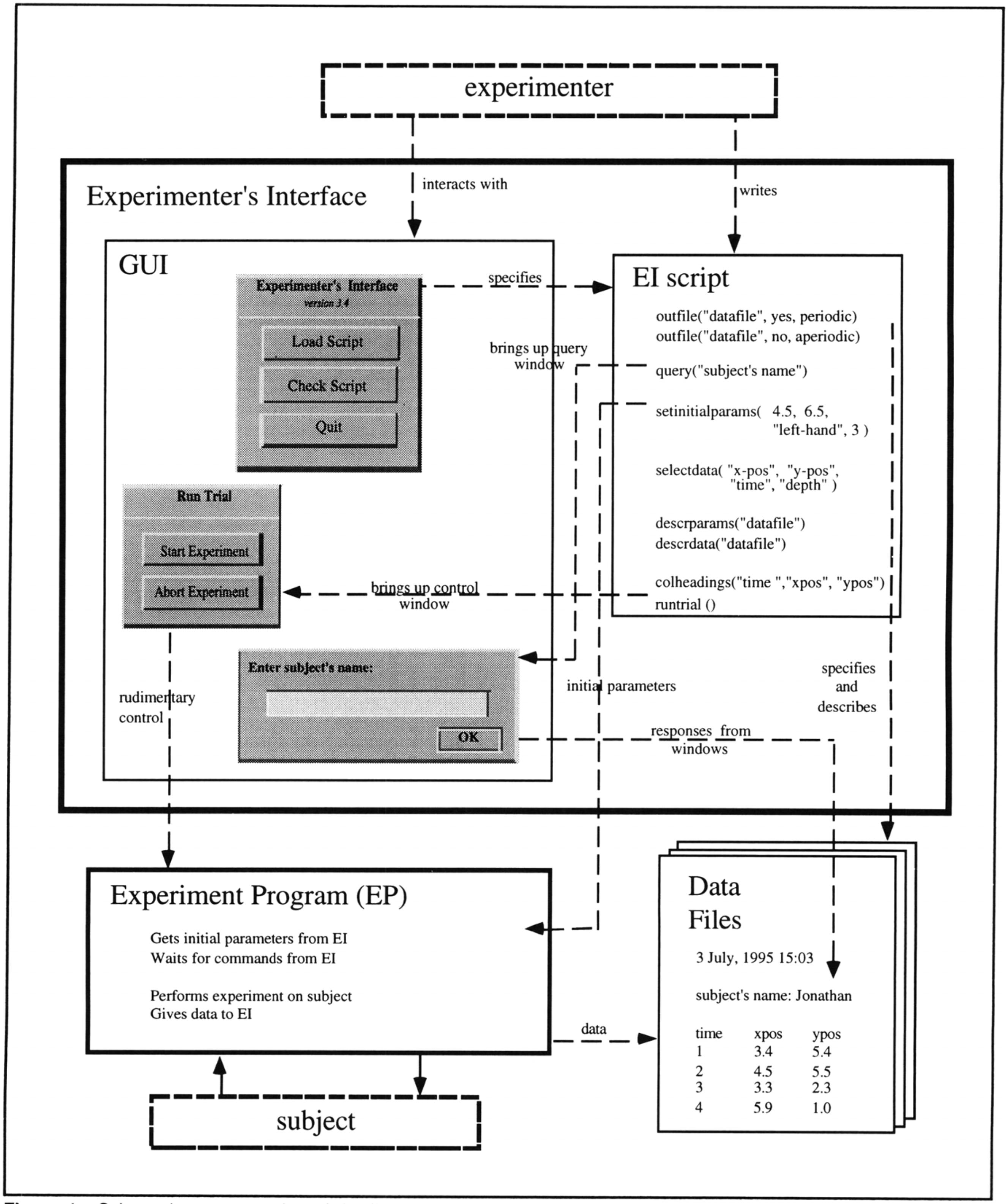

Figure 1. Schematic representation of the Experimenter's Interface. The Experimenter's Interface presents a uniform interface between the experimenter and the VE simulation that is presented to the subject, through the use of a standard set of script commands and a consistent Graphical User Interface. This approach facilitates the separation of simulation development from experimental design.

meant to drive an experiment has certain typical features. By using the EI for experiment specification, experiment control, and data collection, time and resources can be freed for other tasks. The EI also encourages lucid and consistent software design techniques since software engineers need to sufficiently modularize their code to efficiently utilize the EI interface. An additional advantage of using the EI is that more features will continue to be added over time, whereby it will evolve into a comprehensive top level interface for future experiments. Figure 1 depicts a schematic representation of the EI as currently used in the OOD experiments.

Completion of a Baseline Simulation for the OOD Task

In the OOD task, most information is received visually by observing the current configuration of shoreline and navigation aids relative to the submarine (see figure 2). Occasionally, verbal course information is sent by the chief navigator, and spoken acknowledgments of the OOD's commands are given by the helmsman. However, the OOD must be able to independently confirm the suggestions of the piloting team, and in addition, must be trained to perform this task independent of the piloting team in the event of on-board communication problems. Therefore, in the VE, the primary mode of output to the trainee is visual imagery, and spoken output to the trainee is secondary.

On the other hand, the OOD is only able to have an effect on the environment through a well-defined set of verbal commands that are issued to the helmsman. The implications for the VE system, then, are that the primary mode of input from the trainee is speech, which requires a speech recognition system and some degree of natural language understanding. Finally, tracking the head movements of the trainee is necessary to update the visual display from the appropriate point of view.

Creating a true-to-life model of every facet of this simulation in a single pass is an intractable undertaking. To facilitate software development, an evolutionary prototyping approach was employed in which a modular task analysis forms the framework from which incremental subtasks and added levels of detail can be derived. For example, it was decided that the first OOD simulator prototype would include the sequential task elements of centerline alignment, turning aid identification, and turn execution, while leaving secondary features such as water currents and moving traffic to be incorporated later.

Within these initial constraints, and in consultation with various domain experts, including submarine officers serving as instructors from the U.S. Navy Submarine School, the initial goals of the baseline simulation have been achieved. Overall, these officers showed considerable enthusiasm for the OOD simulator, including its potential for use as an onboard mission rehearsal system. We are currently in the process of analyzing results of the training experiments conducted with the completed baseline simulation, and are planning extensions to the experimental protocol.

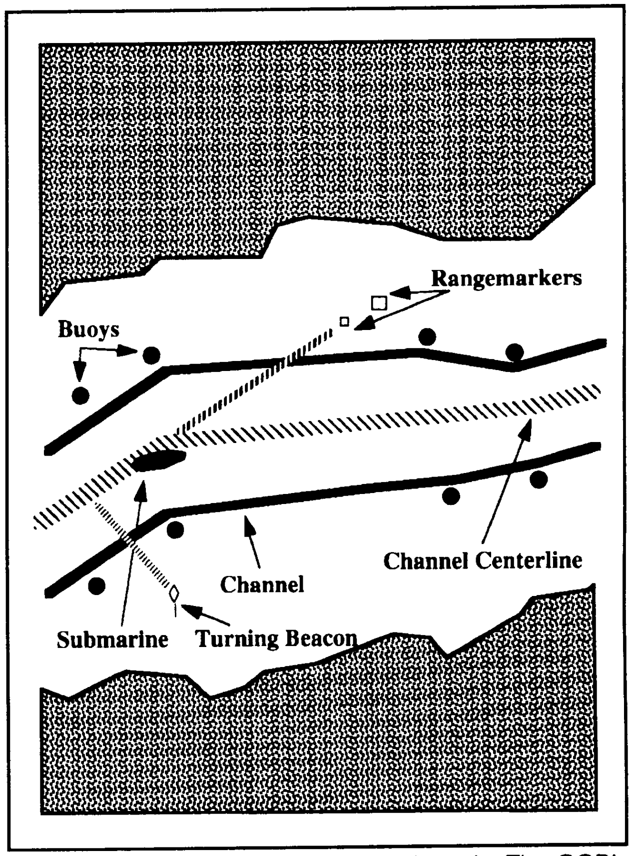

Figure 2. Map of channel to be navigated. The OOD's task in large measure is to verify the course of the ship as it negotiates the marked channel, using navigation aids such as buoys, which mark the channel boundaries, and rangemarkers, which are used to align the vessel with the channel centerline.

1.9.2 Sensorimotor Satellite

Work in the sensorimotor satellite involves research on human responses to alterations in sensorimotor loops associated with the use of virtual environment (VE) systems. Such alterations are likely to occur because of equipment limitations (resulting in unintended distortions, delays, or variability), or in some cases, because of deliberate attempts to emphasize or de-emphasize certain features of a simulation in order to improve training efficiency.

During the past year, we have continued a series of experiments concerned with altered relations between seen and felt hand position in the plane facing the user and initiated a new series of experiments related to the perception of depth in VEs. In the first series of experiments, we have examined the effects of rotation on both closed-loop and open-loop movements of the hand, as well as compared the effects of rotation on velocity control to those on position control. In the second series of experiments, we have begun to examine (1) stereopsis and the effects of altered interpupillary distance on depth perception in VEs, (2) the manner in which depth perception, achieved by means of motion parallax, degrades as temporal delay is increased, and (3) the effect of stereoscopic depth cues on brightness and contrast induction in VEs.

Most of these experiments were conducted using existing equipment such as the Sensorimotor Testbed and the Virtual Workbench, both previously described in *RLE Progress Report 137*. In order to perform the new motion parallax experiments, we have designed a device to passively limit an observer's head motion to a controlled maximum linear velocity. This system is constructed around a standard "bite bar" which can slide freely along a 20 cm track (see figure 3). In the experiments, the

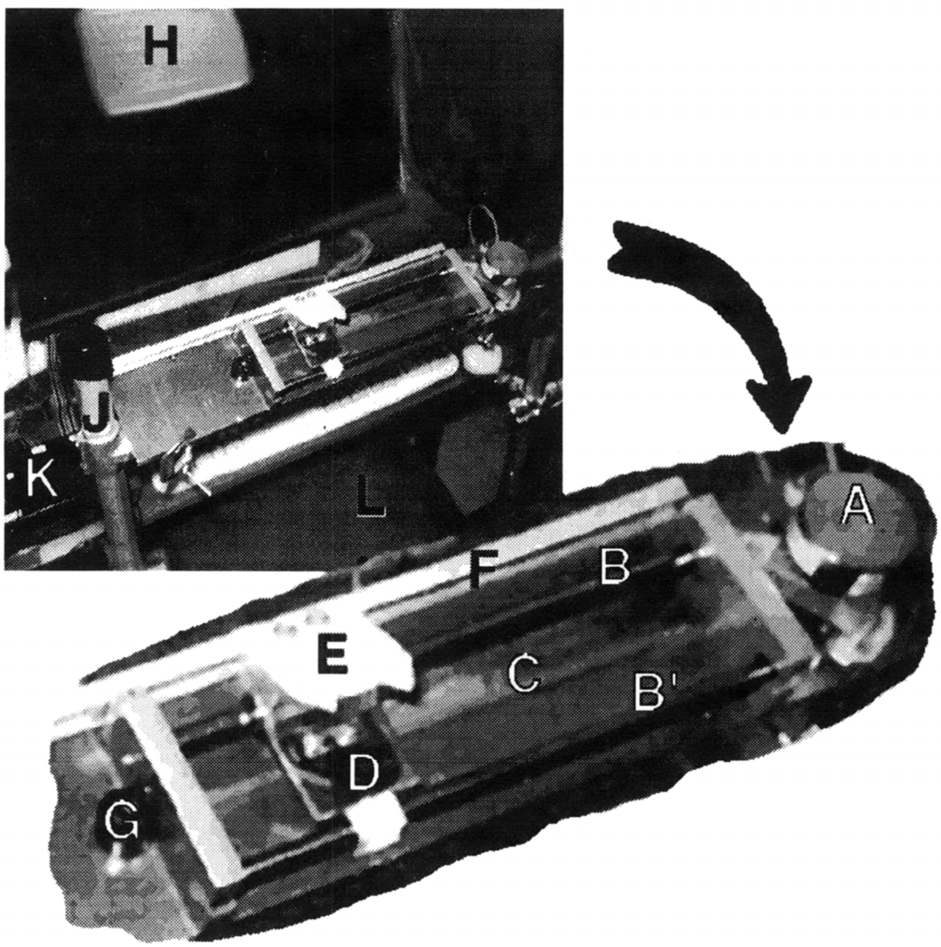

Figure 3. Motion Parallax Headguide. A: Magnetic particle brake; B: Guide rails; C: Toothed coupling belt; D: Carriage; E: Bite-piece; F: Precision linear potentiometer; G: Idler cog; H: CRT display screen; J: SpeedRail support framework; K: Interface control unit; L: Subject's seat.

observer's self-generated lateral motion can be limited through the frictional drag imparted by a servo-controlled magnetic particle brake. We are thus able to safely achieve changes in the velocity profile of the observer's viewpoint by completely avoiding components that are capable of generating motion on their own.

Subjects viewed a split-field random dot display, and their ability to identify the nearer field was measured. Changes in this ability were measured (discrimination thresholds) as a function of delay (between position measurement and display update), temporal quantization (of position readings and display writes), and spatial quantization (of position and display resolution). These measures have particular relevance with respect to limits on acceptable interprocess delays and the required resolution of VE tracking systems.

Figure 4 shows a plot of discrimination performance under various delay conditions. The data suggest that degradation of depth perception is relatively small for delays of less than 167 msec. A precipitous increase in discrimination threshold occurs when delay increases beyond 200 msec. Current VE systems commonly exhibit tracker-to-display delays of under 100 msec (e.g., an SGI Onyx with a Polhemus FASTRAK head tracker can achieve delays in the 67 msec range while displaying scenes of nontrivial complexity). Thus, in the absence of additional degradations of temporal and spatial quantization, high-end VE performance lies within an adequate range to support depth perception from motion parallax. We are currently mapping out the effects of delays and temporal and spatial quantization, singly and in combination.

In general, the results being obtained in our research are being used both to support development of predictive models of human performance and to determine requirements for cost-effective VE interfaces.

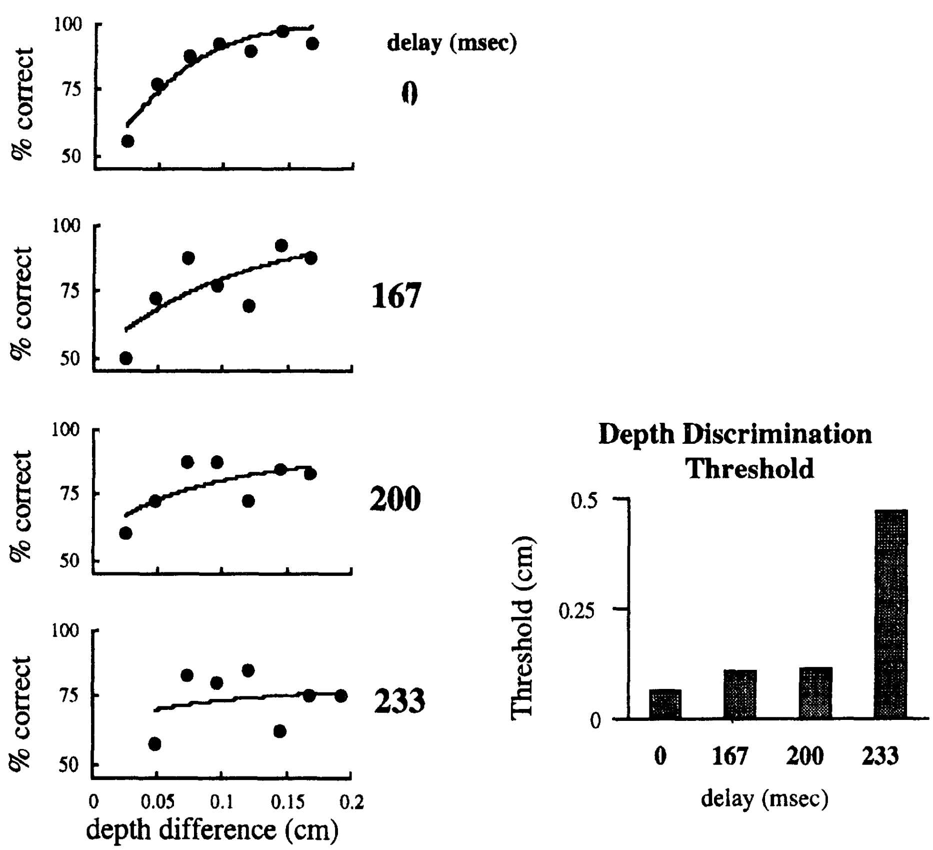

Figure 4. Results for four delay conditions. Left column: forty trials were run for each depth. The abscissa shows separation of the (virtual) surfaces in depth, while the percentage of trials on which the observer correctly indicated the location of the nearer surface is shown on the ordinate. Weibull functions fit to the data sets are shown as solid lines. Right column: thresholds for correct identification of the nearer surface, as estimated from the Weibull fits for the four different delay conditions.

1.9.3 Haptics Satellite

Over the past few years, we have developed device hardware and interaction software and performed psychophysical experiments pertaining to haptic interactions with virtual environments.[26] Two major devices for performing psychophysical experiments, the linear and planar graspers, have been fitted with additional sensors for improved performance. The linear grasper is now capable of simulating fundamental mechanical properties of objects such as compliance, viscosity, and mass during haptic interactions. Virtual wall and corner software algorithms were developed for the planar grasper, in addition to the simulation of two springs within its workspace. Another haptic display device developed previously, the PHANToM, has been used to prototype a wide range of force-based haptic display primitives. A variety of haptic rendering algorithms for displaying the shape, texture, and friction of solid surfaces have been implemented on the PHANToM. All the three devices have been used to perform psychophysical experiments aimed at characterizing the sensorimotor abilities of the human user and the effectiveness of computationally efficient rendering algorithms in conveying the desired object properties to the human user.

Human Sensorimotor Performance

We have completed a series of psychophysical experiments with the linear grasper designed to characterize human performance in the discrimination of fundamental mechanical properties of objects. These included several experiments that measured manual just noticeable difference (JND) in size, force, stiffness, viscosity, mass, velocity and acceleration.[27] We have also completed an analysis of the subjects' motor data during the discrimination of viscosity and mass. This analysis revealed that subjects used statistically the same initial temporal force profiles when manually discriminating these object properties. Based on this observation, a mathematical model has been developed to show that this motor performance gives rise to different spatial variations of force and motion, depending on the mechanical property of the object being squeezed. From this model, relationships have been derived that enable predictions for the manual resolution of object properties to be made solely on the basis of force discrimination. A comparison of these predicted JND values with the actual JND results has lead to the postulation of a single sensorimotor strategy capable of explaining the sensory and motor performance observed in all viscosity and mass experiments. This hypothesized strategy has been found not only to explain the observed motor and sensory data in these discrimination experiments, but it is also consistent with our previous work on the manual resolution of mechanical compliance.[28]

Force Shading Algorithm

We have developed novel algorithms for the haptic display of shape, texture, and friction of solid surfaces using the PHANToM.[29] We have also performed psychophysical experiments to determine the effectiveness of such rendering algorithms in conveying the desired object properties to the human user. The results show that even when the user is exploring nominally flat surfaces, he can be made to feel as if he is interacting with a shaped and textured surface by appropriately varying the direction of the force reflected back to the user. These new algorithms, called "force shading," permit the mapping of a shape or texture onto a polygon, so that they may be used in haptic rendering in the same way that texture mapping and color shading are used in graphics rendering.

We have performed psychophysical experiments to determine the simplest polyhedral representation that is acceptable to the user as a smooth continuous surface. The PHANToM, which is a high performance haptic interface device developed as a

---

[26] M.A. Srinivasan, "Virtual Haptic Environments: Facts Behind the Fiction," *Proceedings of the Eighth Yale Workshop on Adaptive and Learning Systems,* Center for Systems Science, Yale University, New Haven, June 1994; M.A. Srinivasan, "Haptic Interfaces," in *Virtual Reality Scientific and Technical Challenges,* eds. N.I. Durlach and A.S. Mavor, Report of the Committee on Virtual Reality Research and Development, National Research Council (Washington, D.C.: National Academy Press, 1995); M.A. Srinivasan, "Haptic Interfaces: Hardware, Software, and Human Performance," *Proceedings of the NASA Workshop on Human-computer Interaction and Virtual Environments,* NASA Conference Publication 3320: 103-121 (1995).

[27] G.L. Beauregard, M.A. Srinivasan, and N.I. Durlach, "Manual Resolution of Viscosity and Mass," *Proceedings of the ASME Dynamic Systems and Control Division* DSC-Vol. 57-2: 657-662 (New York: ASME, 1995).

[28] H.Z. Tan, N.I. Durlach, G.L. Beauregard, and M.A. Srinivasan, "Manual Discrimination of Compliance Using Active Pinch Grasp: the Roles of Force and Work Cues," *Percept. and Psychophys.* 57(4): 495-510 (1995).

[29] H.B. Morgenbesser, *Force Shading for Shape Perception in Haptic Virtual Environments,* M.Eng. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1995.

part of the VETT project, was utilized for the experiments. It is capable of reflecting back three-dimensional force vectors within a three-dimensional workspace. In the first set of experiments in each trial, the user is asked to stroke the stylus of the PHANToM over a continuous cylindrical surface and a polyhedral approximation to it (both are virtual surfaces). The polyhedral surface is associated with a force-shading scheme such that the subject can vary the effective average curvature of the polyhedral surface. The subject is asked to modify the force-shaded polyhedron until it perceptually matches the feel of the continuum surface. These experiments have been conducted for polyhedra with one to three faces approximating cylinders of several radii of curvature. The results show that the force-shading algorithm is successful in creating the illusion of a continuous surface and that the polyhedral surface needs to have a higher average curvature than that of the continuous cylinder that is its perceptual equivalent. In the second set of experiments, subjects manually explored the haptic display of various virtual polyhedral approximations of a circular cylindrical bump, with or without force shading. They then indicated the perceived shape by selecting one from a menu of shapes displayed on a computer monitor. Without force shading, the subjects accurately identified the polyhedral nature of the virtual bumps. With force shading, however, the subjects identified the polyhedral approximations as feeling closer to that of a smooth cylinder. The higher the number of polygons, the higher the percentage of trials in which the shape was identified with a smooth cylinder. The results from both experiments imply that the same polyhedral approximations to object shape used in visual displays can be used for haptic displays, provided the haptic renderer does the appropriate force-shading to make it feel continuous and smooth. Therefore, one can envision a data architecture for multimodal VEs where the core geometric information about object shapes relevant for the specific virtual environment of interest is independent of visual or haptic display modalities.

Visual-Haptic Interactions

Psychophysical experiments were also conducted with the planar grasper to investigate the influence of visual information on the perception of stiffness of virtual springs. In each trial, the subjects could manually press and feel the stiffness of two springs programmed to be side-by-side within the workspace of the planar grasper. At the same time, visual images of the springs were displayed on a computer monitor such that the amount of deformation of the spring that was pressed was displayed in real-time. The subjects were asked to judge which of the two springs was stiffer. The results showed that both in the absence of visual information (when the monitor was turned off) as well as when the visual displacement was exactly equal to the haptic displacement, the subjects were almost 100 percent correct in their judgements. However, when the visual displacement of the two springs were interchanged (i.e., for a given force applied by the subject, the visual displacement of the stiffer spring was equal to the haptic displacement of the softer spring and vice versa) the subjects were almost 100 percent wrong in their judgments. These and other experiments involving different scaling values between the visual and haptic displacements showed that in computing the stiffness of objects, when subjects have a choice of associating visual or haptic displacements with their haptic perception of force, they consistently choose the visual displacements, i.e., the visual information dominates over kinesthetic information in this context. An important implication for virtual environments is that by skewing the relationship between the haptic and visual displays, the range of object properties that can be effectively conveyed to the user can be significantly enhanced. For example, although the range of object stiffnesses that can be displayed by a haptic interface is limited by the force-bandwidth of the interface, the range perceived by the subject can be effectively increased by reducing or eliminating visual deformation of the object.

1.10 Training for Remote Sensing and Manipulation

Sponsor

U.S. Navy - Office of Naval Research
Subcontract 40167[30]

Project Staff

Walter A. Aviles, Nathaniel I. Durlach, Francis G. Taylor, John Yoon, Dr. David Zeltzer

This project is concerned with a special type of remotely operated vehicle (ROV), namely, teleoperated minisubmarines. A few thousand such vehicles are currently operating around the world. They are used by oil companies to inspect and maintain offshore oil platforms, by various civil authorities to inspect underwater structures such as dams and sewage outflow pipelines, and by the U.S. Navy for undersea salvage and submerged mine counter measures (MCM).

The objective of the training for remote sensing and manipulation (TRANSOM) research program is to design, develop, and evaluate a prototype ROV pilot training and mission rehearsal system for a shallow-water MCM mission incorporating intelligent tutoring system (ITS) techniques within a virtual environment (VE). Such a system will benefit the efficiency of MCM operations, resulting in the preservation of life and property as well as significant cost savings in the training of operators. In addition, it is foreseen that the results of the proposed program will have applicability to ROV missions outside the MCM scenario and to other remote sensing and manipulation systems.

Five organizations are collaborating on this effort. Imetrix Corporation, Cataumet, Massachusetts, is the prime contractor, overseeing the overall effort, providing the ROV and control expertise, and serving as the main conduit for product transition. Our research group is providing the lead in human/machine interface research and VE technology. McDonnell Douglas Training Systems (MDTS), St. Louis, Missouri, is providing expertise in ITS design and leading the training transfer studies. Bolt Beranek and Newman, Inc., Cambridge MA, is developing the ITS and providing training research support. Finally, LRDC, Pittsburgh PA, is providing training expertise in team interaction.

Work in the Sensory Communication Group is being conducted in three areas. This work is being implemented on SGI Indys acquired specifically for the TRANSOM program. We also make use of the SGI ONYX acquired for the VETT program.

(1) We have developed simulations of the ROV, the ocean bottom over which it operates, and the ROV control console. The ROV geometry model is based in specifications provided by Imetrix. ROV dynamics are simulated using a package developed at Imetrix and adapted for use in the VE developed at MIT. Validation and verification of this simulation software is currently getting underway.

(2) We have implemented a "developers' interface" that provides a rapid-prototyping environment for the development of intelligent VE tutoring components. This uses a Scheme language "front-end" to the SGI open Inventor Language.

(3) Prototypical part-task teaching components of the tutor have been integrated with the simulated ROV, the ROV console, and ocean elements as part of a joint effort with BBN and MDTS to identify effective training interventions.

1.11 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, Dr. Robert H. LaMotte,[31] Dr. Robert D. Howe,[32] Jyh-Shing Chen, Kiran Dandekar, Rogeve J. Gulati, Louise Jandura, Steingrimur P. Karason, Frederick L. Roby, Sudeep Rangaswamy

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 side we wish to understand

---

30 Imetrix Corporation.

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

32 Harvard University, Cambridge, Massachusetts.

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 humans bring to bear on task performance we seek guidelines on how to better build robots. From the robotic side we wish to understand how mechanism and sensor design choices can best be made to maximize grasping and manipulative competences. By better understanding the mechanical demands of task performance we seek to understand the performance demands which underlie skillful human manipulation.

The main components of the research conducted under this project are (1) development of new hardware for robotic and human studies, (2) processing of robot sensor signals and task-level control of the devices, (3) investigation of the mechanics of human fingerpad-object contact, (4) experiments on human perception and control of forces using some of the devices, and (5) development of a computational theory of haptics that is applicable to both robots and humans. The subsections to follow provide descriptions of the results obtained mainly by the Human Haptics Group at the Touch Lab in RLE.

1.11.1 Biomechanics of Human Fingerpad-Object Contact

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. 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 goals of this research are (1) to determine the growth and motion of contact regions and the associated force variations over time between the human fingerpad and carefully chosen transparent test objects whose microtexture, shape or softness is varied in a controlled manner, (2) Experimental measurement of the surface deformations of human fingertips under shaped indentors, and (3) Characterization of the mechanical properties of the human fingerpad. The results obtained are being used to gain a deeper understanding of the neurophysiological and psychophysical data we have already obtained for the same test objects.

To measure the *in vivo* surface deformations of the fingerpad under various tactile stimuli, we have designed a videomicroscopy system together with a high precision tactile stimulator. The videomicroscopy system consists of a set of video zoom lenses attached to a high-resolution CCD camera, whose output can either be digitized into the computer system memory in real time at about 20 frames/s, or stored on a laserdisk at 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. The tactile stimulator is composed of a linear stepper motor with a microstepping drive. The motor is controlled by a 80386 PC, with a specified indentation velocity commanded by a 80486 PC via a digital link. To record the contact force, a strain gage based single degree of freedom force sensor is mounted on the motor to which a transparent test object can be attached for both biomechanical and psychophysical experiments. This method allows the force and video data to be synchornized and stored in the 80486 PC. With this setup, we are able to investigate how the skin-object contact region changes with indentation velocity and force. In active touch experiments the subject contacts a stationary specimen, whereas in passive touch experiments the stimulator moves the specimen to indent a stationary finger at a given velocity. High contrast images of the contact interface are achieved with coaxial and other fiberoptic lighting.

Videomicroscopy of the Fingerpad-object Contact Regions

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 non-uniform 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 non-contact 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 previous psychophysical, biomechanical and neurophysiological experiments.[33]

Measurement of Surface Deformation of Human Fingerpads

The finite element models described previously need to be verified by comparing the experimentally observed skin surface deformations with those predicted by the finite element models under the same mechanical stimuli. The experimental data was obtained by indenting human fingerpads with several cylindrical and rectangular indentors and acquiring images of the undeformed and deformed fingerpad using the videomicroscopy setup.[34] Fine markers were placed on the fingerpad and the skin surface deformation was measured by tracking the displacements of the markers in the high resolution video images. The same experiment was simulated using the finite element models of the human fingertip and the displacements of corresponding points were compared with the experimental data. The displacements predicted by the multilayered 3D model matched the experimental data quite well.[35]

Force Response of the Human Fingerpad to Indentation

A 2-DOF robot designed by Dr. Howe of the Harvard group was modified to serve as a "Tactile Stimulator" capable of delivering static and dynamic stimuli to the human fingerpad.[36]

Three types of indentors (point, flat circular, and a flat plate) attached to the stimulator imposed a variety of constant velocity ramps (1 to 32 mm/s), depths of indentation (0.5 to 3mm), and sinusoids (0.25 mm to 0.5 mm amplitude; 0.125 to 16 Hz frequency) under displacement control (resolution ~ 20 microns). The resulting normal and shear forces were measured by a 2-axis force sensor (resolution ~ 10 mN). The results showed a pronounced nonlinear force-indentation depth relationship under both static and dynamic conditions, viscoelastic effects of force relaxation under constant depth of indentation, and hysteresis under sinusoidal displacements. There was wide variability in the magnitude of response for the five subjects who were tested, and their fingertip diameter or volume did not account for the observed variability.

A piecewise linear, lumped parameter model with spring and dashpot elements was developed to identify the mechanical parameters causing the nonlinear response. The model predictions with only one set of parameters for each subject matched the empirical data well under a wide variety of stimuli. The model represents a compact description of the data and will be used to verify and tune our finite element models of the fingertip.

Force Response of the Human Fingerpad to Shear Displacement

The 2-DOF Tactile Stimulator was also used to deliver shear displacement ramp (0.5 to 16 mm/sec for a total shear displacement of 7 mm) at various depths of indentation of the fingerpad (0.5 to 3 mm) by a flat, smooth Aluminum plate.[37] Only one subject has been tested so far under these stimuli, and the resulting data has been analyzed with a view towards fine tuning the experimental protocol and parameters. The results show that at each depth of indentation, the shear displacement initially caused increasing skin stretch and shear force, fol-

---

33 M.A. Srinivasan and R.H. LaMotte, "Tactual Discrimination of Softness," *J. Neurophys.* 73(1): 88-101 (1995); M.A. Srinivasan and R.H. LaMotte, "Tactual Discrimination of Softness: Abilities and Mechanisms," in *Information Processing in the Somatosensory System,* eds. O. Franzen, R. Johansson, and L. Terenius. (Basel, Switzerland: Birkhauser Verlag AB, forthcoming).

34 F.L. Roby, K. Dandekar, and M.A. Srinivasan, "Study of Fingertip Deformation Under Indentations by Circular and Rectangular Indentors," Report to the MIT Summer Research Program, 1994; F.L. Roby and M.A. Srinivasan, "Study of Fingertip Deformations Under Indentations by Circular Indentors," Report to the MIT Summer Research Program, 1995.

35 K. Dandekar, *Role of Mechanics in Tactile Sensing of Shape,* Ph.D. diss., Dept. of Mech. Eng., MIT, 1995.

36 R.J. Gulati, *Determination of Mechanical Properties of the Human Fingerpad, in vivo, using a Tactile Stimulator,* M.S. thesis, Dept. of Biomedical Eng., Boston University, 1995; R.J. Gulati and M.A. Srinivasan, "Human Fingerpad Under Indentation I: Static and Dynamic Force Response," *Proceedings of the 1995 Bioengineering Conference,* eds. R.M. Hochmuth, N.A. Langrana, and M.S. Hefzy, BED-29: 261-262 (1995).

37 J.D. Towles and M.A. Srinivasan, "Frictional Properties of the Human Fingerpad," Report to the MIT Summer Research Program, 1994.

lowed by slipping of the plate across the skin surface. The shear force-shear displacement was almost linear and slip occurred at around 3 mm shear displacement at all velocities. Low velocities tended to cause stick-slip (as indicated by oscillatory shear force during slip), whereas the shear force decreased smoothly at higher velocities. At increasing depths of indentation, slip occurred at larger shear displacements, as is to be expected. The coefficient of static friction was obtained by measuring the slope of the normal and shear forces at the incipience of slip for a given shear velocity. To a first approximation it was found to be independent of shear velocity. More experiments on different subjects are being initiated.

1.11.2 Human Perception and Control

To quantify human perception and control of forces, we have conducted experiments under a wide variety of conditions: (1) Tracking of visual displays of static and dynamic force traces with a stationary fingerpad (isometric case), (2) Maintaining constant contact force on a moving robot end-effector (isotonic case), (3) sensing and control of torque applied on the shaft of an "Instrumented Screw Driver", and (4) Control of grasp forces on an "Active Instrumented Object". The data from each of these experiments have been analyzed with the viewpoint of developing engineering specifications of human haptic performance. Brief descriptions of each of the experiments and models described above are given in the sections below.

Isometric Force Tracking Ability of Humans

In the experimental setup, a human subject tracks visual images of force traces displayed on the monitor by applying appropriate normal forces through a fingerpad that is in contact with a force sensor.[38] In these experiments, the finger moves by only fractions of a mm, thus approximating isometric muscular contraction conditions. We had completed one set of experiments on such force tracking last year, and now we have expanded the range of target forces and their frequency. During tracking constant and sinusoidal force targets (three subjects with 3 trials per stimulus), the mean absolute error increased with constant force magnitude, target sinusoid frequency and amplitude. The errors for a sinusoid of a given amplitude are 5 to 40 times higher than those for constant force targets with the same magnitude. Even at relatively low frequency of 2 Hz, the errors can be higher than 50 percent of the sinusoid amplitude at all amplitudes.

Isotonic Force Control Ability of Humans

We employed the high precision "Glass Smooth Actuator" to measure the ability of human subjects to maintain a constant force (0.1 to 0.8N) by pressing their fingerpads on the actuator's end effector while it was moving sinusoidally (2 to 16 degrees in amplitude; 0.5 to 16 Hz frequency).

During each trial, the robot maintained a constant position for the first 10 seconds, and during the first 8 seconds, the subjects tracked a constant target force displayed as a line on a monitor. The monitor screen was then blanked out, and after 2 seconds, the actuator started moving sinusoidally at a preprogrammed frequency and amplitude, but the subjects were asked to maintain the same force as before. All the subjects were able to perform the task with very little drift in mean force. However, the deviations from the mean were in phase with the actuator motion, and various error measures changed differently with stimulus parameters. The mean absolute error increased with frequency and amplitude almost linearly, but remained constant with respect to target force magnitude.

Torque Sensing and Control

The human ability to sense and control torque was investigated in experiments with the Instrumented ScrewDriver (ISD).[39] The ISD is comprised of a single shaft, which is supported by low friction bearings, and 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 all cases the subjects grasped the handle of the ISD between the thumb and index finger of their dominant hand and turned the shaft clockwise for 180 degrees against a constant resistive torque applied by the magnetic particle brake. The magnitude of this resistive torque was varied across different trials. Two types of experiments were conducted: discrimination experiments to determine the human resolution in sensing

[38] 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-49 (New York: ASME, 1993).

[39] L. Jandura and M.A. Srinivasan, "Experiments on Human Performance in Torque Discrimination and Control," in *Dynamic Systems and Control,* Vol. 1, ed. C.J. Radcliffe, DSC-55-1 (New York: ASME, 1994).

torque and control experiments to determine the human motor capability in controlling torque.

All torque discrimination experiments used a one-interval, two-alternative, forced-choice paradigm with no feedback to the subject. The reference torque value was 60 mN-m and the comparison values were equal to 5, 10, 20, and 30 percent of the reference torque. In addition, training runs were conducted with a comparison value of 50 percent of the reference torque until the subject response was 90 percent correct. The Just Noticeable Difference for torque was found to be 12.7 percent for the reference torque of 60 mN-m. During some of the trials, in addition to recording the stimulus and the subject's response, the resistive torque, the output of the torque sensor and the angular position of the shaft over time were also recorded. These data are used to make comparisons between the motor performance in the discrimination task and the control task.

For the control experiments, subjects were asked to maintain a constant angular velocity while turning against the constant resistive torque. The value of the angular velocity was up to the subject to choose, but they were asked to try and use the same value for each trial. Because of the physics of the ISD, attempting to maintain a constant angular velocity is directly related to attempting to apply and maintain a constant torque during shaft motion. The constant resistive torque values used were the same as for the discrimination experiments. As before, the resistive torque, the output of the torque sensor, and the angular position of the shaft were recorded over time. Comparison of the time profiles of angular velocity indicate that even when subjects were trying to maintain a constant angular velocity in the control experiments, their performance was not significantly better than when they were trying to discriminate the torques.

A curious phenomenon observed rather consistently in all of the data is the occurrence of peaks in the velocity and acceleration profiles at about 0.1 second intervals. To further investigate this observation, the power spectral density of the middle third of the angular velocity profile was calculated. Although there is some tendency for the discrimination PSDs to be single-peaked while the control PSDs are double-peaked, this was not observed consistently across all subjects. However, in all subjects, most of the frequency content was less than about 15 Hz for both the discrimination and control experiments.

Passive Grasp Control of an Active Instrumented Object

Most of the investigations of human grasp control reported in the literature pertain to passive objects. In order to test the limitations of the human motor system in compensating for sudden disturbances, we fabricated an "Active Instrumented Object".[40] It mainly consisted of a pneumatic cylinder whose piston could expand or contract through the operation of computer controlled valves. A position sensor monitored the motion of the piston while two 2-axis force sensors measured the normal and shear forces applied by a subject's fingerpads on two circular plates when the object was held in a pinch grasp. The plate surfaces were either polished aluminum or sandpaper depending on the experiment. A visual display on a computer monitor indicated the force of pinch grasp the subjects were required to apply, and when subjects achieved it, after a random time-delay the object contracted suddenly (170 to 240 mm/s velocity; 19 mm displacement). When initial grasp forces were less than 1.8 N, the subjects dropped the object all the time (3 subjects; 10 trials per grasp force value per subject), whereas they held it in increasing nuber trials as the grasp forces increased. No significant difference between aluminum and sand paper surfaces were observed. In one trial where there was a small slip, a perturbation in force applied by the thumb is seen at around 100 ms, but the object did not drop. In another trial where the object was dropped, oscillations in force applied by the middle finger began at 60 ms and those in force applied by the thumb began at about 80 ms. Both these oscillations continued until the object was dropped at about 100 ms.

A third order lumped parameter model of the finger was proposed and the parameters of the model corresponding to the stiffness and damping of the fingerpad and muscle together with finger mass were identified using the recorded force and displacement data. This model of the finger was shown to predict the experimental data quite well, indicating that the subjects' fingers are effectively passive during rapid haptic events that occur within 60 to 80 ms. This model also related the grasping performance of each subject to the strategy the subject used in the stiffness modulation of the finger muscles before each trial. Furthermore, the

[40] S.P. Karason and M.A. Srinivasan, "Passive Human Grasp Control of an Active Instrumented Object," *Proceedings of the ASME Dynamic Systems and Control Division,* DSC-57-2: 641-647, (New York: ASME, 1995).

identified parameters governing the passive dynamic properties of the fingers provide a starting point for more complex models which account for the active nature of the finger muscles.

1.11.3 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. Therefore the mechanistic analyses needed to solve 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) are sufficiently similar for human and robot tactile sensing systems that a theory common to both systems can be developed.

We first developed such a "computational theory" using a simplified 2D half-space model of the human or robot finger subjected to arbitrary pressure or displacement loading conditions normal to the surface, and gave explicit formulae for the coding and decoding problems.[41] We have now expanded these results to a more general 3D half-space model where the load direction can be completely arbitrary.[42] Explicit solutions for the coding problem are given and enable the selection of a useful set of relevant stimuli as well as the choice of sensors appropriate for maximizing the information about the stimulus on the skin surface. The solution of the decoding problem is also given, both for the idealized noise-free case and for the realistic case with measurement noise. For the latter, the solutions are shown to be numerically stable and optimal.

In our work during the previous years, we were successful in answering basic identification and control issues that arise during manipulation of compliant objects using compliant fingerpads.[43] In order 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 modelled the dynamic interactions that occur between compliant end-effectors and deformable objects by a class of nonlinear systems.[44] It was shown that standard geometric techniques for exact feedback linearization techniques were inadequate. New algorithms were developed by using adaptive feedback techniques which judiciously employed the stability characteristics of the underlying nonlinear dynamics. In both theoretical and simulation studies, it was shown 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.

1.11.4 Publications

Annaswamy, A.M., and M.A. Srinivasan. "The Role of Compliant Fingerpads in Grasping and Manipulation: Identification and Control." Chapter in a book to be published by the Institute of Mathematics. New York: Springer Verlag, forthcoming.

Beauregard, G.L., M.A. Srinivasan, and N.I. Durlach. "Manual Resolution of Viscosity and Mass." *Proceedings of the ASME Dynamic*

[41] M.A. Srinivasan, "Tactile Sensing in Humans and Robots: Computational Theory and Algorithms," *Newman Laboratory Technical Report,* Department of Mechanical Engineering, MIT, 1988.

[42] S.P. Karason, M.A. Srinivasan, and A.M. Annaswamy, "Tactile Sensing of Shape," *Proceedings of the Workshop Organized by the Center for Information Driven Mechanical Systems (CIDMS),* Dept. of Mechanical Engineering, MIT, 1994.

[43] A.M. Annaswamy and M.A. Srinivasan, "A Study of Dynamic Interactions between Haptic Interfaces and Compliant Fingerpads," *Proceedings of the Motion Control Workshop,* Berkley, California, March 1994; A.M. Annaswamy and M.A. Srinivasan, "The Role of Compliant Fingerpads in Grasping and Manipulation: Identification and Control," Chapter in a book to be published by the Institute of Mathematics (New York: Springer Verlag, forthcoming).

[44] A.M. Annaswamy and D. Seto, "Object Manipulation Using Compliant Fingerpads: Modeling and Control," *ASME J. Dynamic Syst., Measure., Cont.* (1993).

Systems and Control Division. DSC-Vol. 57-2: 657-662. New York: ASME, 1995.

Dandekar, K., and M.A. Srinivasan. "A 3-dimensional Finite Element Model of the Monkey Fingertip for Predicting Responses of Slowly Adapting Mechanoreceptors." *Proceedings of the 1995 Bioengineering Conference.* Eds. R.M. Hochmuth, N.A. Langrana, and M.S. Hefzy. BED-29: 257-258 (1995).

Gulati, R.J., and M.A. Srinivasan. "Human Fingerpad Under Indentation I: Static and Dynamic Force Response." *Proceedings of the 1995 Bioengineering Conference.* Eds. R.M. Hochmuth, N.A. Langrana, and M.S. Hefzy. BED-29: 261-262 (1995).

Karason, S.P., and M.A. Srinivasan. "Passive Human Grasp Control of an Active Instrumented Object." *Proceedings of the ASME Dynamic Systems and Control Division.* DSC-57-2: 641-647. New York: ASME, 1995.

LaMotte, R.H., C. Lu, and M.A. Srinivasan. "Peripheral Neural Representation of the Shapes and Orientations of Three-dimensional Objects Stroked Across the Monkey Fingerpad." *Soc. Neurosci. Abstr.* 21: 1162 (1995).

LaMotte, R.H., C. Lu, and M.A. Srinivasan. "Tactile Neural Codes for Shapes and Orientations of Objects." In *Information Processing in the Somatosensory System.* Eds. O. Franzen, R. Johansson, and L. Terenius. Basel, Switzerland: Birkhauser Verlag AB. Forthcoming.

Roby, F.L., and M.A. Srinivasan. "Study of Fingertip Deformations Under Indentations by Circular Indentors." Report to the MIT Summer Research Program, 1995.

Srinivasan, M.A. "Haptic Interfaces." In *Virtual Reality Scientific and Technical Challenges.* Eds. N.I. Durlach and A.S. Mavor. Report of the Committee on Virtual Reality Research and Development. Washington, D.C.: National Academy Press, 1995.

Srinivasan, M.A. "Haptic Interfaces: Hardware, Software, and Human Performance." *Proceedings of the NASA Workshop on Human-computer Interaction and Virtual Environments*, NASA Conference Publication 3320: 103-121 (1995).

Srinivasan, M.A., and K. Dandekar. "An Investigation of the Mechanics of Tactile Sense Using Two Dimensional Models of the Primate Fingertip." *J. Biomech. Eng.* 118: 48-55 (1996).

Srinivasan, M.A., and R.H. LaMotte. "Tactual Discrimination of Softness." *J. Neurophys.* 73(1): 88-101 (1995).

Srinivasan, M.A., and R.H. LaMotte. "Tactual Discrimination of Softness: Abilities and Mechanisms." In *Information Processing in the Somatosensory System.* Eds. O. Franzen, R. Johansson, and L. Terenius. Basel, Switzerland: Birkhauser Verlag AB. Forthcoming.

Tan, H.Z., N.I. Durlach, G.L. Beauregard, and M.A. Srinivasan. "Manual Discrimination of Compliance Using Active Pinch Grasp: the Roles of Force and Work Cues." *Percept. Psychophys.* 57(4): 495-510 (1995).

Theses

Dandekar, K. *Role of Mechanics in Tactile Sensing of Shape.* Ph.D. diss., Dept. of Mech. Eng., MIT, 1995.

Gulati, R.J. *Determination of Mechanical Properties of the Human Fingerpad, in vivo, Using a Tactile Stimulator.* S.M. thesis. Dept. of Biomedical Eng., Boston University, 1995.

Morgenbesser, H.B. *Force Shading for Shape Perception in Haptic Virtual Environments.* M.Eng. thesis. Dept. of Electr. Eng. and Comput. Sci., MIT, 1995.

1.12 Role of Skin Biomechanics in Mechanoreceptor Response

Sponsor

National Institutes of Health
Grant R01-NS33778

Project Staff

Dr. Mandayam A. Srinivasan, Kiran Dandekar, Balasundara I. Raju, Kimberly J. Voss, Suvranu De

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 mechanoreceptive 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 generation of 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 skin and subcutaneous tissues, and 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.

1.12.1 Determination of the Shape and Compressibility 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. The three-dimensional (3D) external geometry of the primate fingertips 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.[45] We have reconstructed several human and monkey fingertips by using this method.

For mechanistic 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.[46] 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 also developed an experimental setup involving a computer controlled linear actuator for fingertip volume change measurements under dynamic conditions.[47] The results show that reductions in fingertip volume are in phase with stimulus variations, with an increase in their mean value over time. The volume changes during the ramp phase increase linearly with indentor displacement and are independent of velocity; during sawtooth stimulations, however, the nature of the hysteresis loops depend on velocity of indentation.

1.12.2 Fingertip Models and Finite Element Analysis

We have performed linear and nonlinear finite element analysis of a series of mechanistic models of the fingerpad under a variety of mechanical stimuli.[48]

The models range from a semi-infinite medium to a 3D model based on the actual finger geometry, and composed of a homogeneous elastic material, a thick elastic shell containing a fluid or a multilayered medium. 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 and 3D models. We have also solved the nonlinear contact problem of indentations by cylindrical objects and sinusoidal step shapes. The large number of numerical calculations needed even for the linear two dimensional models necessitated the use of the Cray-C90 at the NSF Pittsburgh Supercomputer Center.

---

45 T.R.R. Perez, K. Dandekar, and M.A. Srinivasan, "Videomicroscopic Reconstruction of the Human Finger," Project report to the MIT Summer Science Research Program, 1992.

46 M.A. Srinivasan, R.J. Gulati, and K. Dandekar, "*In vivo* Compressibility of the Human Fingertip," *Advances in Bioengineering*, ASME 22: 573-576 (1992).

47 W.E. Babiec, *In vivo Volume Changes of the Human Fingerpad Under Indentors*, S.B. thesis, Dept. of Mech. Eng., MIT, 1994.

48 M.A. Srinivasan and K. Dandekar, "Role of Fingertip Geometry in the Transmission of Tactile Mechanical Signals," *Advances in Bioengineering*, ASME 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," *J. Biomech. Eng.* 118: 48-55 (1996); K. Dandekar and M.A. Srinivasan, "Tactile Coding of Object Curvature by Slowly Adapting Mechanoreceptors," in *Advances in Bioengineering*, ed. M.J. Askew, BED-28: 41-42 (New York: ASME, 1994); K. Dandekar and M.A. Srinivasan, "A 3-dimensional Finite Element Model of the Monkey Fingertip for Predicting Responses of Slowly Adapting Mechanoreceptors," *Proceedings of the 1995 Bioengineering Conference*, eds. R.M. Hochmuth, N.A. Langrana, and M.S. Hefzy, BED-29: 257-258 (1995); K. Dandekar, *Role of Mechanics in Tactile Sensing of Shape*, Ph.D. diss., Dept. of Mech. Eng., MIT, 1995.

The results show that the model geometry has a significant influence on the spatial distribution of the mechanical signals and that the elastic medium acts like a low-pass filter in causing blurring of the mechanical signals imposed at the surface. Multi-layered 3D models of monkey and human fingertips accurately predicted the surface deformations under a line load, experimentally observed by Srinivasan.[49] The same models predicted the experimentally observed surface deformations under cylindrical indentors as well. These 3D finite element models were used to simulate neurophysiological experiments involving indentation by rectangular bars, aperiodic gratings, cylindrical indentors and step shapes. Several strain measures at typical mechanoreceptor locations were matched with previously obtained neurophysiological data to determine the relevant mechanical signal that causes the receptors to respond. In all simulations, the strain energy density at the receptor location was found to be directly related to the static discharge rate of the slowly adapting afferents. In addition, strain energy density is a scalar that is invariant with respect to receptor orientations and is a direct measure of the distortion of the receptor caused by loads imposed on the skin. We have therefore hypothesized that the strain energy density at the receptor site is the relevant stimulus to the slowly adapting receptors.

In order to further improve the spatial resolution, a three-dimensional finite element model with high mesh density near the loading region has been developed. To model the internal geometry more accurately, we have investigated the use of magnetic resonance imaging (MRI) to visualize the internal structures of the fingerpad. In preliminary experiments, we have achieved a resolution of 300 microns/pixel. Efforts are underway to improve the resolution to 60 microns/pixel. Also, to account for the dynamic behavior of the fingertip, the models are being enhanced to include viscoelastic effects. Once the models achieve sufficient spatial resolution and simulate temporal effects, they can be used to generate hypotheses on how the central nervous system might infer object shape from mechanoreceptor signals.

1.13 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, Professor Anuradha M. Annaswamy, Dr. Robert H. LaMotte

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 etched on plane glass plates.[50] 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.[51]

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 at various orientations across the fingerpads of anesthetized monkeys, and evoked responses in cutaneous mechanoreceptive primary afferent nerve fibers were recorded.[52] Major geometrical properties of the

---

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

50 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).

51 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).

52 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, "Cutaneous Neural Codes for Shape," *Can. J. Physiol. Pharmacol.* 72: 498-505 (1994); R.H. LaMotte, C. Lu, and M.A. Srinivasan, "Peripheral Neural Representation of the Shapes and Orientations of Three-dimensional Objects Stroked Across the Monkey Fingerpad," *Soc. Neurosci. Abstr.* 21: 1162 (1995); R.H. LaMotte, C. Lu, and M.A. Srinivasan, "Tactile Neural Codes for Shapes and Orientations of Objects," in *Information Processing in the Somatosensory System,* eds. O. Franzen, R. Johansson, and L. Terenius (Basel, Switzerland: Birkhauser Verlag AB, forthcoming).

shapes were well represented in the spatiotemporal 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 until now: 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.

The 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 the response, such as the outline of the region of mechanoreceptor activity in the skin, encoded the size, shape and orientation of the 2-D outline of the object in contact with the skin. The third dimension of shape was represented best in the shape of the distribution of spatial discharge rates primarily in the SA fiber population. The shapes of the spatial discharge rates of RAs were more irregular and variable than those of SAs and exhibited poor or no representations of object shape. It was hypothesized that the distribution of slopes (in impulses/sec/mm)of the spatial discharge rate profile evoked by an object in the SA population encoded the distribution of curvatures on the surface of the object in contact with the skin. This is 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 without differentiating (which is a noise enhancing process), but by exploiting its relationship 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 distribution, thus encoding the shape of the object in terms of its surface curvatures.[53]

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.[54] 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. The mechanistic data for rubber specimens indicates that the basis for the perception of softness of rubber-like objects is likely to be the spatio-temporal variation of pressure on the skin (or, equivalently the skin displacement and its derivatives). Neurophysiological data shows that the resulting responses from slowly adapting type I afferent population within the skin might encode the compliance of such objects.

1.14 Inertial Head Tracking

Sponsor

National Aeronautics and Space Administration
Grant NCC 2-771

Project Staff

Brian Clarkson, Nathaniel I. Durlach, Eric M. Foxlin, Owen D. Johnson, Christopher R. Richardson, John J. Rodkin, Nathan R. Shnidman

The objective of this effort is to develop a new head-tracking technology for HMD applications which offers large working volume, high fidelity, and unencumbered operation.

53 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).

54 M.A. Srinivasan and R.H. LaMotte, "Tactual Discrimination of Softness," *J. Neurophys.* 73(1): 88-101 (1995); M.A. Srinivasan and R.H. LaMotte, "Tactual Discrimination of Softness: Abilities and Mechanisms," in *Information Processing in the Somatosensory System,* eds. O. Franzen, R. Johansson, and L. Terenius (Basel, Switzerland: Birkhauser Verlag AB, forthcoming).

Commercially available tracking systems suffer from a variety of limitations which restrict their use in human-machine interfaces. Mechanical trackers offer fast response times and good accuracy, but require physical attachment to the head which limits range of motion and user comfort. Optical methods are expensive, require a clear line-of-sight, and usually necessitate a trade-off between precision and working volume. Magnetic trackers (such as those available through Polhemus or Ascension Technologies) are reasonably priced, but have a small working volume and suffer from noise and magnetic field distortions caused by metallic objects. Finally, acoustic systems based upon a transmitted signal's time-of-flight are available at reasonable cost and offer a large working volume; however, their speed is limited by the speed of sound, and they are sensitive to acoustic interference.

Our system departs from prior art in head-tracking by relying on inertial sensors for its primary measurements. This offers the potential advantage of a self-contained head-mounted measurement device that does not rely on signals from other equipment in the lab. This makes possible large range and freedom from interference. The main difficulty with inertial systems is drift, which we correct with an inclinometer and compass. The orientation tracker therefore retains its large range and relative invulnerability to interference.

Due to the difficulty of tracking translational position by inertial means, the first two prototypes were designed to measure orientation only. In the past year, we have developed a hybrid 6-DOF tracking system which uses a Kalman filter to combine ultrasonic position updates with a 6-DOF inertial tracker. Other projects begun during the past year include development of a high performance inertial measurement unit (IMU) based on miniature dynamically-tuned gyroscopes and a low-cost ultra-miniature IMU based on solid state angular rate sensors and accelerometers.

1.14.1 Ultra-Miniature Solid-State Inertial Measurement Unit (IMU)

A third prototype of the MIT inertial tracker has been constructed which reduces the volume by a factor of three compared to the second prototype. The new sensor, based on state-of-the-art micromachined angular rate sensors and accelerometers, has dimensions of 1 × 1.24 × 1.34 inches and weighs 2.1 ounces. The biggest change to the design of the sensor assembly is that the fluid inclinometer was eliminated and repaced by a triad of orthogonal accelerometers which can determine the gravimetric tilt regardless of attitude. In addition, the 2-axis fluxgate magnetometer was replaced with a much less bulky solid-state magnetometer. New interface circuitry was developed for the new sensor to interface it to the PC ISA bus for development and also to interface to a small microcontroller card for ultimate application in wearable and wireless VR systems.

To achieve this level of miniaturization, lower performance angular rate sensors had to be used. To compensate, more thorough methods of calibration were developed, making use of a servo-motor driven tilt table/rate table apparatus.

1.14.2 6-DOF Inertial Tracker

The largest effort during the past year has been to expand the tracking capabilities from 3-DOF to 6-DOF. In addition to providing all-attitude gravimetric tilt correction, the triaxial acclerometers in the new IMU were included for this purpose. Background reading and coursework have been completed in the field of strapdown inertial navigation to identify any existing theory that can be applied to the head-tracking problem. A hybrid 6-DOF tracking system has been designed which uses a Kalman filter to achieve efficient sensor data fusion of the signals from the three rate gyroscopes, three linear accelerometers, and an ultrasonic time-of-flight position triangulation system. The portion of the Kalman filtering algorithm which has been implemented so far enables improved estimation of orientation, which is vital to extracting linear acceleration in a confounding gravity field. The details of this Kalman filter have been sumbmitted for publication in VRAIS96. Preliminary efforts in combining the ultrasonic position data with the double integrated acceleration data have shown promise of achieving the smoothness, update rate, and predictive capability of the accelerometers, with the accuracy of the ultrasonic system.

1.14.3 High-Performance IMU using Dynamically-Tuned Gyroscopes

To determine the ultimate accuracy that might be achieved by inertial body-trackers, we have begun to construct an IMU using today's state-of-the-art tactical-grade avionics technology which might represent the level of performance that will be achieved with tomorrow's low-cost micromachined technology. A final selection of a combination of 2-DOF DTGs and silicon servo-accelerometers has been made, and the development of the drive electronics for these sophisticated inertial instruments is well under-way.

1.14.4 Publication

Foxlin, E. "Inertial Head-Tracker Sensor Fusion by a Complementary Separate-Bias Kalman Filter." *Proceedings of VRAIS*, Santa Clara, California, March 30-April 3, 1996.

1.15 Supernormal Auditory Localization for Improved Human-Machine Interfaces

Sponsor

U.S. Air Force - Office of Scientific Research
Grant F49620-94-1-0236

Project Staff

Douglas S. Brungart, John Crouch, Nathaniel I. Durlach, Professor Richard M. Held, Gregory G. Lin, Kinuku Masaki, John Park, Dr. Barbara G. Shinn-Cunningham

This work constitutes a continuation of work described previously in *RLE Progress Reports 135* and *136*.

Research in this area conducted during the past year has focused on theoretical modeling of the results obtained on adaptation to altered azimuthal cues and on preliminary experimental work concerned with azimuthal coding of distance. Comments on each of these topics are contained in the following subsections.

1.15.1 Modeling of Adaptation to Transformed Azimuthal Cues

The most important accomplishment to date has been the development of a model capable of describing the experimental results for nearly all of the experiments performed so far. In particular, the model predicts the changes in both bias and resolution as subjects adapt to supernormal localization cues.

This model is significant for both sensorimotor adaptation and psychoacoustics. Its importance in the field of sensorimotor adaptation arises from two main facts. First, the model can make quantitative predictions of most aspects of performance over the time course of the adaptation process. Not only are most previous adaptation models qualitative rather than quantitative, but other models are generally restricted to describing only how mean performance changes over time; no attention is given to how resolution varies over time. As such, the current model is more powerful than any existing model of sensorimotor performance found in the literature. Second, the necessary assumptions of the model have important implications for sensorimotor adaptation in general.

The model assumes that (1) subjects cannot adapt to nonlinear transformations of auditory localization cues, but instead adapt to a linear approximation of the imposed nonlinear transformation, and (2) given the linear constraint on adaptation, performance asymptotes to levels very close to the ideal levels achievable for the imposed nonlinear transformation. The fact that these assumptions allow the model to describe the results from the current experiments leads to a number of important questions about other sensorimotor adaptation studies. For instance, is it a general principle that humans cannot adapt to nonlinearities in sensorimotor rearrangement? Can it be shown that the failure of humans to completely adapt to other types of rearrangement is really a by-product of their inability to adapt to nonlinear transformations? How universal are the principles found in the current model? These questions are of great importance both for scientific understanding of sensorimotor adaptation and practical applications in which human subjects must adapt to sensorimotor rearrangements.

The importance of the model in the field of psychoacoustics arises from its ability to describe how performance evolves over time when subjects are provided with specific feedback or training that pushes them to interpret physical stimuli in new ways. No existing psychophysical models provide insight into how aspects of performance might change over time.

The current model incorporates ideas from the fields of sensorimotor adaptation and psychoacoustics to lead to a model that significantly adds to both areas. Whereas previous models of sensorimotor adaptation provided only qualitative descriptions of how one aspect of performance changed over time, previous psychophysical models provided detailed quantitative predictions of all aspects of performance at one point in time. The current model of adaptation to supernormal auditory localization cues provides quantitative predictions of all aspects of performance over time.

In the model, it was assumed that the range of stimuli attended to at any given time (determined by the slope relating mean response to acoustic stimulus) determined the amount of noise in the human perceptual system. For larger ranges of stimuli, noise increased, whereas for smaller stimulus ranges, it decreased. In the model, as subjects adapted to supernormal cues, the range of stimulus values to which the subject attended also increased. This increase in range made the model predict a decrease in resolution over time for the same physical stimuli.

1.15.2 Distance Coding

It is well known that sound source distance is not well perceived naturally. Furthermore, very little effort has been made to explore the extent to which perception of distance could be substantially improved for use in acoustic displays by means of artificial coding.

We have completed an initial study concerned with subjects' abilities to identify various filter transfer characteristics,[55] using characteristics that we might use for coding distance. This work differs from past work performed at a variety of laboratories, including our own, on the perception of spectral shape[56] in that we are interested in selecting shapes and experimental paradigms related to the coding of distance rather than in modeling of cross-frequency intensity comparisons. In particular, in most cases our attention is focused on (1) transfer characteristics that are related (at least loosely) to those naturally encountered and that do not interfere too seriously with the perception of the transmitted signal (i.e., with the "message") and (2) the task of identification rather than discrimination. Our initial series of experiments examined absolute identification (AI) performance using the three-dimensional set of "single-echo" filters given by the following equations:

$$|S_{A,m,\tau}(\omega)|^2 = A^2\left[1 + m^2 + 2m\cos(\omega\tau)\right]$$

$$\text{phase}[S_{A,m,\tau}(\omega)] = \tan^{-1}\left[\frac{-m\sin(\omega\tau)}{1 + m\cos(\omega\tau)}\right]$$

We have investigated the ability of listeners to perceive information encoded as the strength and delay of a single echo of the source. Although the work focused on how much information listeners can extract when distance-like cues are presented rather than on the perception of distance per se, it is a first step toward developing simple but reliable distance cues for a virtual-environment system.

The most important result from these experiments was that the amount of information transfer (IT) was startingly small. Whereas many unidimensional stimulus sets lead to an IT of 2 to 3 bits (and two-dimensional stimulus sets to an IT of 3-5 bits), in these experiments the value of IT obtained fell in the range of 0.2 - 2 bits. In general, these small values of IT appeared to result from two factors: (1) large JNDs in the variables m and τ and (2) lack of perceptual independence between m and τ (i.e., discrimination or identification of one variable increased substantially when the value of the other variable was randomized). Thus, it appears that encoding distance solely by means of the single-echo modulation parameters m and τ cannot lead to good distance resolution.

1.15.3 Publications

Shinn-Cunningham, B.G. "A Dynamic, Psychophysical Model of Adaptation in Localization Experiments." *J. Acoust. Soc. Am.* 97(5): 3411 (1995).

Shinn-Cunningham, B.G., and A. Kulkarni. "Applications of Virtual Auditory Space." In *Virtual Auditory Space.* Ed. S. Carlile. New York: Landes Publishing Company. Forthcoming.

Shinn-Cunningham, B.G., H. Lehnert, G. Kramer, E.M. Wenzel, and N.I. Durlach. "Auditory Displays." In *Spatial and Binaural Hearing.* Eds. R. Gilkey and T. Anderson. New York: Erlbaum. Forthcoming.

Shinn-Cunningham, B.G., P.M. Zurek, E.R. Stutman, and R. Berkovitz. "Perception of Azimuth for Sources Simulated Using Two Loudspeakers in Natural Listening Environments." *Mid-Winter Meeting of the Association for Research in Otolaryngology*, St. Petersburg, Florida, February 4-8, 1996.

---

[55] D.S. Brungart, *Distance Information Transmission Using First Order Reflections,* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1994.

[56] D.M. Green, *Profile Analysis* (Oxford University Press, 1988); N.I. Durlach, L.D. Braida, and I. Ito, "Toward a Model for Discriination of Broadband Signals," *J. Acoust. Soc. Am.* 62(4): 940-947 (1986); C.L. Farrar, C.M. Reed, Y. Ito, N.I. Durlach, L.A. Delhorne, P.M. Zurek, and L.D. Braida, "Spectral-shape Discrimination. 1. Results from Normal-hearing Listeners for Stationary Broadband Noises," *J. Acoust. Soc. Am.* 81(4) (1987).

1.16 Auditory Cues to Combat Spatial Disorientation

Sponsor

U.S. Air Force - Office of Scientific Research
Agreement with Brandeis University

Project Staff

Nathaniel I. Durlach, Professor Richard M. Held, Dr. Barbara G. Shinn-Cunningham

Spatial disorientation (SD) in flight leads to the death of many flight crew members and wastes millions of dollars per year. Disorientation occurs because the aerial environment includes novel features to which our sensorimotor systems are not appropriately attuned.

In most past work directed towards correcting this problem, attention has focused primarily on the visual system, the vestibular system, and nonvestibular kinesthesia, as well as on interactions among these systems. In contrast, our work focuses on the possible use of auditory stimulation to alleviate SD problems.

To evaluate the efficacy of using auditory stimulation for this purpose, it is necessary to determine the extent to which and manner in which auditory perception is altered by acceleratory force fields. Although it is already known that certain illusory effects in audition can be induced by such fields (the audiogravic and audiogyral illusions), these effects have not yet been adequately characterized. Also, because such little previous research has been done in this area, it is possible that other important illusory affects also exist but have not yet been noted.

This research program is a collaborative effort by the Graybiel Spatial Orientation Laboratory at Brandeis University and the Sensory Communication Group of the Research Laboratory of Electronics at MIT. Whereas the former has personnel and facilitates appropriate to the creation and study of acceleratory force fields, the latter has personnel and facilities appropriate to the study of audition.

The objectives of our work can be summarized as follows:

(1) Configure currently available equipment for providing accelerative forces and binaural audio in a manner suitable for studying the effects of such forces on auditory perception of azimuth (as well as auditory perception of pitch, loudness, and subjective duration);

(2) Make use of this facility to study how specified auditory perceptions are altered by application of various types of accelerative force fields (chosen to simulate the force fields encountered in aircraft flight);

(3) Construct a quantitative theoretical model to interpret the resulting data and to extend our ability to predict auditory function in non-1g force fields beyond the experimental conditions.

Work in the first eight months of the program (the period covered by this progress report) consisted of developing our experimental facilities and performing initial experiments examining how localization of a single sound source is affected by acceleratory force fields.

The initial experiments assessed changes in perceived localization of a sound in the azimuthal plane of the head during changes of the linear, gravitoinertial force field (G) in the head's azimuthal plane produced by centrifugation in a rotating room. Subjects were restrained in a bed, 2 m from the center of the rotating room, that kept the body's long axis in an Earth-horizontal plane and the interaural axis oriented along a radius of the room, right ear toward the center. Standard trials were 704 sec long. They involved 100 sec with the room stationary (normal 1 g condition), 152 sec of acceleration up to a speed that produced a G level two times the magnitude of normal g and tilted 60 deg toward the center of the rotating room, relative to normal g, 100 sec at this constant speed, 152 sec of deceleration and 200 sec with the room stationary again. In alternate trials, subjects set a sound to appear in the median plane of their head or set a joystick to indicate haptically the perceived median plane of the head. In sound localization trials, subjects listened to wide band noise gated at 4 Hz (50 percent duty cycle), spatialized with a Convolvotron II and played through earphones. The sound came on in a random location (azimuth) every 20 sec and the subject had 15 sec to set it to the apparent median plane of the head by adjusting a bidirectional switch that controlled the HRTF used by the Convolvotron. In haptic trials there was no sound except for a beep every 20 sec signaling the subject to set the joystick to coincide with head's perceived median plane.

Ten subjects participated in a series of three related experiments involving different G profiles. In the Experiment 1, the bed was flat, in an Earth-horizontal plane. Thus, when the room reached its dwell velocity, G was equal to 2 g in magnitude and was tilted 60 deg to the subject's right. Relative to the pre- and post-rotation periods, the average sound settings during rotation were 8.8 deg to right. (Put differently, subjects heard a physically constant sound shift to the left during acceleration.) The shift developed gradually during acceleration. There was no evidence of a return to baseline (adaptation) during 100 sec of high G, and when the room decelerated auditory settings returned to pre-rotation baseline without overshoot or delay. Joystick settings were mirror symmetric to the sound settings, shifting to the left of baseline (Mean= 11.7 deg) during rotation. Thus, changes in both G magnitude and azimuthal orientation of G produce a change in the perceived, head-relative location of a constant auditory stimulus. The next two experiements showed that neither factor (magnitude or direction) alone is sufficient. In Experiment 2, the bed was pre-positioned to hold the subject's in a 60 deg right ear down postion in 1 g so that during rotation the 2.0 g force field would be parallel to the median plane. There were no differences between auditory settings in 1.0 g with the body Earth horizontal and in 2.0 g with the body horizontal relative to the resultant G field. Thus, a change in G magnitude without change in G orientation is insufficient to alter auditory localization. Experiment 3 was run entirely in stationary, 1 g conditions. Subjects made auditory and haptic settings with the bed tilted to orientations between 75 deg left ear down and 75 deg right ear down, at 15 deg increments, in random order. There was no effect of tilt in 1 g on sound localization. Thus, a change in azimuthal orientation relative to the normal 1 g field is not sufficient to produce an auditory localization shift. Head-relative changes in auditory localization do occur when G both increases in magnitude and tilts away from the median plane.

Supplementary observations confirmed and extended these observations. Experiment 4 confirmed the results of Experiment 1 with a different psychophysical technique. We provided a constant auditory stimulus throughout an entire 704 sec trial and asked subjects to point to the apparent location with the joystick. Four subjects perceived a 13 deg leftward shift with a time course similar that of the settings in Experiment 1. Perceived shift to the left of a constant sound is consistent with settings to the right of the perceived auditory midline. In Experiment 5, subjects were oriented with their left ear toward the center of the room in order to reverse, relative to Experiment 1, the direction of change in G orientation relative to the head during centrifugation. The results were that auditory settings were reversed relative to Experiment 1. In Experiment 6, we assessed whether auditory localization would remain constant during prolonged exposure to an augmented, tilted G field. During 10 minutes (as compared to 100 sec in Experiemnt 1) of exposure to a 2 g, 60 deg tilted field, there was no change in auditory or haptic settings relative to the median plane.