Section 2 Sensory Communication

Chapter 1 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. Donald K. Eddington, Dr. Susan L. Goldman, Dr. Julie E. Greenberg, Dr. Lynette A. Jones, Dr. Jeng-Feng Lee, Dr. William M. Rabinowitz, Dr. Christine M. Rankovic, Dr. Charlotte M. Reed, Dr. Wendelin L. Sachtler, Dr. J. Kenneth Salisbury, 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, Lorraine A. Delhorne, Seth M. Hall, Dorrie Hall

Visiting Scientists and Research Affiliates

Dr. Paul Duchnowski, Geoffrey L. Plant, Dr. Matthew Power

Graduate Students

Walter A. Aviles, Gerald L. Beauregard, Maroula Bratakos, Jyh-Shing Chen, Kiran Dandekar, Joseph Desloge, Eric M. Foxlin, Joseph A. Frisbie, Rogeve J. Gulati, Rakesh Gupta, Louise Jandura, Jean C. Krause, David S. Lum, Hugh B. Morgenbesser, Philip M. Nadeau, Michael P. O'Connell, Nicholas Pioch, Hong Z. Tan, Daniel P. Welker, Craig B. Zilles

Undergraduate Students

Walter E. Babiec, Stephen V. Baird, James H. Bandy, Susan E. Born, Erika N. Carmel, Gail Denesvich, Ashanthi Gajaweera, Dorrie Hall, Gabrielle Jones, Steingrimur P. Karason, Danielle G. Lemay, David C. Lossos, Jonathan Pfautz, Frederick L. Roby, Tomas Rodriguez-Perez, Jonathan R. Santos, Matthew Sexton, Ranjini Srikantiah, Joseph D. Towles, Lukasz A. Weber, Evan F. Wies

Technical and Support Staff

Ann K. Dix, Eleanora M. Luongo, Mabayoje Tuyo

1.1 Research on Hearing, Hearing Aids, and Tactile Communication of Speech

1.1.1 Hearing Aid Research

Sponsor

National Institutes of Health
Contract 2 R01 DC00117

Project Staff

Professor Louis D. Braida, Dr. Paul Duchnowski, Dr. Matthew Power, Dr. Christine M. Rankovic, Dr. Anne H. Takeuchi, Seth M. Hall, Dr. Patrick M. Zurek, Joseph A. Frisbie, Ann K. Dix, Danielle G. Lemay, Matthew Sexton, Ranjini Srikantiah, Jean C. Krause, David S. Lum

Characteristics of Sensorineural Hearing Impairment

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 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.[1] More recently, Duchnowski and Zurek[2] have shown that a type of multiband amplitude expansion proposed by

[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.*, forthcoming.

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. Current research is directed towards understanding the ability of the two simulation approaches to account for the impaired perception of nonspeech sounds. To evaluate the ability of the expansion simulation to reproduce the effects of hearing impairment for such sounds, we have developed a version of the simulator that operates in real time. This simulator uses a digital signal processor to filter signals into fourteen bands (with widths comparable to critical bands), estimates the short-term energy in each band, applies a level-dependent attenuation based on this estimate, and sums the resulting attenuated band signals. Measurements[4] of simultaneous narrowband masking patterns and forward-masked psychophysical tuning curves indicate that both simulation techniques produce reductions in frequency selectivity that are similar to those observed in listeners with sensorineural hearing impairments (e.g., Dubno and Schaefer[5]). By contrast, the expansion simulation preserves the microstructure of the normal-hearing listener's audiogram, while the masking simulation does not.

Characteristics of the Speech Signal

Token Variability

The goal of this research is to understand the limitations on speech intelligibility associated with natural variations in the utterances of the same speech sound. Although differences in the acoustic realization of speech elements associated with the identity of the speaker are typically large enough to have stable perceptual effects, subtler differences associated with context and variability of production in a fixed context can also be discerned[6] and effect the ability to identify speech sounds, particularly when the speech signal is degraded by noise or other distortions. Although the difficulties posed by such variation for automatic speech recognition systems are increasingly appreciated, their effects on human speech reception have not been studied systematically.

To obtain insight into the combined effects of multiple sources of acoustic variability, Takeuchi and Braida[7] measured the ability of listeners to identify /b/-V-/t/ syllables in which the vowel was selected from 9 monophthongs of American English. To make the identification task nontrivial, the syllables were distorted by a multiplicative noise process. In one stimulus set, the three tokens of each syllable were produced by a single speaker, but each token was distorted by three different noises. In a second set, the tokens were produced by three talkers, but each token was distorted by a single noise. The effects of speaker variability were generally greater than the effects of token variability or noise variability. Token variability tended to have a greater effect when the stimulus set varied in distorting noise than when the stimulus set varied in speaker, suggesting that the degree of variability effect depends upon the context provided by the other stimuli in a set. The effects of speaker and token variability, as well as of token and noise variability, were generally additive. Blocking the stimuli by holding one or both sources of variability constant reduced, but did not eliminate, these effects.

Clear Speech

We have initiated a new study of the changes that occur in the acoustic speech signal associated with the attempt to speak clearly. Previous studies (e.g.,

---

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 D.S. Lum, *Evaluation of a Hearing Loss Simulator,* Advanced Undergraduate Project Report, Dept. of Electr. Eng. and Comput. Sci., MIT, 1994.

5 J.R. Dubno, and A.B. Schaefer, "Frequency Selectivity for Hearing-Impaired and Broadband-Noise-Masked Normal Listeners," *Quart. J. Exp. Psych.* 43A: 543-564 (1991).

6 R.M. Uchanski, 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.

7 A.H. Takeuchi and L.D. Braida, "Effects of Multiple Sources of Variability on the Accuracy of Vowel Identification," submitted to *J. Acoust. Soc. Am.*

Uchanski et al.)[8] have demonstrated that clear speech is generally enunciated more slowly than conversational speech, but the degree to which the reduction in rate is essential to clarity has not been established. To improve our understanding of the importance of reduced speaking rate to improved intelligibility, we plan to measure the intelligibility of speech produced by "professional" talkers under nine different speaking modes. The talkers will be individuals, such as debaters, public speakers, and actors who have been trained to adopt speaking styles that differ from everyday conversational speech. The speaking modes will differ with respect to both speaking rate and other speech characteristics. Four of the modes will be at a conversational speaking rate: soft, loud, conversational, and clear speech at conversational rate. Two modes will be at a quick speaking rate: quick speech and clear speech at quick rates. Three modes will be at a slow rate: slow speech, clear speech, and conversational speech with pauses between words (as if speaking to a discrete-word automatic speech recognition system).

Preliminary measurements of the intelligibility of speech produced in the traditional conversational and clear styles suggests that there may be a simple relation between the word score W and average speaking rate R (with sentence durations corrected for pauses between words): $W = a - bR$, where a but not b depends on the identity of the talker. Based on this observation we plan to select talkers with relatively high values of a and attempt to train them using techniques similar to those developed by Chen[9] to increase their speaking rate while maintaining the intelligibility of clear speech and intelligibility while maintaining the speaking rate of conversational speech.

Models of Speech Intelligibility

Speech Transmission Index

Unlike the Articulation Index (AI), (e.g., Dugal et al.[10]) which bases predictions of speech intelligibility on the long-term power spectrum of speech, the Speech Transmission Index (STI), (e.g., Houtgast et al.[11]) 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,[12] 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.

Although the STI is typically computed from estimates of modulated spectra derived from theoretical considerations or from measurements of modulation transfer functions made with intensity modulated noise, we are attempting to base this calculation on measurements of real speech waveforms. However, these attempts have been obstructed by artifacts that distort the envelope spectra of speech that has been degraded by noise and reverberation. Specifically, we observe a "noise floor" in the envelope spectra if any background noise is present. For power spectra of white noise, the expected power spectral level is proportional to the variance of the noise. In our case, we compute power spectra of filtered noise envelopes. The variance of the envelopes is inversely proportional to the bandwidth of the octave filters we use. We have confirmed that the noise floors we observe vary in level according to the bandwidth of the filters. We are currently examining various ways to remove or compensate for these noise floor effects.

8 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.

9 F.R. Chen, *Acoustic Characteristics and Intelligibility of Clear and Conversational Speech at the Segmental Level,* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1980.

10 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).

11 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).

12 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(4): 4344-446 (1986).

Even if they work on noise alone, none of the strategies tested have resulted in speech in noise spectra as we would expect.

We have also examined some alternative techniques to calculate the STI. Ludvigsen et al.[13] have developed a technique in which the squared envelopes of octave bands of undistorted speech are correlated with those of degraded speech using a linear regression analysis. The slope of the regression is used, along with the squared envelope means, to compute a transmission index for each octave band. When we implemented their algorithm, we found that it is extremely sensitive to the existence of delays between the original and corrupted speech envelopes and would thus be difficult to apply to reverberant speech. Since, in many other ways this technique holds promise, we will continue studying this idea.

Publications

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

Lum, D.S. *Evaluation of a Hearing Loss Simulator.* Advanced Undergraduate Project Report, Dept. of Electr. Eng. and Comput. Sci., MIT, 1994.

Maxwell, J.A., and P.M. Zurek. "Feedback Reduction in Hearing Aids." *IEEE Trans. Speech Audio Proc.* Forthcoming.

1.1.2 Enhanced Speechreading

Sponsor

National Institutes of Health
Contract 1 R01 DC02032

Project Staff

Professor Louis D. Braida, Dr. Susan L. Goldman, Charlotte M. Reed, Lorraine A. Delhorne, Dr. Paul Duchnowski, Maroula Bratakos, Gabrielle Jones, Mabayoje Tuyo, Philip M. Nadeau, Danielle G. Lemay, Matthew Sexton

Basic Studies of Audiovisual Integration

The McGurk effect[14] demonstrates that visual lip movements can influence perception of auditory speech even when auditory and visual cues are in conflict. For example, when auditory /ma/ is dubbed onto visual lip movements of /ta/, the auditory-visual stimulus is often perceived as "na." Whereas this effect has been found to be robust for American perceivers, it is generally much weaker for Japanese perceivers when the stimuli are produced by Japanese speakers.[15] Although this finding suggests that linguistic and/or cultural factors affect the integration process, it may also reflect the properties of the utterances of the single Japanese speaker who produced the utterances used in these studies.

We conducted an experiment with new stimuli to explore this issue. These stimuli consisted of utterances of the syllables /ba,da,ga,pa,ta,ka,ma,na/ produced by two Japanese and two American speakers. Recorded sound and videotaped lip movements from each speaker were combined to create 16 pairs, half of which consisted of concordant AV components (e.g., A=/ba/, V=/ba/) the remainder consisted of discrepant components (e.g., A=/ba/, V=/ga/). Since Japanese vowels are typically shorter than English vowels, the Japanese speakers were instructed to produce sounds with the same duration as those produced by the English speakers. This is in contrast to our previous study in which the American speaker was instructed to produce shorter vowels. The slower Japanese articulations were more easily lipread than those used in earlier studies.

The stimuli were presented in auditory-visual (AV), visual only (V), and auditory only (A) conditions. Subjects wrote what they heard (AV and A conditions), or what they thought the speaker was saying (V condition), with no restrictions on the response set. The size of the McGurk effect was estimated as the increase in percentage of auditory place of articulation errors (labial versus nonlabial) caused by discrepant visual cues. Results indicated that although the group difference was replicated, the Japanese subjects showed a stronger McGurk effect than before for auditory labials, although

[13] C. Ludvigen, *Acta Otol.* Suppl 469: 190-195 (1990).

[14] H. McGurk and J. MacDonald, "Hearing Lips and Seeing Voices," *Nature* 264: 746-748 (1976).

[15] K. Sekiyama and Y. Tohkura, "McGurk Effect in Non-English Listeners," *J. Acoust. Soc. Am.*, 90: 1797-1805 (1991); K. Sekiyama, "Differences in Auditory-Visual Speech Perception Between Japanese and Americans," *J. Acoust. Soc. Japan (E)* 15: 143-158 (1994).

there were large individual differences. Unlike our previous study,[16] there was no difference between the Japanese and English stimuli. The strong effect of the visual stimulus on Japanese subjects appears to be related to the auditory ambiguity and visual robustness of the stimuli for these listeners. For the American perceivers, on the other hand, the size of the McGurk effect was comparable to that found previously and relatively unaffected by the ambiguity of the auditory stimuli.

Supplements Based on Automatic Speech Recognition

The goal of this work is to develop systems for producing and displaying discrete speechreading supplements that can be derived from the acoustic signal by speech recognition technology. Current efforts are focused on supplements related to the Manual Cued Speech System used in some educational and communication settings by deaf individuals.

Although previous studies of the reception of Manual Cued Speech and of the performance of ASR systems suggests that current technology may provide an adequate basis for the design of an automatic cueing system,[17] 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 have initiated a study of the reception of synthetically generated cues characterized by errors and delays. The visual cues are static handshapes that correspond to the nominal shapes and positions of manual cued speech but lack fluid articulation. These cues are dubbed onto video images of a face that were recorded simultaneously with the acoustic speech signal. In preliminary tests on one skilled receiver of Manual Cued Speech, scores for words in IEEE sentences were 48 percent for speechreading alone, 92 percent for Manual Cued Speech, 84 percent for error-free static handshapes, and 72-74 percent for cues derived from phones recognized with 10-20 percent errors. Although scores for error-free cues are lower than for manual cued speech, this may result from the reduced speaking rate that accompanies cue production (100 versus 150 wpm). The scores obtained with cues derived from a simulation of recognition errors are sufficiently encouraging to warrant further study of this approach.

Supplements Based on Signal Processing

Listeners who have severe to profound hearing impairments typically rely heavily on speechreading to communicate and often receive little benefit from conventional hearing aids in the absence of speechreading. We are attempting to develop aids that present simplified representations of the speech waveform acoustically to the impaired ear. These simplified signals consist of tones that are amplitude modulated by the amplitude envelopes of filtered bands of speech. Research in this area is focused on the effect of background noise on the benefit provided by the signals, and with perceptual interference between pairs of modulated tones.

Envelopes Derived from Noisy Speech

The amplitude envelopes of filtered bands of speech, when derived from clean speech and presented as amplitude modulations of tones located at the centers of the bands, substantially improve the speechreading ability of listeners with normal hearing.[18] To assess the effectiveness of these supplements when the band-envelopes are extracted from noisy speech, Lemay and Braida[19] tested four young adults with normal hearing and superior speechreading abilities on word reception in both highly contextual (CUNY) and low-context sentences (IEEE). The interference was additive Gaussian noise filtered to have a speech-like power spectrum. The noisy speech was filtered into

[16] K. Sekiyama, "Differences in Auditory-Visual Speech Perception Between Japanese and Americans," *J. Acoust. Soc. Japan E* 15: 143-158 (1994).

[17] R.M. Uchanski, L.A. Delhorne, A.K. Dix, L.D. Braida, C.M. Reed, and N.I. Durlach, "Automatic Speech Recognition to Aid the Hearing Impaired. Prospects for the Automatic Generation of Cued Speech," *J. Rehab. Res. and Dev.* 31: 20-41 (1994).

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

[19] D.G. Lemay and L.D. Braida, "Band Envelope Speechreading Supplements Derived From Noisy Speech," *J. Acoust. Soc. Am.* 95: 3014 (1994).

octave bands centered at 500, 1600, and 3300 Hz, fullwave rectified, and low pass filtered (50 Hz bandwidth) to create band-envelope signals. The resulting low-bandwidth signals were then used to amplitude-modulate one or more tones centered in the passbands.

For the CUNY sentences, word intelligibility scores increased from 31.4 percent for speechreading alone to 67.4 percent, 68.3 percent, and 47.8 percent when single band supplements derived from noiseless speech were presented at 500, 1600, and 3300 Hz respectively, and to 93.9 percent when all three envelope signals were presented. For the IEEE sentences, these scores were 14.5 percent for speechreading alone; 36.0 percent, 26.5 percent, and 18.3 percent for the single envelope signals, and 82.0 percent for the multiple envelope signal. These results are similar to the findings of Grant, Braida, and Renn (1991).[20] When the envelopes are derived from noisy speech, scores in the audiovisual condition are reduced, but higher than for speechreading alone for $S/N > -4$ dB (CUNY) and $S/N > 0$ dB (IEEE) for the multiple envelope supplement. Roughly half of the increase in score (relative to speechreading alone) is retained at $S/N = 0$ dB (60 percent words correct) for the CUNY sentences. These results suggest that the band envelope signals can be successfully adapted for presentation to listeners with severely restricted dynamic ranges, since at a S/N of 0 dB, excursions of the component envelope signals are limited to roughly 12-15 dB by the relatively constant background noise. Moreover, all subjects were able to utilize the envelopes regardless of lipreading ability while the best lipreaders derived the greatest benefits.

Supplements Derived from Multiple Band Envelopes

Although speechreading can be considerably enhanced when a normal hearing listener receives a single tone modulated by the amplitude envelope of a single octave speech band, even larger gains are possible when additional envelopes are used to modulate tones at the center frequencies of the bands from which they were derived. To benefit a listener with high-frequency hearing loss, these modulated signals must be presented at low frequencies in close spectral proximity. However, when two or more modulated tones are simultaneously presented, regardless of the frequency separation between the tones, there may be perceptual interactions that can enhance or interfere with the perception of the modulation patterns.

In previous work, Takeuchi and Braida[21] reported that the ability of normal-hearing listeners to compare amplitude modulation patterns was adversely affected by the presence of simultaneous ipsilateral amplitude-modulated distractors. Perceptual interference effects, although generally small in magnitude, were greatest when the distractor envelope was uncorrelated with the target pattern and was greater when the separation between carrier and distractor carrier frequencies was small (200 and 250 Hz) than when it was large (200 and 500 Hz). To determine whether this finding results from the effects of peripheral masking (which typically decreases with spectral separation), Takeuchi and Braida[22] extended these experiments to include conditions in which the distractor consisted of an unmodulated tone or was presented dichotically. In general, the amount of interference found in the present experiment was smaller than that found in the earlier one, but the dependence on correlation and separation was similar for replicated conditions. Dichotic presentation produced roughly the same amount of interference as diotic presentation, for both correlated and uncorrelated distractors. This indicates that interference is unlikely to be of peripheral origin and also suggests that listeners can perceive correlations between envelopes equally well under diotic and dichotic presentation. Unmodulated distractors did not degrade envelope comparisons when the frequency separation was large and reduced performance when the frequency separation was small only for two of the four listeners tested.

Field Studies

To obtain insight into the benefits provided to hearing impaired listeners by band-envelope supplements in everyday situations, we plan to provide

---

20 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).

21 A.H. Takeuchi and L.D. Braida, "Recognition of Amplitude Modulation Patterns in the Presence of a Distractor: I. Effects of Correlation and Frequency Relation," *J. Acoust. Soc. Am.*, forthcoming.

22 A.H. Takeuchi and L.D. Braida, "Recognition of Amplitude Modulation Patterns in the Presence of a Distractor: II. Effects of Dichotic Presentation and Unmodulated Distractors," *J. Acoust. Soc. Am.*, forthcoming.

wearable DSP-based SiVo Aids[23] that have been programmed to produce these supplements to listeners with severe to profound impairments. In preliminary work, digital oscillators and filters have been implemented on the TMS320C50 processor used in the SiVo Aid and found capable of producing signals similar to those studied previously. Preliminary evaluations of these amplitude envelope signals as aids to speechreading have been performed by one hearing-impaired listener using tests of word intelligibility in (IEEE) sentences. For that listener, who is an effective hearing aid user, the increase in intelligibility provided by the single-band envelope cues was found to be roughly equal to that provided by her own aid. Detailed planning of the field study and recruitment of participants is now underway.

Publications

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

Takeuchi, A.H., and L.D. Braida. "Recognition of Amplitude Modulation Patterns in the Presence of a Distractor: I. Effects of Correlation and Frequency Relation." *J. Acoust. Soc. Am.* Forthcoming.

Takeuchi, A.H., and L.D. Braida. "Recognition of Amplitude Modulation Patterns in the Presence of a Distractor: II. Effects of Dichotic Presentation and Unmodulated Distractors." *J. Acoust. Soc. Am.* Forthcoming.

Uchanski, R.M., L.A. Delhorne, A.K. Dix, L.D. Braida, C.M. Reed, and N.I. Durlach. "Automatic Speech Recognition to Aid the Hearing Impaired. Prospects for the Automatic Generation of Cued Speech." *J. Rehab. Res. Dev.* 31: 20-41 (1994).

1.1.3 Cochlear Implants

Sponsor

National Institutes of Health
Contract 2 P01 DC00361[24]
Contract N01 DC22402

Project Staff

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

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

During the past year, work has involved (1) analysis of cue integration in audiovisual speech reception, (2) studies of interactions among different implant channels, and (3) alternative speech processing for improved implant performance. The work in (3) is performed with Joseph Tierney and Dr. Marc Zissman (see also part 5 section 3 chapter 1.7).

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

A series of experiments on audiovisual integration with speech segments has been essentially completed. Closed-set identification for a set of ten vowels, a set of the 12 most frequently occurring consonants, and the full set of 24 (initial-position) consonants was tested. Results have been

[23] 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: 2136-215 (1992).

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

obtained using vision alone, audio (the implant) alone, and audiovisually. Analyses to determine the efficiency with which the information available separately from vision and audition is utilized in the combined audiovisual condition are underway.[25] Initial analysis of the 12-consonant results suggests that (1) confusion patterns for the vision-alone condition are more consistent and more highly structured than are those for the audition-alone condition, and (2) most subjects make efficient use of the unimodal results in the combined condition.

The work in channel interactions is designed to assess the perceptual independence among different channels of an implant. Current spread from intracochlear electrodes is extensive[26] and psychophysical studies with simple stimuli reveal substantial interelectrode interactions.[27] However, the consequences of these interactions for speech reception are not clear.

In one set of experiments, consonant identification was measured using each of four individual intracochlear electrodes (driven from their corresponding individual processor channels), using different pairs of electrodes, and using all four electrodes (the normal system). Analysis of multielectrode integration uses the same models being applied to study audiovisual intergration. These models quantify the extent to which the results from multielectrode conditions reflect optimum use of the information available from the separate electrode conditions versus interference from electrode interactions. The models indicate that predicted optimum performance with multiple electrodes sometimes exceeds observed performance by substantial amounts (e.g., 85 percent-correct predicted versus 65 percent-correct observed). In these cases, however, large response biases are evident in the results from the separate electrode conditions. When the integration models are constrained to maintain these same biases, the predicted and observed multichannel results are generally more similar.

Another set of experiments assessed multichannel saliency with the Ineraid prosthesis by altering the normal tonotopic mapping between the sound-processor's four filter channels and the intracochlear electrode array.[28] A single high-performing subject was tested with (1) the normal four channel processing system, (2) a "single-channel" processor, formed by summing the outputs of the four processor filters and delivering the resulting current to a single electrode, and (3) a tonotopically "reversed" system, formed by connecting the four processor filter outputs to four intracochlear electrodes in reversed order. When using the implant in conjunction with speechreading, all three mappings showed a large improvement on the recognition of words in sentences over speechreading alone. When using the implant alone (without speechreading), tests of consonant and vowel recognition, and the recognition of words in isolation and in sentences all showed a substantial decrease in performance across the three mappings: normal > single-channel > reversed. The patterns of segmental confusions and the relations among scores on different tests were highly consistent.[29] Overall, the results indicate that, despite the extensive spread of current associated with monopolar intracochlear stimulation, the Ineraid electrode array affords a degree of perceptual selectivity that substantially aids speech reception.

The work on altered speech processing involves evaluation of the recently developed continuous-interleaved-stimulation (CIS) strategy. Each channel of a CIS processor uses the compressed envelope of its bandpass filtered output to modulate biphasic current pulses that are delivered to an intracochlear electrode. Pulses are interleaved across channels to avoid simultaneous field interactions, and pulse rates are high (~2000 pps/channel) to preserve temporal waveform cues. Using up to six monopolar electrodes directly accessible with the Ineraid implant, this strategy has shown considerable promise in acute evaluations conducted in the labo-

---

[25] L.D. Braida, "Crossmodal Integration in the Identification of Consonant Segments," *Quart. J. Exp. Psychol.* 43A: 647-677 (1991).

[26] C. van den Honert and P.H. Stypulkowski, "Single Fiber Mapping of Spatial Excitation Patterns in the Electrically Stimulated Auditory Nerve," *Hear. Res.* 29: 195-206 (1987).

[27] R.V. Shannon, "Multichannel Electrical Stimulation of the Auditory Nerve in Man: II. Channel Interaction," *Hear. Res.* 18: 135-143 (1983).

[28] W.M. Rabinowitz and D.K. Eddington, "Effects of Channel-to-Electrode Mappings on Speech Reception With the Ineraid Cochlear Implant," *Ear Hear.*, forthcoming.

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

ratory.[30] In collaboration with a group in Innsbruck, a prototype portable real-time system has been developed (based on a DSP56001) that can realize some CIS implementations. Two subjects with nine years of experience using the standard Ineraid analog sound processor are now wearing the CIS system on a full-time basis. After several weeks, one subject prefers the CIS strategy; however, objective measures of speech reception show no gain (re the standard processor). The second subject shows a large gain with CIS and some evidence of continuing improvement. On relatively difficult sentences presented in quiet and without lipreading cues, the subject scores near perfectly with CIS versus 65 percent with the standard processor. For speech reception in noise, he shows a deficit of 7 dB re normal-hearing listeners, which compares favorably to deficits of 10 to 15 dB obtained with the standard analog processor.

Publication

Rabinowitz, W.M., and D.K. Eddington. "Effects of Channel-to-Electrode Mappings on Speech Reception With the Ineraid Cochlear Implant." *Ear Hear.* Forthcoming.

1.1.4 Binaural Hearing

Sponsor

National Institutes of Health
Grant R01-DC00100[31]

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.

Recent work has examined the precedence effect that is observed when leading and lagging sounds occupy different frequency regions.[32] Subjects were asked to match the intracranial lateral position of an acoustic pointer to that of a test stimulus composed of two binaural noise bursts with asynchronous onsets, parametrically-varied frequency content, and different interaural delays. The precedence effect was measured by the degree to which the interaural delay of the matching pointer was independent of the interaural delay of the lagging noise burst in the test stimulus. Results show an asymmetric frequency effect in which the lateralization influence of a lagging high-frequency burst is almost completely suppressed by a leading low-frequency burst, whereas a lagging low-frequency burst is weighted equally with a leading high-frequency burst. This asymmetry is shown to be the result of an inherent low-frequency dominance that is seen even with simultaneous bursts. When this dominance is removed (by attenuating the low-frequency burst), the precedence effect operates equally effectively both upward and downward in frequency.

1.1.5 Multimicrophone Hearing Aids

Sponsor

National Institutes of Health
Grant R01-DC00270

Project Staff

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

The goal of this research is to determine the improvements that can be provided to hearing aids through the use of multiple microphones. The work is directed toward 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. Array processing algorithms are first implemented and evaluated in terms of signal-to-noise improvement in computer simulations. The most promising approaches are then implemented in real-time with wearable devices (tethered to a computer) for laboratory evaluations

30 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).

31 Subcontract from Boston University, Professor H. Steven Colburn, Principal Investigator.

32 B.G. Shinn-Cunningham, P.M. Zurek, N.I. Durlach, and R.K. Clifton, "Cross-frequency Interactions in the Precedence Effect," *J. Acoust. Soc. Am.*, forthcoming.

in terms of speech reception in noise and sound localization by normal and hearing-impaired listeners. Some devices will be made portable for field studies.

Work in the past year has focused on algorithmic design and analysis. One particular area concerns controlling the adaptation of time-varying processing schemes to minimize signal cancellation and misadjustment. These problems arise when the array is not perfectly aligned to the desired source and in the presence of strong source signals. They are especially bothersome in the hearing aid application because: (1) hearing aid users cannot be expected to hold their heads rigidly pointing at the signal source; and (2) often the signal is larger than the interference and so it degrades the precision of interference cancellation that can be achieved. These issues were outlined previously, along with ad hoc solutions, by Greenberg and Zurek.[33] Recent work has thoroughly analyzed and refined those ad hoc solutions, producing a modified adaptive algorithm that includes two additional features. One modification is to adjust the steps taken by the adaptive algorithm in order to reduce misadjustment when the output power is strong, and the other is to inhibit adaptation based on the correlation between microphone signals. Guidelines have been established for selecting the relevant processing parameters in a variety of acoustic environments.

Another problem with adaptive systems is that reverberation results in a signal that does not arrive from the source direction, but that is correlated with the source signal. A simple solution to the problem of reverberation-induced signal cancellation has been studied.[34] By reducing the amount by which the adaptive filter can "look ahead in time" (by a bulk processing delay), signal cancellation can be avoided with little loss in noise-reduction performance.

The result of this design and analysis work is a specification for a relatively simple and robust broadside adaptive array.[35] Computer simulations have evaluated this system in a variety of acoustic environments. Steady-state results show that it provides very large improvements in relatively anechoic environments. Substantial benefits are obtained in moderate reverberation, particularly if relatively long filters are used. In extreme reverberation, performance is comparable to that of the underlying nonadaptive microphone array. Transient results indicate that convergence of the adaptive algorithm is sufficiently rapid for processing speech signals. In addition, this work has considered the use of directional microphones and the number of microphones required in a practical system.

Work to date on multimicrophone arrays for hearing aids has been aimed at processing the microphone signals to form a single output signal. Such monaural-output systems contain none of 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 these natural benefits of binaural hearing along with improvements from multimicrophone array processing, we have been exploring ways of combining them. In one project, we have been exploring 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).[36] 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, we are exploring the design of a two-microphone adaptive binaural array, with the microphones worn at the ears.[37]

In both projects a very promising approach is a particularly simple one that exploits the psychoacoustic finding that sound localization is dominated by low-frequency interaural delay cues. Using this fact, we allocate the lowpass part of the signal spectrum for transmitting natural binaural cues by simply passing

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

34 J.E. Greenberg and P.M. Zurek, "Preventing Reverberation-Induced Target Cancellation in Adaptive-array Hearing Aids," *J. Acoust. Soc. Am.* 95: 2990-2991 (1994).

35 J.E. Greenberg, *Improved Design of Microphone-Array Hearing Aids,* Ph.D. diss., Division of Health Sciences and Technology, MIT, 1994.

36 J.G. Desloge, *A Fixed Microphone Array Hearing Aid with a Binaural Output,* M.S. Thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1994.

37 D.P. Welker, *A Binaural-Output Adaptive-Beamforming Hearing Aid,* M.S. Thesis, Dept. of Electr. Eng. and Comp. Sci., MIT, 1994.

the lowpass-filtered microphone signals to the ears. The highpass part is the output of the array, presented diotically. This approach allows source localization via the low frequencies and an enhanced signal via the high frequencies.

1.1.6 Tactile Communication of Speech

Sponsor

National Institutes of Health
Grant 5 R01 DC00126

Project Staff

Lorraine A. Delhorne, Gail Denesvich, Nathaniel I. Durlach, Ashanthi Gajaweera, Geoffrey L. Plant, William M. Rabinowitz, Dr. Charlotte M. Reed, Dr. Mandayam A. Srinivasan, Hong Z. Tan, Jonathan R. Santos

The ultimate goal of this research is to develop tactual aids for the deaf and deaf-blind that can serve as substitutes for hearing in speech communication. Research is conducted in three major areas summarized below.

Basic Study of Encoding and Display Schemes

This research is designed to develop methods of displaying acoustic signals to the tactual sense for optimal information transfer. Research in this area includes measurements of basic resolution and information transfer for tasks involving active finger motions and for those involving motional or vibratory stimulation of the finger, in addition to work on the design and construction of a multifinger tactual stimulator. Studies employing active finger motions include measurements of the ability to discriminate and identify joint-angle rotation, measurements of the ability to discriminate the thickness of materials with different bending moduli,[38] and measurements of manual perception of compliance with work cues eliminated and force cues minimized.[39] Studies employing motional or vibratory stimulation include measurement of amplitude and frequency discrimination for motional (i.e., low-frequency, high-amplitude) signals[40] and a study of the reception of Morse Code sequences employing motional and vibratory stimulation.[41] In addition, work has been conducted on the design and construction of a multifinger tactual device capable of providing stimulation along the frequency continuum from motion to vibration.

Tactual Supplements to Speechreading

This research is designed to lead towards the development of tactual aids to supplement information available through speechreading. Recent research in this area has been concerned with describing and understanding differences in performance achieved through auditory and tactual presentation of a simple acoustic-based supplement to speechreading. Data comparing the two modes of presentation were obtained for various tasks, including sentence, prosodic, and segmental reception.[42] In addition, related nonspeech psychophysical measurements for tactual and auditory presentation were obtained to increase our understanding of the differential results observed in aided speechreading tasks.[43]

---

38 A. Gajaweera, *Tactual Discrimination of Thickness,* B.S. thesis, Dept. of Mech. Eng., MIT, 1994.

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

40 G.L. Beauregard, W.M. Rabinowitz, H.Z. Tan, and N.I. Durlach, "Amplitude and Frequency Resolution for Motional Stimulation," Third International Conference on Hearing Aids, Tactile Aids, and Cochlear Implants, Miami, Florida, May 2-5, 1994.

41 H.Z. Tan and N.I. Durlach, "A Study of the Tactual Perception of Motor Input Sequences," Third International Conference on Hearing Aids, Tactile Aids, and Cochlear Implants, Miami, Florida, May 2-5, 1994; H.Z. Tan, N.I. Durlach, W.M. Rabinowitz, and C.M. Reed, "Tactual Performance with Motional Stimulation of the Index Finger," (Abstract) *J. Acoust. Soc. Am.* 95: 2986 (1994).

42 J.M. Besing, 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," submitted to *Seminars in Hearing,* 1994; M.S. Bratakos, *Supplements to Speechreading: A Comparison of Auditory and Tactile Presentation of a Single-band Envelope Cue,* B.S. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1993; C.M. Reed, M.S. Bratakos, L.A. Delhorne, and G. Denesvich, "A Comparison of Auditory and Tactile Presentation of a Single-band Envelope Cue as a Supplement to Speechreading," Third International Conference on Hearing Aids, Tactile Aids, and Cochlear Implants, Miami, Florida, May 2-5, 1994.

43 W.M. Rabinowitz, C.M. Reed, L.A. Delhorne, and J.M. Besing, "Tactile and Auditory Measures of Modulation Resolution," (Abstract) *J. Acoust. Soc. Am.* 95: 2987 (1994); W.M. Rabinowitz, C.M. Reed, and L.A. Delhorne, Tactile and Auditory Measures of Modulation Resolution," Third International Conference on Hearing Aids, Tactile Aids, and Cochlear Implants, Miami, Florida, May 2-5, 1994.

Evaluation of Practical Aids

This research is designed to evaluate the speech-reception performance of experienced users of portable, wearable tactual aids and to compare this performance to that of users of other types of auditory prostheses. Research in this project area includes (1) laboratory evaluations of tactual devices with artificially deafened subjects, (2) a field study of the performance of deaf adult users of tactual aids,[44] (3) a study of the effects of training on adult performance with a tactile device, and (4) a study of the effects of training on the use of a tactile device by deaf children.[45] The tactile devices used in this research include two commercially available, wearable tactile aids: Tactaid 2 and Tactaid 7 (both products of Audiological Engineering Corporation), with research focused primarily on the Tactaid 7 as a multichannel display of the first and second speech formants. The Tactaid 2, by comparison, is a two-channel tactile display of one lowpass frequency region of speech and one highpass band. In addition, some data have been obtained on several nonwearable laboratory-based systems, including a nine-channel laboratory implementation of the Queen's University vocoder (Brooks and Frost, 1983) and a high-performance single-channel vibrator (the Alpha-M AV6 Minishaker) used to present speech-envelope supplements to speechreading.

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." Submitted to *Seminars in Hearing* (1994).

Rabinowitz, W.M., D. Henderson, C.M. Reed, L.A. Delhorne, and N.I. Durlach. "Development and Evaluation of an Improved Synthetic Tadoma System." *J. Acoust. Soc. Am.* Forthcoming .

Reed, C.M., and L.A. Delhorne. "Current Results of a Field Study of Adult Users of Tactual Aids." Submitted to *Seminars in Hearing* (1994).

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.* Forthcoming.

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

Thesis

Gajaweera, A. *Tactual Discrimination of Thickness.* B.S. thesis, Dept. of Mech. Eng., MIT, 1994.

1.2 Haptics Research

1.2.1 Introduction

In the following subsections, we describe work performed during the past year on a wide variety of projects concerned with manual sensing and manipulation (i.e., haptics). Further work on haptics, oriented specifically toward the use of virtual environments for training, is discussed in section 1.3 below.

1.2.2 Mechanistic Modeling of the Primate Fingerpad

Sponsor

National Institutes of Health
Grant R29-DC00625

Project Staff

Dr. Mandayam A. Srinivasan, Kiran Dandekar, Walter E. Babiec, Rogeve J. Gulati

When we touch an object, the source of all tactile information is the spatio-temporal distribution of mechanical loads on the skin at the contact interface. The relationship between these loads and the resulting stresses and strains at the nerve terminals within the skin plays a fundamental role in the neural coding of tactile information. Although empirical determination of the stress or strain state

[44] C.M. Reed, L.A. Delhorne, and N.I. Durlach, "Results obtained with Tactaid 2 and Tactaid 7," *Proceedings of the 2nd International Conference on Tactile Aids, Hearing Aids, and Cochlear Implants,* eds. A. Risberg, S. Felicetti, G. Plant, and K.-E. Spens, (Stockholm, Sweden: Dept. of Speech Comm. Music Acoust., KTH, 1992), pp. 149-155; C.M. Reed and L.A. Delhorne, "Current Results of a Field Study of Adult Users of Tactual Aids," submitted to *Seminars in Hearing,* 1994.

[45] G. Plant, "The Development of a Training Program for Profoundly Deaf Children Fitted with Tactaid 7," Third International Conference on Hearing Aids, Tactile Aids, and Cochlear Implants, Miami, Florida, May 2-5, 1994.

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. The hypotheses can then be verified 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.

Determination of the Shape and Compressibility of the Primate Fingertip (distal phalanx)

The first step in performing mechanistic analyses of the primate fingertip is to determine its geometric and material properties. The three-dimensional (3D) external geometry of 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.[46] We have reconstructed several human and monkey fingertips 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.[47] Our results show that increases in either the depth of indentation or the contact area with the indentor increased the compressibility of the fingertip. The highest change in fingertip volume was about 5 percent.

We have now developed an experimental setup involving a computer controlled linear actuator for fingertip volume change measurements under dynamic conditions.[48] We measured volume changes of the fingerpad under three types of indentors (point, circular flat and flat plate) imposing constant velocity ramps (1 to 8 mm/s), several depths of indentation (1 to 4 mm) and sawtooth stimulation (1 and 2 mm amplitude; 0.25 to 2Hz frequency). The fingerpad, encased in a thin latex fingercot, was placed in a fluid-filled chamber and indented by using a computer controlled stepper motor. The resulting fingertip volume changes were calculated (resolution ~ 1/100 ml) by video recording the fluid level in a small diameter pipette, transcribing the data frame-by-frame, and subtracting out the fluid level changes due to indentor displacement. The results show that reductions in fingertip volume are small and 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. We have analyzed such data for four subjects.

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.[49] The models range from a semi-infinite medium to a three-dimensional (3D) model based on the actual finger geometry, and composed of a

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

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

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

49 M.A. Srinivasan and K. Dandekar, "Role of Fingertip Geometry in the Transmission of Tactile Mechanical Signals," *Advanes in Bioengineering*, ed. M.W. Bidez (Chicago: ASME, 1992), vol. 22, pp. 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.*, forthcoming; K. Dandekar and M.A. Srinivasan, "Tactile Coding of Object Curvature by Slowly Adapting Mechanoreceptors," in *Advances in Bioengineering*, ed. M.J. Askew. Chicago: ASME, 1994, vol. 28, pp. 41-42; K. Dandekar and M.A. Srinivasan, "A Three-dimensional Finite Element Model of the Monkey Fingertip for Predicting Responses of Slowly Adapting Mechanoreceptors," ASME summer annual meeting, Beaver Creek, Colorado, 1995.

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 two-dimensional (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.

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. Multilayered 3D models of monkey and human fingertips accurately predicted the surface deformations under a line load, experimentally observed by Srinivasan.[50] 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 the simulations, the strain energy density at the receptor location was found to be directly related to the static discharge rate of the slowly adapting (SA) 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 summary, this research project has resulted in mechanistic models of the primate fingertip whose predictions match a variety of biomechanical and neurophysiological data. The models can also be used to generate hypotheses to be tested in future biomechanical and neurophysiological experiments.

1.2.3 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,[51] Steingrimur P. Karason

We have been collaborating with Dr. Robert LaMotte, 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 nm-high parallel bars on plane glass plates.[52] 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 detection of the microtexure.[53]

In studies of the tactile perception of shape, a series of two- and three-dimensional objects (e.g., cylinders, spheres, ellipsoids and wavy surfaces) were pressed or stroked across the fingerpads of anesthetized monkeys and evoked responses in cutaneous mechanoreceptive primary afferent nerve fibers were recorded.[54] Major geometrical properties of the shapes were well represented in the spatio-

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

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

52 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).

53 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).

54 M.A. Srinivasan and R.H. LaMotte, "Tactual Discrimination of Softness," *J. Neurophys.* 73(1): 88-101 (1995); R.H. LaMotte and M.A. Srinivasan, "Responses of Cutaneous Mechanoreceptors to the Shape of Objects Applied to the Primate Fingerpad," *Acta Psychologica* 84: 41-51 (1993); R.H. LaMotte, M.A. Srinivasan, C. Lu, and A. Klusch-Petersen, A., "Cutaneous Neural Codes for Shape," *Can. J. Physiol. Pharm.* 72: 498-505 (1994).

temporal responses of SA and RA afferent fiber populations, particularly those of the SAs. The results show that the following hypothesis explains qualitatively the data we have obtained: the depth of indentation and the change in curvature of the skin surface are encoded by the discharge rates of SAs; in addition, the velocity and the rate of change in skin surface curvature are encoded by the discharge rates of both SAs and RAs.

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

Based on a theoretical analysis of the mechanics of contact, we have proposed a mechanism by which shapes of objects within contact regions are perceived through the tactile sense. The curvature of the skin surface under an object, which we know from differential geometry is approximated by the second spatial derivative of surface deflection, is coded without differentiating (which is a noise enhancing process), but by exploiting its relation to surface pressure. Pressure peaks occur where the depths of indentation and/or changes in the skin surface curvature are large. 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.[55]

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.[56] 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.

Development of a Computational Theory of Haptics

Our research on the 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 similar for human and robot tactile sensing systems.

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 then gave explicit formulae for coding and decoding problems.[57] We have now expanded these results to a more general 3D half-space model where the load direction can be completely arbitrary.[58] 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 informa-

55 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).

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

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

58 S.P. Karason, M.A. Srinivasan, and A.M. Annaswamy, "Tactile Sensing of Shape," Center for Information Driven Mechanical Systems (CIDMS) *Workshop Proceedings*, Dept. of Mech. Eng., MIT, 1994.

tion 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.[59] 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 modeled the dynamic interactions that occur between compliant end-effectors and deformable objects by a class of nonlinear systems. It was shown that standard geometric techniques for exact feedback linearization 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.

1.2.4 Biomechanics of Human Fingerpad-Object Contact

Sponsor

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

Project Staff

Dr. Mandayam A. Srinivasan, Jyh-Shing Chen, Kiran Dandekar, Rogeve J. Gulati, Tomas Rodriguez-Perez, Frederick L. Roby, Joseph D. Towles

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

The goals of this project are (1) determination of 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 high precision force sensors. The videomicroscopy system consists of a set of video zoom lenses attached to a high-resolution CCD camera, whose output can either be digitized directly at about 5 frames/s, or stored on a laserdisk at real-time frame rates (30 frames/s) for off-line digitization. The zoom lenses allow continuous variation of magnification, with the field of view covering the entire fingerpad, or just a few fingerprint ridges. High contrast images are achieved with coaxial and other fiberoptic lighting. In collaboration with our colleagues at the Artificial Intelligence Laboratory at MIT, we designed and built two 6-axis force sensors that are customized to our application. These sensors have much higher resolutions (10 bit) than commercial sensors operating in a comparable range of forces (5 Newtons). Transparent test objects can be attached to these sensors for both biomechanical and psychophysical experiments.

---

[59] A.M. Annaswamy and D. Seto, "Object Manipulation Using Compliant Fingerpads: Modeling and Control," *ASME J. Dynam. Syst. Measure, Control* (1993); A.M. Annaswamy and M.A. Srinivasan, "A Study of Dynamic Interactions between Haptic Interfaces and Compliant Fingerpads," *Proceedings of the Motion Control Workshop,* Berkeley, California, March 1994; A.M. Annaswamy and M.A. Srinivasan, "The Role of Compliant Fingerpads in Grasping and Manipulation: Identification and Control," Institute of Mathematics (New York: Springer Verlag, 1995).

During the past year, we enhanced the videomicroscopy system to be able to acquire images when the human subject's finger remained passive. A linear stepper motor with a microstepping drive was added to the system. Transparent compliant specimens can be mounted on the motor to indent a stationary finger at a given velocity. The motor is controlled by a 80386 PC, with a specified indentation velocity commanded by a 80486 PC via a digital link. A strain gage based single degree of freedom force sensor is mounted between the specimen and the motor to record contact force. With this setup, we will be able to investigate how the skin-object contact region changes with indentation velocity and force. During the indentation, the contact force and the video signal generated by the CCD camera focused on the contact region are recorded. A new video grabber was incorporated into the system to allow video data to be digitized into the computer system memory in real time. This method allows the force and video data to be synchornized. The video data can be digitized at rates up to 20 frames/sec. For the processing of contact images, a computer program which uses mouse cursor input to manually extract the border of an contact image was developed. Although this method is slow, it serves to calibrate more automated image processing methods under development.

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 nonuniform illumination, we implemented homomorphic image processing algorithms with or without image decimation. The processed images showed that contact regions consisted of discontinuous "islands" along each finger ridge, with clear distinction between contact and noncontact regions over the entire field of view.

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

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.[60] 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.

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

60 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.

stimuli to the human fingerpad.[61] 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 3 mm), 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 very 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.[62] 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, followed 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.

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

Sponsor

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

Project Staff

Dr. Mandayam A. Srinivasan, Dr. J. Kenneth Salisbury, Nathaniel I. Durlach, Dr. Robert H. LaMotte, Dr. Robert D. Howe,[63] Jyh-Shing Chen, Kiran Dandekar, Louise Jandura, Steingrimur P. Karason

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 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 build better 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 during the past year are (1) development of new hardware for robotic and human studies, (2) processing of robot sensor signals and task-level control of the devices, (3)

61 R.J. Gulati and M.A. Srinivasan, "Human Fingerpad under Indentation. I: Static and Dynamic Force Response," ASME summer annual meeting, forthcoming.

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

63 Harvard University, Cambridge, Massachusetts.

experiments on human perception and control of forces using some of the devices. The subsections to follow provide descriptions of the results obtained in each of the three topics.

Development of New Hardware for Robotic and Human Studies

During the past year, we have completed a high-precision glass smooth actuator (GSA) test bed that was used for experiments in human perception and dual actuator control. A force sensing fingertip has been integrated with a 3-degree-of-freedom force exerting device (the PHANToM) to provide an apparatus for conducting force-based touch perception experiments. A passive three-finger hand, known as the Talon, has been instrumented with a coarse array of tactile sensors and multiple grasp force sensors for use in grasp quality assessment experiments. A high-performance 8x8 tactile array has been developed and interfaced. Several one- and two-axis force sensors were fabricated for experiments to determine mechanical properties of the human fingerpad as well as ability of humans to control forces. For psychophysical and neurophysiological experiments on tactile information obtained through tool contact, an instrumented stylus was constructed to measure three-axis force signals. An apparatus capable of measuring human fingerpad volume changes to a resolution of 1/100 ml was built to investigate the compressibility of the fingerpads. An "Instrumented Active Object" which could be made to expand or collapse rapidly under computer control was developed and used in experiments on human grasp control.

Robot Sensing, Actuation, Control and Planning

Sensing and Interpretation of Tactile and Contact Force Information

Significant effort in the development of sensors and interpretation methods has been undertaken at both Harvard University and at MIT's Artificial Intelligence (AI) Laboratory. Dr. Howe's group at Harvard has developed greatly improved tactile array sensors. Work on the manufacturing process has created sensors that are easily made with a photolithography-etching process. The resulting sensors are thin and flexible and easily attached to a wide range of robot finger tip shapes and sizes. In combination with new readout and signal processing electronics, this has resulted in very fast and sensitive sensors: less than 0.1 gram noise per element with complete readout of an 8x8 array in 5 msec.

At the Artificial Intelligence Laboratory, we have focused on perception of contact states observable from contact temporal force data. This work has progressed in four areas. Low frequency measured force data reflects information about contact constraints. Estimating the form of these constraints from force and motion measurements is essential to unstructured robot force control. A procedure was developed for estimating the directions of constraint, and the geometric parameters defining the constraint equations, for points in a Cartesian space and for the interaction of planar polygons. These results may be applied to three-dimensional problems by taking planar slices of three-dimensional space.

A procedure using the force and motion measurement residual from the sensor was also developed for selecting the type of the contact constraint. After each new force and motion measurement, the best estimate of the geometric parameters can be computed. These can then be used in conjunction with an assumed form of the contact constraints to compute a measurement error or residual. By testing the size of the error for all possible measurement models, the model which best fits the data can be determined.

A theoretical foundation and set of definitions for representing the forces and motions in contact tasks was defined. The forces and motions are represented as a network of measurement models (task network). Each node in the network represents a particular measurement model. A set of points in configuration space can be associated with each and every node. The points in configuration space then define the connections in the network. These definitions provide a starting point for learning and computing networks for given tasks from models of the geometry and desired actions.

A task state observer was developed that determines the probable current state of a task given a task network for the task. An observer was coded for a network and set of measurement models for the energy in the derivative of the strain energy. This signal is a good representation for high frequency short duration events. The observer was implemented in real-time and achieved 72 percent labeling accuracy on a test data set. A simple demonstration event driven robot program was also written. The robot program uses the most probable robot state, determined by the observer, to select an action.

High-Precision Actuation

The high precision actuator project has designed and built a prototype glass smooth actuator (GSA).

The basic design relies on a dual actuator concept, in which a small actuator runs a high-frequency torque control loop, and a large actuator creates a low frequency force bias using a large spring in the transmission. The GSA has been used both as a stimulation device for psychophysical experiments and as a platform for advancing robot actuator concepts. It has been used in an initial set of experiments in human force control in which the subject attempts to track a moving endpoint, while exerting constant force. Another set of experiments is underway in which the actuator follows a moving environment, while acting to maintain constant force. The current performance of the GSA is very good. The actuator has a force control dynamic range of 500:1. It is also capable of controlling force at 40 Hz bandwidth.

Grasp Quality Assessment

In conjunction with Dr. Salisbury's group, Dr. Pratticizzo from the University of Pisa (while on a five-month fellowship at the MIT Artificial Intelligence Laboratory) has developed an approach for measuring the quality of whole hand grasps. The work combines a solid theoretical approach to measuring the "distance" from losing the grasp with a practical computational approach to real-time measurement of grasp quality. The approach utilizes contact geometric and friction information to define the maximum possible grasp quality; with the measurement of contact forces, it also permits assessment of the instantaneous quality as a percentage of the maximum attainable quality. The approach provides a means for estimating the maximum quality possible even in the absence of full sensory information, provides a method for optimizing internal grasp forces and can be used to specify limits on dynamic motions to avoid grasp failure. Real-time processing of touch and force sensor data from the talon hand has been implemented to demonstrate online grasp quality assessment.

Grasp Gaits

We have developed an approach for planning multi-finger grasp gaits. Truly dexterous manipulation of objects with the fingers requires repeated sequences of moving the object and repositioning of the fingers. Our approach permits planning sequences of repositioning and regrasping an object with two and three finger grasps. It considers loss of mobility due to figure range limits and loss of constraint due to friction and force closure limitations. The enormous search tree has been sufficiently reduced by a series of planning heuristics that real-time planning for rotation of an unknown planar object is possible. A real-time graphic simulation permits viewing progress of the planned actions.

Robot Control

Professor Howe's group at Harvard has implemented a Cartesian stiffness controller on the two-fingered Planar Manipulator in the Harvard Robotics Laboratory. This controller permits accurate control of the apparent stiffness and the center of compliance of a grasped object. The controller will be used in planned experiments on the role of tactile information in real-time control of precision manipulation tasks.

Human Perception and Control

The Human Haptics Group at MIT has investigated (1) the mechanics of the human fingerpad and its role in tactile sensing, and (2) human perception and control of forces exerted by the fingerpads. The former, described earlier and partially supported by this grant, includes experiments to determine the mechanical behavior of the fingerpad under normal and shear forces, as well as the compressibility of the fingerpad. Nonlinear lumped parameter models were developed and finite element analysis of multilayered 3D fingertip models were carried out to match biomechanical and neurophysiological data.

To quantify human perception and control of forces, experiments were conducted 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 "Instrumented Active Object". The data from each of these experiments have been analyzed with the viewpoint of developing engineering specifications of human haptic performance. The Yale group has performed psychophysical and neurophysiological experiments on perception and control of tactile information obtained by touching an object with a tool. An instrumented stylus was used in both active and passive touch experiments on the discrimination of softness by human subjects and in recording cutaneous neural responses to the same stimuli applied to monkey fingerpads. 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. In these experiments, the finger moves by only fractions of a millimeter, 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 three 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 (GSA) to measure the ability of human subjects to maintain a constant force (0.1 to 0.8 N) 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 two 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).[64] 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 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 percent, 10 percent, 20 percent 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.

---

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

A curious phenomena 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 phenomena, 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.

Grasp Control of an Instrumented Active 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 "Instrumented Active Object."[65] 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 (three subjects; ten trials per grasp force value per subject), whereas they held it in increasing number 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 in which the subject dropped the object, oscillations in force applied by the middle finger began at 60 ms, and those oscillations in force applied by the thumb began at about 80 ms. Both these kinds of oscillations continued until the object was dropped at about 100 ms. These data are being analyzed further for a deeper understanding of the dynamics of grasping.

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

The capacity of humans to discriminate the softness of rubber specimens differing in compliance by means of a tool (stylus) was measured under a variety of experimental conditions. Under active touch (AT), the observer held the stylus in a pinch grip and either tapped, pressed or bounced it against the rubber specimens during tests requiring either ranking or pairwise discrimination. Under passive touch (PT), one end of the stylus indented the stationary fingerpad with a base force maintained by a torque motor while the other end was tapped, pressed or bounced (dropped onto) by the rubber specimen. Discrimination of softness was superior during AT than during PT for both tapping and pressing, suggesting that proprioceptive cues contribute to optimal tactual performance. The fact that discrimination was more accurate for tapping than for pressing under PT suggests the importance of tactile cues generated by the initial impact of the stylus against the surface of the rubber. Bouncing the specimen (attached to a pivoted arm) against the stylus under PT produced characteristic oscillatory contact forces on the skin that subjects could readily discriminate indicating the usefulness of vibratory signals generated by the impact of a lightly held tool against an object.

The same stimuli used in the PT tests in humans were applied via the stylus to the fingerpad of the anesthetized monkey. Evoked nerve impulse activity was recorded electrophysiologically from slowly-adapting type I and rapidly-adapting type I (RA) and type II (PC) mechanoreceptive peripheral nerve fibers. Increases in the hardness of the specimen tapped or bounced against the stylus elicited nerve-impulse trains with shorter interspike intervals and shorter durations for all three fiber types. Fiber responses of SAs during pressing were related only to the force rate (velocity of pressing) and did not discriminate between specimens. Thus, the psychophysical measures can be accounted for by the responses of cutaneous mechanoreceptors. An implication of this research for the development of robotic hands is the need to develop sensors and decoding algorithms for transducing and recognizing dynamic changes in contact force produced by the impact of a tool against an object.

[65] S.P. Karason and M.A. Srinivasan, "Human Grasp Control of an Instrumented Active Object," submitted to the 1995 ASME winter annual meeting, San Francisco, California.

1.2.6 Publications

Annaswamy, A.M., 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.

Annaswamy, A.M., and M.A. Srinivasan. "The Role of Compliant Fingerpads in Grasping and Manipulation: Identification and Control." Institute of Mathematics, Springer Verlag. Forthcoming.

Bicchi, A., and D. Prattichizzo. "A Standard Form for the Dynamics of General Manipulation Systems." Submitted to the IEEE International Conference on Robotics and Automation, 1995.

Bicchi, A., P. Petreschi, and D. Prattichizzo. "Control of Force Decomposition in Enveloping Grasping." In *Proceedings of the Symposium on Robot Control,* 1994.

Bicchi, A., and D. Prattichizzo. "Controllability of Whole-Arm Manipulation." 33rd CDC, forthcoming.

Dandekar, K., and M.A. Srinivasan. "Tactile Coding of Object Curvature by Slowly Adapting Mechanoreceptors." In *Advances in Bioengineering.* Ed. M.J. Askew. Chicago: ASME, 1994, BED-Vol. 28, pp. 41-42.

Dandekar, K., and M.A. Srinivasan. "A three-dimensional Finite Element Model of the Monkey Fingertip for Predicting Responses of Slowly Adapting Mechanoreceptors." ASME summer annual meeting. Beaver Creek, Colorado. Forthcoming.

Eberman, B., and J.K. Salisbury. "Application of Change Detection to Dynamic Contact Sensing." *Int. J. Robotics Res.* 12(5): (1994).

Gulati R.J. and Srinivasan M.A. "Human Fingerpad under Indentation. I: Static and Dynamic Force Response." ASME summer annual meeting. Forthcoming.

Jandura, L., and M.A. Srinivasan. "Experiments on Human Performance in Torque Discrimination and Control." In *Dynamic Systems and Control,* Ed: C.J. Radcliffe. Chicago: ASME, 1994, DSC-Vol. 55-1, pp. 369-375.

Karason, S.P., M.A. Srinivasan, and A.M. Annaswamy. "Tactile Sensing of Shape." *Center for Information Driven Mechanical Systems (CIDMS) Workshop Proceedings*, Dept. of Mech. Eng., MIT, September 1994.

Karason, S.P., and M.A. Srinivasan. "Human Grasp Control of an Instrumented Active Object." Submitted to the 1995 ASME winter annual meeting.

LaMotte, R.H., M.A. Srinivasan, C. Lu, and A. Klusch-Petersen. "Cutaneous Neural Codes for Shape." *Can. J. Physiol. Pharm.* 72: 498-505 (1994).

LaMotte, R.H., J. Zhang, and J. Yan. "Tactile Signals Generated by a Tool Used to Discriminate the Softness of Objects." *Soc. Neurosci. Abstr.* 20: 1381 (1994).

Peine, W.J., J.S. Son, and R.D. Howe. *A Palpation Device for Artery Localization in Laparoscopic Surgery.* Harvard University, Robot Sensor Laboratory Technical Report, September 1994.

Petreschi, P., D. Prattichizzo, and A. Bicchi. "Articulated Structures with Tendon Actuation for Whole-Limb Manipulation," *Proceedings of the 1994 International Conference on Robotics and Automation,* San Diego, California, May 1994.

Roby, F.L., 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.

Son, J.S., E.A. Monteverde, and R.D. Howe. *A Tactile Sensor for Localizing Transient Events in Manipulation.* Harvard University, Robot Sensor Laboratory Technical Report, May 1994.

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.* Forthcoming.

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

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

Thesis

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

1.3 Virtual Environment Technology for Training

1.3.1 Research Testbed

Sponsor

U.S. Navy - Naval Air Warfare Center Training Systems Division
Contract N61339-94-C-0087

Project Staff

Nathaniel I. Durlach, Dr. David Zeltzer, Dr. Jeng-Feng Lee, Walter A. Aviles, James H. Bandy, Rakesh Gupta, Dorrie Hall, Jonathan Pfautz, Nicholas Pioch

During the second year of the Virtual Environment Technology for Training (VETT) program, the major goal of the Testbed project was to continue development of the basic facilities and architecture of the Testbed so that meaningful training experiments could be initiated. The following efforts were completed or are underway during the second year.

Implementation of Research Support Methods and Tools

Testbed personnel have worked closely with personnel of the Satellite projects in an effort to refine the software and hardware interfaces that each Satellite will require to conduct experimental research using the Testbed. We have recently completed the Phase-1 implementation of the Experimenters' Interface (EI), a distributed software package that will allow experimenters—either VETT investigators, or personnel from other military or government laboratories—to make use of the VETT Testbed as a research tool.

Performance Characterization

Test hardware and software for measuring the performance of the Testbed have been completed. Using these tools the performance of the Testbed (e.g., graphics, auditory, and haptic update rates and delays) has been precisely measured. This is important to characterize the system itself for development purposes, but it is also necessary to provide baseline data for Satellite experiments which will be run on the Testbed.

Development of Requirements and Specifications for Software Architecture and Multimodel VE Database Required to Support Physically-based Modeling of Objects and Actions within Virtual Environments

In collaboration with the Haptics and Sensorimotor groups, the design of a common representational framework is underway, beginning with physically based, articulated, rigid bodies. This representational framework will be utilized to model and render the visual, auditory, and haptic interactions which characterize virtual environments.

Upgrade Testbed Graphics Capabilities and Performance

A RealityEngine2 graphics subsystem was purchased and integrated into our SGI Onyx system. This system not only allows improved graphics performance but also increases compatibility with Naval Air Warfare Center Training Systems Division (NAWCTSD) systems. The Core Testbed VE-prototyping software, known as 3D, now makes use of the anti-aliasing capabilities of the RealityEngine2. In addition, work is currently underway to utilize more than one processor within our Onyx system to support graphics and general computation. As one aspect of this, Core Testbed personnel are exploring the substitution of the SGI Performer software library for the SGI GL graphics library used in the 3D system. The Performer library directly supports the use all RealityEngine2 features and the use of multiple processors for graphics processes. A Virtual Research VR4 head-mounted display was recently purchased and has been integrated into the Testbed. This LCD-based HMD has a larger horizontal field-of-view ($\sim 40^\circ$ horizontal) than our previous HMD.

Integration and Evaluation of Improved Haptic Interfaces, Speech Interaction Systems, and Auditory Displays

The PHANToM haptic display device was incorporated into the Testbed, and several demonstration applications have been implemented. A BBN Hark speaker-independent, connected speech recognition system was purchased and integrated. The auditory event representation, modeling, and rendering process received particular attention during the second year. As part of the aforementioned multimodal object database efforts, initial methods of representing and modeling acoustic events have been developed. In addition, an audio-event server and sound spatialization system were developed.

The audio-event server allows the playback of prerecorded audio sounds (i.e., based either upon recordings or generated by physical models). Hundreds of sounds may be simultaneously generated, mixed, and output on eight physical output channels. These output channels in turn are spatialized using commercial sound spatializers.

A Crystal River Engineering BEACHTRON 2-channel sound spatializer has been integrated into the Core Testbed. Three new spatialization systems (the CtronII and the Gargantutron from Crystal River Engineering and the system from Tucker-Davis Technologies) were acquired under Grant N61339-93-C-0104. The goal of this grant was to create a "next-generation" system with expanded capabilities to further the field of sound spatialization for auditory virtual environments. Since no single system appeared capable of delivering all of the envisioned needs of the field (e.g., enhanced computational throughput, increased internal storage, use of off-the-shelf rather than special-purpose hardware, greater programmability, etc.), three solutions were simultaneously developed. The Gargantutron, based on mature hardware and software, was developed to ensure continuity between existing technology and new approaches while increasing throughput and effective storage by a factor of four. The Gargantutron will be used primarily for generation of "natural-sounding" spatial sound, including environments with reflective surfaces. The CtronII was developed using off-the-shelf DSP cards and frequency-domain filtering techniques. As such, the CtronII should prove to be a fairly inexpensive system for the simulation of relatively simple echoic environments and a proof of concept for frequency-domain representations of HRTFs in real-time environments. Finally, the system developed by Tucker-Davis technologies is highly modular and extremely flexible.

All three of the new sound spatialization systems will be tested to ensure that they meet their design specifications. Following this initial evaluation, the systems will be employed in various projects under the VETT program. Both the Gargantutron and the CtronII will be used primarily in generating natural-sounding spatial sound for projects running on the Core Testbed. While developed primarily for the simulation of natural-sounding acoustic environments, the Tucker-Davis system's flexibility makes it an ideal tool for use in further auditory-motor experiments in the Sensorimotor Loop project.

Implementation of the Officer of the Deck Training Application and Pilot Experiments

The first training system under development on the Core Testbed involves training a submarine officer of the deck (OOD) to perform in-harbor navigation on a surfaced submarine. This is a task that involves a dozen or more personnel, most of whom make up the below-deck piloting team that operates periscopes, radar and sonar systems, and performs basic navigational calculations. The OOD, stationed on the conning tower of the surfaced boat during this task, is essentially responsible for monitoring the progress of the boat as it negotiates a marked channel. This officer verifies that the navigational suggestions of the piloting team are correct and that the course of the submarine is indeed conforming to the navigational plan.

The U.S. Navy, in fact, currently has simulator systems for the below-deck piloting team, but there is no corresponding simulator or formal training program for junior officers who must learn to perform this task— essentially in an informal on-the-job apprenticeship. As part of the VETT program, a VE system has been developed that allows an OOD trainee to view a particular harbor and associated waterway through a head-mounted display, receive spoken reports from a simulated piloting team, give spoken commands to the helmsman, and receive verbal confirmation of command execution from the helm. The training effectiveness of the system—including instructional aids unique to VE systems—will be evaluated experimentally with a small number of subjects.

The design of the OOD application has been accomplished with the close cooperation of the Training Satellite at BBN, and the members of this group have designed the experimental program for evaluating the OOD training application.

1.3.2 Sensorimotor Satellite

Sponsor

U.S. Navy - Naval Air Warfare Center
Training System Division
Contract N61339-93-C-0055
Contract N61339-94-C-0087

Project Staff

Nathaniel I. Durlach, Professor Richard M. Held, Dr. Thomas E.v. Wiegand, Dr. Wendelin L. Sachtler, Dr. Barbara G. Shinn-Cunningham, Evan F. Wies, David C. Lossos, Susan E. Born, Stephen V. Baird, Lukasz A. Weber, Erika N. Carmel

Figure 1. The virtual workbench concept in schematic form.

Work on the sensorimotor satellite in the VETT program is directed towards evaluating human response to alterations in sensorimotor loops associated with the use of virtual environment (VE) technology. Even in cases in which the design goal is complete realism, alterations in the form of distortions, time delays, and statistical variability ("noise") are inevitable (1) in the presentation of information to the human operator (because of limitations in displays), and (2) in the sensing of the operator's responses via the controls in the human-machine interface. Such alterations will occur in all sensorimotor loops involved in the system, including, for example: (1) those associated with visual and proprioceptive sensing, and motor control of hand position and posture; and (2) those associated with visual and auditory images of computer generated scenes and motor control and proprioceptive sensing of head position and orientation. Furthermore, in some cases, it appears that specific training goals can be achieved more efficiently by purposefully introducing distortions to highlight or magnify certain features or relationships in the training situation. In these cases, as well as those in which the alterations are the result of technological limitations, it is essential to determine the human response to the altered characteristics of the sensorimotor loops involved.

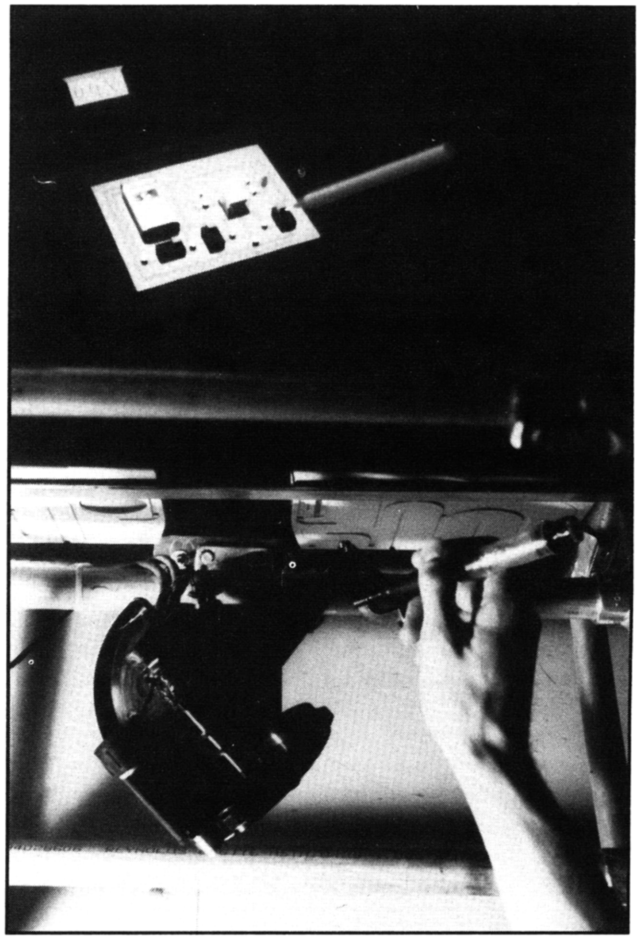

Figure 2. The virtual circuit board display and PHANToM™ unit mounted below the mirror.

In the past year, we have addressed these issues in three areas. First, we have continued our basic psychophysical experiments in the area of adaptation to alterations in the visual-proprioceptive sensorimotor loop (using the apparatus and paradigm described in *RLE Progress Report No. 136*). Second, we have completed a set of experiments elucidating a model of adaptation to alterations in the loop involving auditory localization.

Third, we have begun a new area of research utilizing a VE system that includes an alternative man-machine interface: the virtual workbench. The multimodal virtual workbench was developed to both study human sensorimotor capabilities in near-field space and to explore cognitive performance and training in complex tasks involving spatialization of information in VEs. The system combines state-of-the-art haptic feedback and visual display technologies with audio capabilities to create an alternative "front-end" interface for the core VETT testbed.

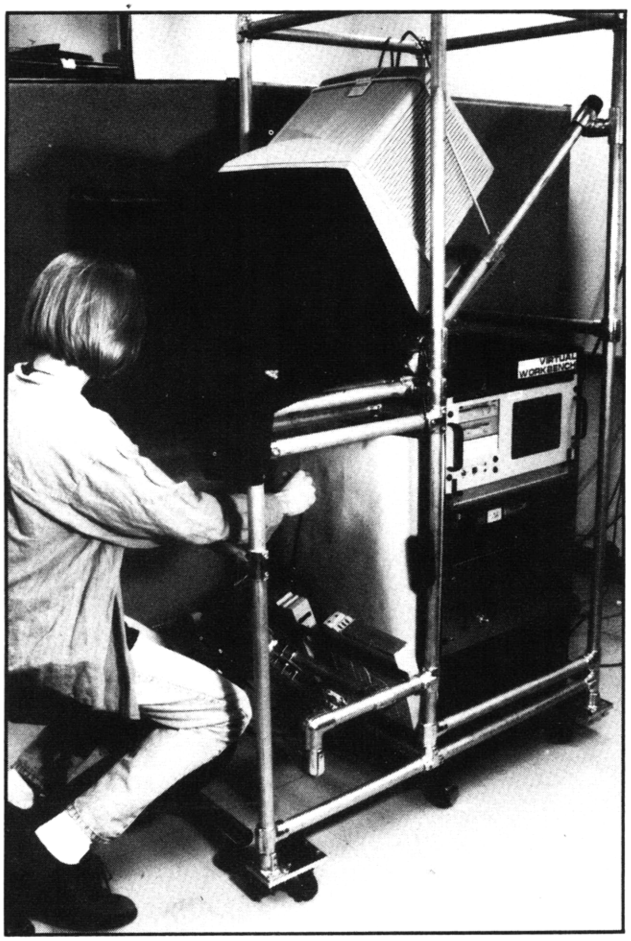

Figure 3. Photograph of subject seated at the virtual workbench.

Currently available HMDs and their associated tracking and graphics rendering systems operate at the limits of current technology. Despite their great cost, these systems are still insufficient to present a virtual world free of performance-reducing and nauseogenic effects. In the virtual workbench, we combine visual display and haptic feedback technologies into a form in which the user looks and reaches into a virtual space (as opposed to trying to immerse the entire body). In this way, we are able to sidestep many of the technical shortcomings of HMDs, enabling the virtual workbench to serve as a precision instrument for measuring human sensory and cognitive performance. The virtual workbench was specifically designed for precise overlay of visual and haptic spaces, providing a cohesive world percept. The effects (on human performance) of deviations from precise registration of the spaces, as well as distortions, delays, and other manipulations (e.g., modification of forces in haptic space), will be studied using this system. In addition to these sensorimotor effects, we are preparing experiments designed to examine the influence of the manipulations mentioned above on cognitive performance (using memory and reasoning tasks). Development efforts are also being directed toward producing a highly veridical training task involving manual interaction with virtual electronic circuit boards and components.

1.3.3 Haptics Satellite

Project Staff

Dr. J. Kenneth Salisbury, Dr. Mandayam A. Srinivasan, Nathaniel I. Durlach, Dr. Lynette A. Jones, Dr. David L. Brock, Gerald L. Beauregard, Hugh B. Morgenbesser, Hong Z. Tan, Craig B. Zilles

In the past year, we have developed device hardware, interaction software and psychophysical experiments pertaining to haptic interactions with virtual environments. Two major devices for performing psychophysical experiments, the Linear and Planar Graspers, have been fitted with additional sensors for improved performance. Another haptic display device developed previously, the PHANToM, has been used to prototype a wide range of force-based haptic display primitives. The three devices were used to perform psychophysical experiments. Copies of the PHANToM, along with demonstration software, have been delivered to the sponsor and the core testbed. In addition, two full-scale tool-handle haptic interface devices have been designed and constructed.

We had previously developed software for the demonstration of haptic displays with the PHANToM, such as interactions with virtual pushbuttons, sliding friction and surface texture displays, virtual screw turning, and virtual grasping as well as two-person haptic "pong" with two PHANToMs. Additional software was developed for visual and haptic rendering of deformable toroids and needle biopsy procedures. Different algorithms for rendering Coulomb and static friction in haptic displays were developed. Stroking of surfaces with varying compliance in the normal direction was demonstrated and experiments revealed a need for preserving passivity of interaction. Stable two-finger grasping and stacking of two virtual blocks was demonstrated with two tool-handle haptic interfaces. Network communication between a haptic renderer (Pentium PC) and graphics renderer (Silicon Graphics) was implemented. The development and specification of a software format to facilitate standard multimodal representation of convex and concave polyhedral objects was started. This format is necessary to allow consistent visual and haptic rendering of objects in virtual environments. Multimodal rendering of a triangular meshed polyhedron of moderate complexity (200 facets) was demonstrated.

Software was also developed for the conduction of psychophysical experiments with the linear grasper, planar grasper and the PHANToM. 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. A variety of haptic rendering algorithms for displaying the shape, texture and friction of solid surfaces have been implemented on the PHANToM. We completed a series of psychophysical experiments with the Linear Grasper to characterize the ability of the human haptic system to discriminate fundamental mechanical properties of objects.[66] These included several experiments to measure the manual resolution of viscosity and mass. In addition, experiments were also performed to study the influence of mechanical work and terminal force on the perception of compliance.

Utilizing the planar grasper, we have completed a set of multimodal psychophysical experiments investigating the influence of visual information on the haptic perception of stiffness in virtual workspaces. Results indicate that the perception of stiffness for objects like virtual pushbuttons can be significantly altered by presenting visually skewed positional information to the subject. We have initiated psychophysical experiments using the PHANToM to determine the feasibility of various haptic display algorithms for shape, texture and friction of solid surfaces. 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 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. Our psychophysical experiments are aimed at characterizing the effectiveness of such computationally efficient simulations and rendering algorithms in conveying the desired object properties to the human user.

Publications

Beauregard, G.L., M.A. Srinivasan, and N.I. Durlach. "Manual Resolution of Viscosity and Mass," Submitted for presentation at the ASME winter annual meeting, San Francisco, California, 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 Research Council, National Academy Press, 1994.

Srinivasan M.A. "Virtual Haptic Environments: Facts Behind the Fiction," *Proceedings of the Eighth Yale Workshop on Adaptive and Learning Systems, Center for Systems Science,* New Haven, Connecticut: Yale University, 1994.

Tan, H.Z., and 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.* Forthcoming.

Tan H.Z., M.A. Srinivasan, B. Eberman, and B. Cheng. "Human Factors for the Design of Force-reflecting Haptic Interfaces." In *Dynamic Systems and Control,* Ed. C.J. Radcliffe. Chicago: ASME, 1994, DSC-Vol.55-1, PP. 353-359.

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

Sponsor

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

Project Staff

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

The human/machine processing system (HMPS) program began August 1, 1993. The overall objec-

[66] 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. Psychophys.*, forthcoming; G.L. Beauregard, M.A. Srinivasan, and N.I. Durlach, "Manual Resolution of Viscosity and Mass," submitted to the 1995 ASME winter annual meeting, San Francisco, California.

tives of this research, as discussed in the proposal, are to:

- explore general techniques for multisensory interaction with sonar data in light of an acoustic propagation model,
- develop a working human-machine processing system (HMPS) testbed embodying the most promising of these techniques, and
- evaluate the effectiveness of the experimental HMPS.

Nearly all the first year objectives have been met. A significant effort was required to become familiar with naval sonar operations; interviews with domain experts (i.e., sonar technicians) and a task analysis based on these interviews were required before system design could begin. At present, the user-centered system design of the HMPS is largely accomplished, and HMPS implementation is underway.

An extensive literature review and a survey of available undersea acoustic propagation models were conducted; this included discussions with undersea acoustics experts at MIT and elsewhere. No extant models were deemed appropriate for the HMPS project. The ray tracing paradigm has been identified as the most appropriate and effective technique for computing undersea sound propagation. A modeling and simulation package based on modifications to ray tracing algorithms reported in Ziomek[67] is currently being implemented.

In order to design a human/machine interface for improved sonar display, it is necessary to understand and analyze the tasks actually performed by sonar crews. To this end, a number of interviews have been conducted with two domain experts, both active-duty sailors, one of whom has served as sonar supervisor, and the other as a sonar technician, on a U.S. Navy fast attack submarine. Based on these interviews, we have developed a task analysis of the functions performed by the sonar supervisor, and we have subsequently designed the functional architecture of the HMPS. In addition, we have been in contact with a number of other submarine officers in the course of the VETT project, and their input has been valuable for the HMPS work as well.

We have chosen to focus the HMPS on the requirements of the fast attack submarine, rather than a ballistic missile submarine or a surface ship, for the following reasons. First, a ballistic missile submarine is a strategic asset and does not pursue offensive tactical operations. Sonarmen on a fast attack submarine, therefore, have a range of offensive tactical operations to be supported by an HMPS that are simply not encountered on a ballistic missile submarine. Surface ships, unlike either class of submarines, can often call on support from aircraft, other naval vessels and satellites in support of tactical operations. The fast attack submarine, however, must operate as a self-contained sensor and weapons platform in a complex and noisy undersea environment.

Once an acoustic contact is detected on a fast attack sub, it is the responsibility of a single individual—the sonar supervisor—to form a coherent picture of the immediate tactical situation, and to make recommendations to the conn regarding control of the vessel. In addition to viewing "waterfall" and similar displays and interacting with the sonar technicians in the sonar shack, the sonar supervisor must construct a "mental model" of the tactical situation with essentially no supporting documentation, hardware, or software. It all must be done "in his head", based on his knowledge and experience. The purpose of the HMPS, then, is to provide tools for enhancing situational awareness and to support "what if" simulations of sensor, vessel, and weapons operations.

We have obtained undersea terrain data from several sources, and we have selected a data set appropriate for initial HMPS implementation. In addition, we have acquired several geometric models of submarines which have also been integrated into the HMPS. Furthermore, we found that ocean eddy data is available from researchers at Woods Hole.

The initial implementation of the HMPS is in progress. The ray tracing model is being coded; a large effort has gone into programming an efficient ray tracing package and programming efficient graphics modules to display ray tracing results and acoustic shadow volumes.

Once the initial HMPS implementation is complete, work will begin on a multimodal operator interface. We expect this work to be underway late in the third quarter of the second year, with subject evaluations to begin in the fourth quarter, as called for in the original proposal.

---

[67] L.J. Ziomek, "The RRA Algorithm: Recursive Ray Acoustics for Three-Dimensional Speeds of Sound, *IEEE J. Oceanic Eng.* 18(1): 25-30 (1993).

We have identified sources for platform and biological signatures, and we are prepared to incorporate such acoustic databases into the HMPS. As time permits, we may investigate cooperative antisubmarine warfare (ASW) techniques. In particular, we will be able to simulate remote active sources which serve as acoustic illuminators for the passive sensors of an attack submarine.

1.5 Development of Inertial Head-Tracking Technology

Sponsor

National Aeronautics and Space Administration/
Ames Research Center
Grant NCC 2-771

Project Staff

Eric M. Foxlin, Nathaniel I. Durlach

The inertial head-tracker program began in September 1992. Our objective is to develop a new head-tracking technology for HMD applications which offers large working volume, high fidelity, and unencumbered operation.

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, the speed of these systems is limited by the speed of sound, and they are sensitive to acoustic interference.

Our system departs from the prior state-of-the-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, making 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. Position can be tracked more accurately than orientation by optical and acoustic systems, but provides greater difficulties for inertial systems. We propose to build a system which combines an inertial orientation tracker with a hybrid acoustic-inertial position tracker because this would overcome most of the problems of either technology used alone.

1.5.1 Orientation Tracker First Prototype

In the first year of the grant, an initial prototype orientation tracker was built. The tracker consisted of three Systron-Donner GyroChips mounted on an aluminum bracket, together with a fluid inclinometer for drift compensation. It was large and heavy, but provided a convincing demonstration of the potential usefulness of inertial orientation tracking. Methods were developed for bench-testing the performance of the prototype and a simple apparatus built for this purpose. Detailed testing of drift, accuracy, resolution, noise, and dynamic response began late in the first year.

Testing of the initial prototype inertial head-orientation tracker was completed early in the second grant year. Three new projects were started to further develop the inertial head-tracking concept: (1) development of a second prototype orientation tracker, (2) construction of a general tracker testing facility, and (3) six-degrees-of-freedom (6 DOF) inertial tracker concept development and research on its underlying device technologies.

1.5.2 Orientation Tracker Second Prototype

The second prototype effort is now nearing completion. Its primary purpose is to take the inertial tracker from the proof-of-concept phase through to the demonstration of a practical tracking device that can be used in real applications including auditory experiments. Several improvements were needed to accomplish this goal. The sensor assembly, entirely redesigned to incorporate a two-axis fluxgate compass for yaw drift compensation, weighs just four ounces. The electronics were improved and transferred to a circuit board which installs in an expansion slot of a PC. Finally, a variety of improvements to the software were made including the addition of rs-232 host interface capability, more systematic sensor error compensation, and a new demonstration and testing program that runs on a separate PC connected to the tracker through the rs-232 interface.

1.5.3 Calibration and Test Facility

A second new project, which is now well underway, is the design and construction of a facility for calibrating our inertial trackers and evaluating all types of head-trackers. We would like to be able to evaluate the performance of our tracker as it would behave in a real VE system, including the effect of the serial interface. In the testing procedure used to evaluate our first prototype, the testing was performed by the same program running on the same PC that integrates the sensor signals. Therefore, any delays that would result from the interface were not included in our measurements. It was also impossible to measure the performance of other common trackers such as those manufactured by Polhemus, Ascension Technologies, and Shooting Star Technologies, and thereby perform a fair comparative evaluation of our inertial tracker versus other tracking systems.

We set out to design an autonomous system which could meet several needs: (1) testing any type of head tracker by stimulating it with a very well-known mechanical input motion and comparing the tracker output with the known mechanical input; (2) obtaining calibration data for inertial trackers by applying well-known mechanical input motions and recording all the sensor outputs simultaneously; and (3) testing the dynamic performance of a complete virtual environment system by applying mechanical motions to the head-tracker and measuring resulting changes directly from the surface of the visual display using photodiodes. For the first task, it is desirable to to provide arbitrary 6 DOF mechanical input motions to identify any nonlinear interactions between the input axes. However, a full 6 DOF rate table was too expensive, so we devised other simpler means for determining the degree of nonlinearity. We designed a mechanical motion stimulus generator that is capable of applying controlled motion to any one of the 6 degrees of freedom at a time. The 6 DOFs comprise three linear translation axes for position and three angular rotation axes for orientation; therefore, one linear motor and one rotary motor, of which one may be operated at any given time, and means for mounting any head tracker in any of three orthogonal orientations on either motor, were deemed sufficient. This equipment is also adequate for the second and third tasks. The motion stimulator package needs to generate significant output up to 15 Hz for both linear and rotational motion, up to 1000 degrees rotation rate; and the linear and rotary axes must be provided with encoders accurate to 0.01 mm and 0.005 degrees respectively.

1.5.4 Six-Degrees-of-Freedom Inertial Tracker

Investigation has begun into the possibility of augmenting the inertial orientation tracker with triaxial accelerometers to track position as well as orientation. Background reading has been done in the field of strapdown inertial navigation to identify any existing theory that can be applied to the head-tracking problem. A thorough review of the state-of-the-art in accelerometer technologies has also been conducted, including both library research and extensive telephone inquiries. A renewed effort to identify small, more accurate gyros has also been made, because the orientation subsystem in the 6 DOF inertial tracker must be accurate to 1 mrad. Drawings and clay models of the five most promising combinations of sensors have been prepared in order to balance size, cost and performance. A combination has been selected to provide three axes of acceleration and three axes of angular rate with navigation-grade performance in a package of only 1.6 cubic inches and three ounces weight. A hybrid 6-DOF tracking system is currently being 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. When completed, this system will offer the highest resolution-range product ever achieved for a human motion-tracking system.

1.5.5 Publications

Foxlin, E. *Inertial Orientation Tracker Apparatus Having Automatic Drift Compensation for Tracking Human Head and Other Similarly Sized Body*. U.S. Patent Application No. U.S.S.N. 261,364, 1994.

Foxlin, E., and N. Durlach. "An Inertial Head-Orientation Tracker with Automatic Drift Compensation for Use with HMDs." Virtual Reality Software and Technology Conference, Singapore, August 23-26, 1994. In *Proceedings VRST '94*. Eds. G. Singh, S.K. Feiner, and D. Thalmann. Singapore: World Scientific, 1994.