Section 2 Sensory Communication

Chapter 1. Sensory Communication

Academic and Research Staff

Professor Louis D. Braida, Professor Richard M. Held, Lorraine A. Delhorne, Nathaniel I. Durlach, Dr. Donald K. Eddington, Dr. Susan L. Goldman, Seth M. Hall, Dr. William M. Rabinowitz, Dr. Christine M. Rankovic, Dr. Charlotte M. Reed, Dr. Mandayam A. Srinivasan, Dr. Annie H. Takeuchi, Dr. Victor W. Zue, Dr. Rosalie M. Uchanski, Dr. Patrick M. Zurek

Visiting Scientists and Research Affiliates

Dr. Richard L. Freyman, Dr. Kenneth W. Grant, Dr. Janet D. Koehnke, Dr. Jack Kotik, Dr. Neil A. Macmillan, Dr. Karen L. Payton, Dr. Patrick M. Peterson, Yun Shao, Min Wei

Graduate Students

Santosh Ananthraman, Jyh-Shing Chen, Belinda Cheng, Kiran B. Dandekar, Joseph G. Desloge, Paul Duchnowski, Joseph A. Frisbie, Eric M. Fuchs, Julie E. Greenberg, Rajesh K. Gupta, Louise Jandura, Gregory R. Martin, Joseph A. Maxwell, Matthew H. Power, Barbara G. Shinn-Cunningham, Robert W. Stadler, Hong Z. Tan

Undergraduate Students

Daniel A. Alvarez, Susan E. Bach, Maroula S. Bratakos, Bridget L. Coffman, John A. Crouch, Swaroop Gantela, Rajashi Ghosh, Andrew C. Goldish, Andrew H. Grant, Rogeve J. Gulati, Darby A. Hailes, Albert G. Hong, Mary A. Hou, Michael T. Keagy, Michael H. Lim, David S. Lum, Sandra Y. Ma, Jason Mueller, Philip P. Nadeau, Mark T. Nadelski, Charlie C. Pan, Prashun R. Patel, Diane E. Ronan, Brian A. Rubin, Sumeet Sandhu, James F. Selph, Andrew Ugarov, Jennifer Wozniak, Ross A. Yu

Technical and Support Staff

Ann K. Dix, Eleanora M. Luongo, Timothy J. Stellmach, Michael T. Tuyo

1.1 Introduction

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

1.2 Hearing Aid Research

Sponsor

National Institutes of Health
Grant 5 R01 DC00117

Project Staff

Professor Louis D. Braida, Lorraine A. Delhorne, Paul Duchnowski, Nathaniel I. Durlach, Joseph A. Frisbie, Dr. Susan L. Goldman, Gregory R. Martin, Matthew H. Power, Dr. Christine M. Rankovic, Dr. Charlotte M. Reed, Dr. Rosalie M. Uchanski, Dr. Victor W. Zue, Dr. Patrick M. Zurek, Daniel A. Alvarez, Bridget L. Coffman, Andrew C. Goldish, Michael T. Keagy, David S. Lum, Jason Mueller, Philip P. Nadeau, Mark T. Nadelski, Jennifer Wozniak

During the past year, our research has focused on (1) linear amplification, (2) the effects of token variability, (3) a computational model of speech intelligibility, and (4) aids to speechreading.

1.2.1 Linear Amplification

Adaptive Control of Gain

In our earlier work, we examined the potential of adaptive filtering for improving speech reception in noise. We found a good association between upward spread-of-masking caused by high-intensity octave bands of noise and speech intelligibility scores when test conditions were characterized by the articulation index (AI).[1] These studies support the use of the AI for estimating the effects of noise interference on speech reception both for normal-hearing and hearing-impaired groups of listeners. Given these positive results, the AI model appears to provide a means for defining a frequency-gain characteristic that will maximize intelligibility for any given case of interfering background noise.

During the past year, we have taken a closer look at two potential limitations to applying the AI model to the clinical problem of selecting hearing aid frequency-gain characteristics for listening in noise. These limitations are excess masking demonstrated by hearing-impaired subects and individual subject differences. With this in mind, we are now conducting a detailed examination of the extensive sets of narrowband noise masking patterns that resulted from our earlier studies.

Our masking pattern corpus includes data obtained from five normal-hearing subjects and seven hearing-impaired subjects. Masking patterns consisted of pure-tone thresholds measured every one-third-octave in the presence of octave bands of noise centered on 0.5, 1.0, or 2.0 kHz and were obtained as a function of masker intensity (70 to 95 dB SPL in 5-dB steps). We have compared the hearing-impaired subjects' masking patterns with the averaged patterns of the normal-hearing subjects. We found: (1) a low-frequency spread-of-masking for both normal-hearing and hearing-impaired subjects that was roughly constant for the masker levels used here; (2) shallower slopes on the high-frequency side of masking patterns for impaired listeners; and (3) for some impaired listeners, substantially greater masking within the masker band. In addition, our results show that decreasing the masker's intensity did not decrease upward spread of masking as rapidly for the hearing-impaired subjects as it did for normal-hearing subjects. These results provide an explanation for the smaller improvements in speech reception in noise demonstrated by some hearing-impaired subjects following attenuation of the frequency region containing a band of masking noise.[2]

We have begun to compare individual subject masking patterns to those predicted from algorithms such as that described by Ludvigsen,[3] which estimates masked thresholds based on the noise spectrum and audiogram. So far, we have found some cases in which Ludvigsen's model underestimates masking substantially, and we are exploring whether these deviations will cause substantial error in prediction of benefit to speech reception. This type of verification is possible only because we have both the masking patterns and the speech intelligibility for each listening condition. As part of this project, we plan to conduct a similar comparison with the recent masking model proposed by Humes et al.[4]

Another area of research has been the utilization of the AI to evaluate the effectiveness of hearing aid gain-assignment schemes which we have extended to include background noise. We calculated AIs for speech in speech-shaped noise and for speech in a high-intensity, low-frequency octave band noise, following the application of frequency-gain characteristics assigned by several popular gain-assignment prescriptions for more than 30 different audiograms. Results support the contention that varying the frequency-gain characteristic as a function of background noise is required in order to maximize speech spectrum audibility.

---

1 C.M. Rankovic, P.M. Zurek, and R.L. Freyman, "Potential Benefits of Varying the Frequency-Gain Characteristic for Speech Reception in Noise for Hearing-impaired Individuals," *J. Acoust. Soc. Am.* 90(Pt.2): 2319 (1991); C.M. Rankovic, R.L. Freyman, and P.M. Zurek, "Potential Benefits of Varying the Frequency-Gain Characteristic for Speech Reception in Noise," *J. Acoust. Soc. Am.* 91: 354-362 (1992).

2 C.M. Rankovic, P.M. Zurek, and R.L. Freyman, "Potential Benefits of Varying the Frequency-Gain Characteristic for Speech Reception in Noise for Hearing-impaired Individuals," *J. Acoust. Soc. Am.* 90(Pt.2): 2319 (1991); J.N. van Dijkhuizen, J.M. Festen, and R. Plomp, "The Effect of Frequency-Selective Attenuation on the Speech-Reception Threshold of Sentences in Conditions of Low-frequency Noise," *J. Acoust. Soc. Am.* 90: 885-894 (1991).

3 C. Ludvigsen, "Relations Among Some Psychoacoustic Parameters in Normal and Cochlearly Impaired Listeners," *J. Acoust. Soc. Am.* 78: 1271-1280 (1985).

4 L.E. Humes, B. Espinoza-Varas, and C.S. Watson, "Modeling Sensorineural Hearing Loss. I. Model and Restrospective Evaluation," *J. Acoust. Soc. Am.* 83: 188-202 (1988).

Real-time Multiband Adaptive Gain Control

The development of multiband adaptive gain-control algorithms is being pursued for two reasons. First, there is the clear desire for an automatic volume control in hearing aids that would obviate the need for the user to constantly adjust the aid manually for different acoustic environments. Multiband control is being studied because the desired compensation is often frequency dependent (in the case, for example, of head shadow). Second, our work[5] and that of others (e.g., van Dijkhuizen et al.[6]) has demonstrated the benefit of increased speech intelligibility gained, under some restricted conditions, from adjusting the frequency-gain characteristic to attenuate spectrally-localized interference. A well-designed algorithm should be capable of both controlling loudness and, when possible, improving speech reception by reducing the upward spread of masking.

This study[7] is an evaluation of two gain-control algorithms, relative to a fixed linear system, for improving intelligibility and reducing noisiness. The fact that the systems under study are time-varying has important implications for how they are implemented and evaluated. One implication is that a real-time system is preferable for the flexibility it offers in introducing transient and time-varying sources of interference. However, because we also seek to evaluate the systems with time-varying interference, this makes traditional intelligibility testing impractical. Such intelligibility tests would also not assess noisiness, the other dimension for which the systems should be evaluated.

In light of these considerations, we implemented a real-time filter bank (fourteen filters with roughly third-octave widths) with adaptive gain controlled independently in each band. A variety of interference sources, both transient, time-varying, and steady, were added to a continuous reading (of a book-on-tape) and presented as the input to each of the systems. Additionally, the input level of speech in quiet was varied. Seven subjects with moderate hearing impairment participated. In an initial session, listeners adjusted the overall level of speech that had been spectrally shaped to match their threshold curves. These adjustments served to determine the gain in the fixed linear system, as well as the target output band levels in the adaptive systems. In the evaluations, subjects rated the outputs of the systems according to both intelligibility of the speech and overall noisiness on scales from 0 to 100.

The first of the two adaptive-gain systems is called Multiband Automatic Volume Control (MAVC) because it is designed to maintain the output of each band at a fixed level, with a limit on maximum gain. Thus, the static input-output characteristic has a fixed gain to a kneepoint and is horizontal above that point. The second system, the Noise Control system, is similar to the MBAVC System, except instead of using fixed band output target levels, its band targets depend on the signal-to-noise ratios, where band SNRs were estimated using an algorithm described by Festen et al.[8] The idea here is to amplify less when the band is dominated by noise than when it is a clear speech signal. In both systems, the level estimation needed for controlling gain is done with a simple RC envelope detector with a one-second time constant.

Generally, results showed the expected effects of gain control, most of which are beneficial. The average intelligibility ratings for the gain-control systems were significantly higher than for the fixed system when speech (in quiet) is presented below the average level. Also, when the input was only noise at a relatively low level, the gain-control systems were rated to be noisier than the fixed system. When the input noise level was greater than the expected speech level, the fixed system amplified the noise above the target level, while the gain-control systems, in the steady state, presented the noise at or below the target level.

Most of the noisiness ratings for the two gain-control systems were lower than those for the fixed system when the environmental noises lasted long enough for the gain-control systems to react. The Noise Control System was rated the least noisy system in most of the high noise level conditions,

---

5 C.M. Rankovic, R.L. Freyman, and P.M. Zurek, "Potential Benefits of Varying the Frequency-Gain Characteristic for Speech Reception in Noise," *J. Acoust. Soc. Am.* 91: 354-362 (1992).

6 J.N. van Dijkhuizen, J.M. Festen, and R. Plomp, "The Effect of Frequency-Selective Attenuation on the Speech-Reception Threshold of Sentences in Conditions of Low-frequency Noise," *J. Acoust. Soc. Am.* 90: 885-894 (1991).

7 G.R. Martin, *Studies of Real-Time Multiband Adaptive Gain Hearing Aids,* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1992.

8 J.M. Festen, J.M. van Dijkhuizen, and R. Plomp, "Considerations on Adaptive Gain and Frequency Response in Hearing Aids," *Acta Otolaryn. Suppl.* 469: 196-201 (1990).

showing the effectiveness of SNR-dependent gain control.

Intelligibility enhancements with gain control are expected for noise sources whose average spectrum is different from the speech spectrum. The gain-control systems react to such noise sources by decreasing the gain more in the bands where the noise is strong, thereby reducing spread of masking. The results indicated that this expectation was realized. There was no significant intelligibility enhancement or degradation when speech plus speech-shaped noise was tested. With other noises that are not shaped like speech, intelligibility ratings were better with gain control in most instances. Those noises whose spectra were the least shaped like speech resulted in the largest increases. Generally, intelligibility ratings of the MBAVC and the Noise Control systems were similar under most conditions.

These results show that there is promise for both the Noise Control and MBAVC Systems to outperform the fixed system with respect to the intelligibility of speech and perceived noisiness of slowly varying background noises that are not spectrally shaped like speech. The results also show that there is promise for the Noise Control System to outperform the MBAVC system with respect to the perceived noisiness of slowly varying background noises that are not shaped like speech, while at the same time, not decreasing the intelligibility of the speech.

One drawback of gain-control systems is the time needed to react to changing sound levels (a time constant of one second was used). For instance, during periods of quiet, the gain for these systems would be large, and, if a noise source suddenly occurs, it will be quite loud for a second or so. To address this problem, we are planning to implement a scheme like Moore's "dual-front end" for quickly suppressing strong onsets (and releasing control quickly if the sound is a transient).[9]

Feedback Cancellation

Several signal processing techniques for reducing feedback have been proposed recently. Thus far, however, evaluations of these methods have been limited to measurements of the increase in stable gain that can be achieved. Furthermore, the conditions under which these measurements have been made are very different from one study to another. In this work, our initial goal was to implement the most promising of the existing techniques and compare them not only for stable gain but also for their effects on sound quality and annoyance. In the process of studying these systems, a novel procedure was developed, and this is also included for comparison.

The five feedback-reduction algorithms being examined are an adaptive notch filter and four variants of adaptive cancellation. The adaptive notch method had been previously proposed for use in public-address systems. This system has not been evaluated in applications to hearing aids, though it has been used in an algorithm for detecting oscillation.[10]

The adaptive cancellation methods have a common goal of estimating feedback path in order to cancel it with another out-of-phase feedback path. The primary difference among methods is in the way that the estimation is performed. In the algorithm described by Bustamante et al.,[11] a delayed version of the output signal of the hearing aid is correlated against the microphone signal. The system of Engebretson et al.[12] is the same but with the intentional addition of noise to the output as a probe signal. In both of these systems, adaptation takes place continuously. Kates' approach differs in that the system waits for oscillation to occur, then interrupts the main signal path and inserts a probe noise for feedback-path estimation. In the novel system developed as part of this work, the estimation is performed by inserting a probe noise during selected time intervals when the input signal is estimated to be low. This approach has three advantages over Kates:[10] (1) it provides a quasi-continuous update of the estimation, not simply waiting for oscillation to occur; (2) it achieves the update without losing large segments of the signal;

9 B.C.J. Moore, "How Much Do We Gain by Gain Control in Hearing Aids," *Acta Otolaryn. Suppl.* 469: 250-256 (1990).

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

11 D.K. Bustamante, T.L. Worrall, and M.J. Williamson, "Measurement of Adaptive Suppression of Acoustic Feedback in Hearing Aids," *ICASSP-89* 2017-2020 (1989).

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

and (3) because the input signal interferes with the estimation of the feedback path, it provides a better estimate by adapting only during intervals when the input signal is small.

All five methods have been implemented with a Motorola DSP96002 chip and are currently being evaluated for increased stable gain, speech quality, and user annoyance. Real-time implementation is necessary for assessing the performance of the algorithm under dynamic conditions. Increased stable gain is measured physically, while sound quality and annoyance are being determined by subjective listening tests. For system evaluation, the acoustic feedback path is simulated with an electrical feedback path. The electrical feedback path provides more control over the experimental variables, a necessity when comparing performance of the feedback suppression algorithms. This path contains a flexible multifilter for shaping the desired transfer function. Dynamic simulation is accomplished by controlled changes in the feedback path.

We are just beginning the evaluation process. It has already become evident, however, that the continuous-adaptation schemes[13] are prone to diverge, even in the absence of feedback. Inspection of these systems shows that they are adaptive recursive filters, for which special precautions must be taken to assure convergence. Both of the other adaptive schemes allow adaptation only during times when the feedback loop is opened, and thereby avoid this problem.

1.2.2 Speech Token Variability

This research consists of three components: (1) characterizing the variability in the acoustic properties of speech sounds, (2) measuring the perceptual effects of token variability on intelligibility in normal and hearing impaired listeners, and (3) developing a model that relates acoustical and perceptual effects. Progress in previous years generally was made in the first component, i.e., measurement of the physical variability in vowels and fricatives. Recently, we have made progress on the second and third components.

The stimuli used in all the perceptual experiments are repetitions of vowels in an /h/-V-/d/ context by the same speaker. The /h/-V-/d/ syllables were digitally distorted in a manner described by Schroeder (1968). Three new experiments have been performed.[14] The first experiment examined the ability of trained listeners to identify different tokens of the same vowel sound as the number of utterances or tokens was varied. For tokens of the vowel /i/, identification scores declined from 85 percent for four tokens, to 64 percent for eight tokens, and 50 percent for 16 tokens. This suggests that the perceptual differences between tokens are not sufficiently salient to permit reliable identification. The second experiment investigated the extent to which variability in the multiplicative distortion of the stimuli affected discriminability. The results show that for the vowel pair /i/-/u/, effects of the distortion are minimal compared to the effects of token variability. For the vowel pair eh-uh, on the other hand, effects of both token and distortion variation are comparable.

In the third experiment, vowel discrimination was measured for the vowel pair ae-ah as a function of the token composition of the stimulus set. In this experiment, eight tokens were used to represent each vowel sound. In general, the relative discriminability of each vowel sound was not strongly affected by the composition of the stimulus set. This permits the relative discriminability to be correlated with the physical characteristics of the stimuli that were made earlier. Linear combinations of the physical parameters computed by Multiple Regression Analysis were generally able to predict more than 60 percent of the variance in the perceptual variables.

1.2.3 Prediction of Speech Intelligibility

This project involves development and evaluation of computational methods for determining the intelligibility of speech subjected to a waveform degradation or signal-processing transformation. The methods employ a model of the human auditory system which reduces a speech waveform to a sequence of discrete symbols, each representing a prototypical vector of parameter values measured in a single frame (e.g., 10 ms) of speech. The perceptual effect of the transformation is estimated by measuring the consistency between the symbol

---

[13] D.K. Bustamante, T.L. Worrall, and M.J. Williamson, "Measurement of Adaptive Suppression of Acoustic Feedback in Hearing Aids," *ICASSP-89* 2017-2020 (1989); A.M. Engebretson, M.P. O'Connell, and F. Gong, "An Adaptive Feedback Equalization Algorithm for the CID Digital Hearing Aid," *Proceedings of the 12th Annual International Conference, IEEE Engineering in Medicine and Biology Society,* 1990, pp. 2286-2287.

[14] D. Ronan, *Further Studies in the Perceptual Effects of Multiple Speech Tokens on Vowel Resolution,* S.B. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1992.

sequence derived from an untransformed (input) speech signal and the symbol sequence derived from a transformed (output) signal. This estimate is implemented via calculation of percent transmitted information (%TI).

Previous research used the energy in each of 20 critical bands as the source of the parameter vector. During the past year, we have focused instead on a different set of 20 parameters, derived from the Ensemble Interval Histogram (EIH) model developed by Ghitza[15] at AT&T Bell Laboratories. The central idea of this model is that speech information is represented in the lengths of intervals between crossings of certain waveform-amplitude values. Implementations of this model (e.g., Ghitza, 1992; Jankowski, 1992)[16] are largely comprised of two major stages: (1) filtering speech into a number of bands, and (2) computing crossing intervals for the individual band waveforms. In the current study, a primary goal was to keep the implementation as similar as possible to that used for computing the band-energy parameter set. In particular, the new computation maintained the frequency region spanned by the bandpass filters, number of bandpass filters, frame rate of the short-time speech analysis, and total number of parameters extracted.

The %TI values obtained via this model can be compared to those obtained via the band-energy computation for each of three types of degradation: additive noise, highpass filtering, and lowpass filtering. Results for the two parameter sets were quite similar in the case of noise. At low SNRs, %TI values for EIH were higher than those for band-energy by about 2 percentage points. At high SNRs, %TI values for EIH were lower than those for band-energy by about 4 percentage points. The crossover point at which the two parameter sets gave equal results was approximately +10 dB SNR. Since the EIH model seemed slightly more resistant to noise than the band-energy model, results were also calculated for negative SNRs, which had not been previously considered. These were much lower than those for the positive SNRs, with %TI values ranging from 17.2 at -5 dB SNR to 3.6 at -20 dB SNR.

For the highpass-filtering degradation, %TI values for EIH were much lower than for band-energy. The average difference was about 17 percentage points; this was fairly uniform across cutoff frequencies. For lowpass filtering, the %TI values were again lower for EIH, but with a smaller discrepancy averaging about 9 percentage points. Differences were greater at higher cutoff frequencies than at lower ones. In fact, at the 500-Hz cutoff frequency, the %TI value for EIH was the higher one (but by less than one percentage point). Because of these differences in the highpass and lowpass conditions, the crossover frequency calculated from the EIH results was much lower than that calculated from the band-energy results: 1000 Hz vs 1750 Hz.

Across all conditions tested, %TI results for the EIH model were usually lower than those for the band-energy model. This suggests that degradations generally affected the consistency of level-crossing information more than they affected the consistency of band-energy information. The absolute numerical values of %TI are of greatest interest in the low-SNR noise conditions. In particular, it would be useful if these values provided an adequate characterization of intelligibility across all noise levels for which humans retain some reasonable speech-reception ability. However, the -20 dB SNR result, 3.6 %TI, is comparable to the 2.0 %TI obtained from randomly generated symbol sequences of the length used in these calculations. Thus, considerably more speech data would need to be processed for reliable %TI results in intense noise degradations.

However, absolute %TI values are not the only data of interest; instead, the relationship between %TI values for noise and filtering is also important. In particular, one characteristic of the result pattern for the band-energy model is that noise degradations of little perceptual significance reduced %TI as much as did filtering degradations that have a somewhat larger perceptual effect. In going from the band-energy model to the EIH model, the %TI values for filtering dropped more than did the %TI values for noise, therefore leading to a result pattern that is more perceptually realistic.

One final observation of the EIH results is the presence of particularly low %TI values for highpass filtering. One possible explanation for this can be found in the selection of the amplitude levels for which crossings are located. These levels span the

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

[16] O. Ghitza, "Auditory Nerve Representation as a Basis for Speech Processing," In *Advances in Speech Signal Processing,* S. Furui and M.M. Sondhi, eds. (New York: Marcel Dekker, 1992), pp. 453-485; C.R. Jankowski, Jr., *A Comparison of Auditory Models for Automatic Speech Recognition,* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1992.

amplitude range found in wideband speech. Highpass-filtered speech is considerably attenuated and rarely crosses the higher levels. This leads to the EIH parameters providing a poorer representation of the speech information. Also, independent of these characteristics of the results, another disadvantage to use of the EIH model is that its computation time is approximately ten times that of the band-energy model.

In addition to work on the EIH model, some supplementary computations were executed during the past year to address the issue of why very slight degradations result in %TI values that are considerably less than 100. Of these, one that seemed insightful had a complete absence of degradation at the waveform level, and instead perturbed the calculation of the symbol sequence by use of a different initial guess in the vector-quantization stage. These experiments helped to confirm the hypothesis that a substantial drop in %TI is caused solely by reassignment of prototype vectors. In some cases of particularly mild degradation, it can be beneficial to omit this reassignment and simply use the prototype vectors from undegraded speech.

Finally, there were some calculations performed to validate portions of the experimental methods. Issues of interest included whether the %TI results were strongly dependent on the particular VQ initialization method, whether accounting for waveform delays due to filtering (either for degradation purposes or within a filter bank) would affect the results, and whether the %TI measure had a reasonable relationship to percentage of exact symbol matches (i.e., a percent-correct measure). None of these uncovered any significant methodological difficulty.

1.2.4 Aids to Speechreading

During the past grant year, our research on aids to speechreading has continued to focus on (1) evaluating the effectiveness of Cornett's (1967) Manual Cued Speech system and (2) developing low bandwidth acoustic signals that enhance speechreading.[17]

Cued Speech

The system of Manual Cued Speech was developed by Cornett (1967) as a supplement for speechreading that could be taught to very young deaf children and used in day-to-day communication. In Cued Speech, eight visually distinct hand shapes are used to distinguish between consonants that are not well resolved via speechreading and four hand positions near the mouth provide distinctions within groups of vowels often confused in speechreading. During the past year, we completed our initial study of the effectiveness of Cued Speech and developed an analysis that can be used to estimate the effectiveness of automatic cueing systems based on automatic speech recognizers.[18]

As previously reported, we have documented the sentence reception ability of four highly-trained receivers of Manual Cued Speech. The availability of cues permitted near perfect reception of high-context everyday sentences spoken at near normal speaking rates, and very high scores (roughly 85 percent of the words were correct) were obtained for low-context sentences.[19] In control tests on cued and uncued sentences, selected speechreaders who had no previous experience with Cued Speech achieved word scores in the range of 24-36 percent on everyday sentences and 15-22 percent on the low-context sentences. Although the levels of speech reception achieved by the trained receivers is impressive, further analysis indicates that a significant fraction (10-20 percent) of speech segments were not correctly perceived. In a related study, we measured how well spectrogram readers and an ASR system could assign segments to cue groups for various cue systems. Well-trained spectrogram readers were able to assign six consonant cues with an accuracy of 85 percent and eleven vowel cues with an accuracy of 62 percent; an expert reader was able to achieve scores of 94 percent and 83 percent on these tasks. An automatic speech recognizer recently developed at MIT achieved accuracies of 77 percent and 65 percent on comparable tasks. In the third study, we applied the Post-Labeling Model of audiovisual integration[20] to these recognizer measurements and data on

17 R.O. Cornett, "Cued Speech," *Am. Ann. Deaf* 112: 3-13 (1967).

18 R.M. Uchanski, K.M. Millier, C.M. Reed, and L.D. Braida, "Effects of Token Variability on Vowel Identification," in *The Processing of Speech: From the Auditory Periphery to Word Recognition.* (Berlin: Mouton de Gruyter, 1992).

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

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

human recognition of consonant and vowel segments via speechreading to estimate the benefit to speechreading provided by such cues. The analysis suggests that with cues derived from current recognizers, consonant and vowel segments would be received with accuracies in excess of 80 percent. This level of accuracy is roughly sufficient to account for the scores our subjects achieved on Manual Cued Speech reception. Use of the Post-Labeling Model provides analytic means for (1) determining optimum cue groups for a given recognizer and speechreader and (2) estimating the cueing performance that might be achieved if the accuracy of current recognizers were improved in various ways, e.g., by reducing the frequency of voicing errors.

Based on these results, it seems appropriate to begin development of an automatic cueing system based on an automatic speech recognizer and a synthetic visual display of cues. Major uncertainties in the design of such a system concern the effects of recognizer errors, appropriate choice of cue groups, and specification of visual display of cue symbols. We are planning a series of experiments to evaluate the effects of recognizer errors on speech reception using an automatic cueing system. If the ASR does not identify a phoneme correctly, the displayed cue will not correspond to the spoken phoneme and therefore will be incorrect. Our strategy is to simulate the performance of automatic cueing systems that use recognizers with various error rates and patterns of error structures. These systems will use the traditional cue groups of manual cued speech so that highly trained cue receivers can evaluate speech reception.

In a preliminary experiment, we tested the percent of phonemes that could be inaccurate and still allow normal listeners to identify speech. We used only auditory stimuli, not the cues from cued speech to give a baseline for further experiments with cued speech. The stimuli were the Clarke sentences,[21] with portions corresponding to selected phonemes replaced by white noise. A given phoneme was replaced by noise in the same proportion that it was misidentified by a known speech recognition system (BBN system). If two percent of the misidentifications that the speech recognition system made were /ah/'s then two percent of the phonemes replaced would be /ah/'s. There were four conditions: 10, 20, 30, and 40 percent of the phonemes in the group of sentences were replaced with noise. Subjects heard the sentences and were instructed to guess what a word was even if they were not sure.

We calculated two measures from a preliminary test. First, we counted the percent of phonemes accurately identified. In general, a higher percent of phonemes were correctly identified in sentences with a lower percent of phonemes replaced by noise. Second, we categorized the guesses subjects made by the number of phonemes from which they differed from the original phoneme. A one phoneme change indicated that a guess was either a change of one phoneme to another phoneme /p/ to /b/, for example, or the addition or deletion of phoneme. We grouped the guesses into one, two, three, or greater than three phoneme changes. The results showed that in general the category of the substitution varied with the percent of phonemes replaced by noise. In sentences with lower percents of phonemes replaced by noise, lower phoneme substitutions were more prevalent than higher phoneme substitutions. Listeners were more likely to guess a word which was acoustically closer to the target word. Similarly, in sentences with higher percents of phonemes replaced by noise, phoneme changes were more likely to be greater than three phonemes from the target than in sentences with lower percents of phonemes replaced by noise. Listeners were less likely to guess a word which was acoustically closer to the target word, and more likely to guess a word which was semantically and syntactically appropriate.

Frequency Lowering of Amplitude Envelopes

Acoustic signals that retain much of the intelligibility of the unprocessed speech signal can be synthesized by extracting the amplitude envelopes of filtered bands of speech and using these envelopes to modulate the amplitudes of tones with frequencies at the centers of the bands. These signals successfully supplement speechreading by normal-hearing listeners.[22] However, when the center frequency of the carrier bands is lowered to match the residual hearing of listeners with severe high-frequency loss, the signals no longer successfully

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

22 M. Breeuwer and R. Plomp, "Speechreading Supplemented with Frequency-Selective Sound-Pressure Information," *J. Acoust. Soc. Am.* 76: 686-691 (1984); 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.* 92: 2345 (1992); K.W. Grant and L.D. Braida, "Evaluating the Articulation Index for Audiovisual Input," *J. Acoust. Soc. Am.* 89: 2952-2960 (1991).

supplement speechreading.[23] In response to a report by Sheft and Yost[24] that the failure of frequency-lowered amplitude envelopes to supplement speechreading may arise because listeners are unable to compare the amplitude envelopes of signals across changes in the center frequency of the signal, Annie Takeuchi has been examining the effect of systematic training on the ability of normal-hearing listeners to match amplitude envelopes of signals at different center frequencies.

The amplitude envelopes were carried by narrow-band noise signals 50-Hz wide, centered at 500, 1600, and 3160 Hz. These baseline frequencies were selected because they are the center frequencies of the filtered bands of speech from which amplitude envelopes have been extracted to supplement speechreading. Subjects were trained on a roving XAB matching-to-sample task, in which the center frequencies of the A and B signals were systematically increased or decreased in increments of 10 percent up to 100 percent and down to -90 percent from the baseline frequency, while the center frequency of the X signal was fixed at the baseline frequency.

For all three baseline frequencies, performance as measured by d' was best when there was no frequency difference between the X, A, and B signals, and d' generally decreased when any frequency difference was present. However, for all frequency differences except decreases of more than 60 percent (-5.7 Barks) from the 1600-Hz baseline frequency, d' remained well above 1.76, which corresponds to 70.7 percent correct unbiased responding on the roving XAB task, the threshold value most commonly measured in psychophysical experiments. For the 500- and 3160-Hz baseline conditions, increases and decreases in frequency produced similar results. For the 1600-Hz condition, increases in frequency produced much smaller decrements in performance relative to no frequency difference than did decreases in frequency. This study demonstrates that with systematic training, listeners are generally able to compare amplitude modulation patterns across large frequency differences.

1.2.5 Human Subjects

There have been no changes in our protocols during the past year, and we do not plan any for the coming year. These protocols were evaluated during the competitive review and have been renewed for the coming grant year.

1.2.6 Publications and Talks

Braida, L.D., P.M. Zurek, K.W. Grant, J.E. Greenberg, and C.M. Rankovic. "Research on Hearing Aids at MIT: Recent Results and Future Directions." In *Proceedings of the International Symposium on Hearing Aids and Speech Training for the Hearing Impaired,* Osaka, Japan, July 16-17, 1991. Eds. H. Levitt and T. Nitta. Forthcoming.

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

Payton, K.L., R.M. Uchanski, and L.D. Braida. "Intelligibility of Conversational and Clear Speech in Noise and Reverberation for Listeners with Normal and Impaired Hearing." Submitted to *J. Acoust. Soc. Am.*

Posen, M.P., C.M. Reed, and L.D. Braida. "The Intelligibility of Frequency-Lowered Speech Produced by a Channel Vocoder." *J. Rehab. Res. Dev.* Forthcoming.

Rankovic, C.M., P.M. Zurek, and R.L. Freyman. "Research on Adaptive Frequency Shaping." Invited lecture presented at the House Ear Institute-sponsored conference *Issues in Advanced Hearing Aid Research,* Lake Arrowhead, California, May 25-29, 1992.

Rankovic, C.M., Freyman, R.L., and P.M. Zurek. "Potential Benefits of Varying the Frequency-gain Characteristic for Speech Reception in Noise." *J. Acoust. Soc. Am.* 91: 354-362 (1992).

[23] 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.* 92: 2345 (1992); K.W. Grant and L.D. Braida, "Evaluating the Articulation Index for Audiovisual Input," *J. Acoust. Soc. Am.* 89: 2952-2960 (1991).

[24] S. Sheft and W.A. Yost, "Spectral Transposition of Envelope Modulation," *J. Acoust. Soc. Am.* 91: S2333 (1992).

Uchanski, R.M., K.M. Millier, C.M. Reed, and L.D. Braida. "Effects of Token Variability on Vowel Identification." In *The Processing of Speech: From the Auditory Periphery to Word Recognition.* Mouton de Gruyter, Berlin, 1992.

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.* Forthcoming.

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

Theses

Martin, G.R. *Studies of Real-time Multiband Adaptive Gain Hearing Aids.* S.M. thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1992.

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

Ronan, D.E. *Effects of Token Variability on Speech Intelligibility for Vowel Sounds.* S.B. Thesis, Dept. of Electr. Eng. and Comput. Sci., MIT, 1992.

1.3 Multimicrophone Hearing Aids

Sponsor

National Institutes of Health
Grant 5 R01 DC00270

Project Staff

Joseph G. Desloge, Nathaniel I. Durlach, Julie E. Greenberg, Dr. William M. Rabinowitz, Robert W. Stadler, Dr. Patrick M. Zurek

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

To enhance monaural speech reception, we envision a microphone array that resolves the incoming signals into simultaneous directional channels, followed by a coding operation that transforms these resolved signals in such a way that resolution is preserved at the perceptual level after the signals are summed for presentation to a single ear.[25] Such a system would permit the monaural listener (like the normal binaural listener) to monitor all directions simultaneously, detect and localize in the same operation, and focus on a single direction. Our current work on microphone arrays is directed toward the creation of a single directional channel containing the target signal (assumed to arise from a target source straight ahead of the listener) and reduction of interference from sources directionally distinct from the target source. Parallel processing of array outputs to achieve simultaneous multiple directional channels will be considered only after further progress on the coding problem has been achieved.

While development of adaptive arrays has continued,[26] considerable effort in the past year has gone into the design of fixed arrays.

1.3.1 Fixed Arrays for Hearing Aids

Microphone arrays with fixed (time-invariant) weights are directed at enhancing a desired signal from one direction (straight ahead) while attenuating spatially distributed interference and reverberation. Using the theory of sensitivity-constrained optimal beamforming,[27] free-field arrays of head-sized extents were studied.[28] The key parameters affecting array design and performance are the set of transfer functions from the target direction to each array microphone $\underline{H}$ (f) and the intermicrophone cross-spectral densities for isotropic noise $S_{zz}(f)$. Design variables included ori-

[25] N.I. Durlach, R.C. Corbett, M.V. McConnell, W.M. Rabinowitz, P.M. Peterson, and P.M. Zurek, "Multimicrophone Monaural Hearing Aids," RESNA 10th Annual Conference, San Jose, California, 1987.

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

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

[28] R.W. Stadler and W.M. Rabinowitz, "On the Potential of Fixed Arrays for Hearing Aids," submitted to *J. Acoust. Soc. Am.*

entation of the array, number and directionality of the microphones within the array,[29] complexity, and robustness of the required processing. For broadside orientation, a variety of arrays based on cardioid and hypercardioid microphones gave very similar performance. They can provide broadband intelligibility-weighted directivities (direct-target gain against isotropic noise) of 7-8 dB with easily implemented weights (simple scalars). For endfire orientation, similar directivities result with weights based on analog gains and pure time delays. However, with weightings that are chosen independently for each frequency, directivities up to ~ 11 dB may be practical. Because of sound diffraction, placement of arrays onto the head potentially impacts both their design and performance. *In-situ* measurements of $\underline{H}$ (f) and $S_{zz}(f)$, as well as simplified theoretical models, are suggested to explore the optimization of head-mounted arrays.

1.3.2 Intelligibility-weighted Measures of Speech-to-Interference Ratio

Many of our multimicrophone subprojects require assessment of the improvement in intelligibility provided by processing systems that are essentially linear. For this purpose, we have developed measures of speech-to-interference ratio and system gain that incorporate factors affecting speech intelligibility.[30] Such measures result from simple modifications of the Articulation and Speech Transmission Indices[31] and are applicable to the same kinds of signal degradations to which the indices apply. Under restricted conditions, such measures are linear transformations of either the Articulation Index or the Speech Transmission Index and are as accurate as these indices in predicting speech intelligibility. This is a point that we have illustrated through a validation study. Depending on the type of signal degradation, measures of reduced complexity can be employed. Such measures may prove convenient in characterizing listening conditions and performance of speech-transmission systems.

1.4 Cochlear Prostheses

Sponsor

National Institutes of Health
Contract 2 P01 DC00361[32]

Project Staff

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

The overall goals of this research are to determine and understand the potential and limitations of cochlear prostheses and develop improved prostheses. Postlingually deafened adults are implanted with the Ineraid multichannel cochlear prosthesis and participate in intensive multifactorial studies. This research capitalizes on the direct accessibility of the implanted electrode array via a percutaneous connector.

During the past year, work has focused on (1) analysis of cue integration in audiovisual speech reception and (2) alternative speech processing for improved implant performance. Progress in (1) is described below. The work in (2) is performed with Joseph Tierney, and Marc Zissmann; progress is described in Part V, Section 3, Chapter 1, section 1.7.2.

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

A series of experiments on closed-set identification of consonants and vowels were begun. The stimuli are presented via computer-controlled videodisc. Results are being obtained using vision alone, audio alone, and audiovisually. At present, most subjects are nearing completion of the testing; asymptotic levels of performance have been reached and sufficient data have been obtained to provide reliable estimates of confusion patterns for each of the three test modes. Analyses to deter-

29 W. Soede, *Improvement of Speech Intelligibility in Noise: Development and Evaluation of a New Hearing Instrument Based on Array Technology,* Ph.D. diss., Delft University of Technology, The Netherlands, 1990.

30 J.E. Greenberg, P.M. Peterson, and P.M. Zurek, "Inelligibility-Weighted Measures of Speech-to-Interference Ratio and Speech System Gain," submitted to *J. Acoust. Soc. Am.*

31 American National Standard Institute, "American National Standard Methods for the Calculation of the Articulation Index," (New York,: ANSI, 1969); H.J.M. Steeneken and T. Houtgast, "A Physical Method for Measuring Speech-transmission Quality," *J. Acoust. Soc. Am.* 67: 318-326 (1980).

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

mine the quantitative efficiency of audiovisual integration are underway. These analyses exploit an analytical framework (Braida, 1990) that quantifies how performance in a combined mode (e.g., audiovisual stimulation) is influenced by integration or interference of cues available in the isolated component modes (e.g., audio and visual stimulation alone).

1.5 Binaural Hearing

Sponsor

National Institutes of Health
Grant 5 R01 DC00100[33]

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.

Experimental research in this area has focused on identifying the stimulus variables responsible for in-head lateralization of auditory images. It is well established that the interaural relations that exist early after the onset of a sound have a strong influence on where the subsequent sound image is heard.[34] It is believed that this effect allows very good sound source localization in highly reverberant environments. Our recent work has confirmed that such effects measured using an acoustic pointer correspond well with measurements of interaural parameter resolution.[35] Further work has shown that the influence of the onset cue can extend over a relatively long (hundreds of milliseconds) subsequent stimulus, depending on both the ambiguity of the interaural cues in the subsequent stimulus and on whether it appears to constitute a separate "auditory object".[36]

1.6 Clinical Applications of Binaural Hearing

Sponsor

National Institutes of Health
Contract 7 R29 DC00428[37]

Project Staff

Dr. Patrick M. Zurek

In this project, we seek to apply our understanding of binaural advantages in speech reception[38] to the prescription of hearing aids, both monaural and binaural, and to test this understanding further. In particular, the experimental work is focused on the performance of both normal hearing and hearing-impaired listeners on tests of binaural detection, localization, and contralateral masking with various amplification patterns applied at both ears.

1.7 Tactile Communication of Speech

Sponsor

National Institutes of Health
Grant 2 R01 DC00126

Project Staff

Lorraine A. Delhorne, Nathaniel I. Durlach, Hong Z. Tan, Dr. William M. Rabinowitz, Dr. Charlotte M. Reed, Dr. Mandayam A. Srinivasan

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 commu-

---

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

34 P.M. Zurek, "The Precedence Effect," in *Directional Hearing,* eds. W.A. Yost and G. Gourevitch (New York: Springer-Verlag, 1987); P.M. Zurek, "A Note on Onset Effects in Binaural Hearing," *J. Acoust. Soc. Am.* 93: 1200-1201 (1993).

35 B.G. Shinn-Cunningham, P.M. Zurek, and N.I. Durlach, "Adjustment and Discrimination Measurements of the Precedence Effect" *J. Acoust. Soc. Am.*, forthcoming.

36 R.L. Freyman, P.M. Zurek, U. Balakrishnan, and C. Yuan-Chuan, "Onset Dominance in Lateralization," submitted to *J. Acoust. Soc. Am.*

37 Subcontract from University of Connecticut. Dr. Janet D. Koehnke, Principal Investigator.

38 P.M. Zurek, "Binaural Advantages and Directional Effects in Speech Intelligibility," in *Acoustical Factors Affecting Hearing-aid Performance,* eds. G.A. Studebaker and I. Hochberg (Boston: Allyn and Bacon, 1993.)

nication.[39] The objectives and specific aims of our research are as follows:

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

1.7.1 Basic Study of Encoding and Display Schemes

A study of joint-angle discrimination (JND) and identification of the proximal interpalangeal (PIP) joint of the index finger has been completed. During the current year, additional data have been collected on the discriminability of the PIP joint angle as a function of metacarpal phalangeal (MCP) joint position. The data indicate that JNDs for PIP joint position are independent of the MCP joint position and average 2.6 degrees. Additional data on PIP joint-angle discrimination as a function of the PIP joint reference position indicate subject-dependent results: JNDs for two subjects are roughly independent of the reference position (averaging 2.4-2.5 degrees), while JNDs for a third subject increased with greater flexion of the PIP joint.

Preliminary work has begun on a study concerned with defining the range of amplitudes and frequencies that can be used to stimulate the tactual system, as well as specifying independent perceptual channels and resolution within these channels. This work includes development of a device for stimulus presentation consisting of a plastic bar attached to a rotary motor with 2000 discrete, specifiable steps. Pilot studies indicate that this device is capable of presenting signals in the range of 1 to 100 Hz over a fairly broad dynamic range. Existing software has been modified to collect data on amplitude and frequency resolution in a one-interval two-alternative forced-choice procedure.

Work has also begun on a study of the manual discrimination of the property of thickness. A series of plates varying in thickness has been prepared for use as the stimuli in this study, and a device has been built for presenting these plates to subjects in experiments concerned with the ability to discriminate thickness as a function of the reference value. Finally, a manuscript summarizing research on the manual discrimination of length, force, and compliance has been prepared.[40]

1.7.2 Tactual Supplements to Speechreading

In the area of acoustic supplements to speechreading, work during the current year has been concerned with developing a more complete understanding of the differential results obtained for auditory and tactile presentation of acoustic-based low-bandwidth supplements to speechreading.[41] This work has included signal processing and software development for a set of psychophysical measurements of modulation resolution. In the area of articulatory supplements to speechreading, work has been concerned with designing an encoding scheme for presenting various types of articulatory information through a seven-channel vibratory array. The vibrators can be controlled through sine-wave inputs and various vibratory patterns can be used to encode articulatory features of consonants and vowels. Preliminary psychophysical studies are being conducted to determine a set of vibratory patterns that provide sufficient information for this problem.

---

[39] C.M. Reed, N.I. Durlach, and L.A. Delhorne, "Historical Overview of Tactile Aid Research," *Proceedings of the Second International Conference on Tactile Aids, Hearing Aids, and Cochlear Implants,* eds. A. Risberg and K.-E. Spens (Stockholm, Sweden: Royal Institute of Technology, forthcoming).

[40] H.Z. Tan, X.D. Pang, and N.I. Durlach, "Manual Resolution of Length, Force, and Compliance," *Winter Annual Meeting of the American Society of Mechanical Engineers,* Anaheim, California, November 8-13, 1992.

[41] M.J. Besing, C.M. Reed, and K.W. Grant, "Tactual Presentation of Low-Bandwidth Envelope Signals: Benefits to Speechreading," *ASHA* 33: 212, 1991.

1.7.3 Evaluation of Practical Aids

Three new subjects have joined our field study of tactile-aid users, bringing the total number of subjects to eight. During the current year, six of the eight subjects visited our laboratory for evaluations of their speechreading performance with the Tactaid VII. Data are now available from three separate testing sessions with one subject (JL), two testing sessions with two subjects (RM, RS), and one session with each of the remaining subjects. The evaluations have been concerned primarily with assessing speechreading ability with and without tactile devices, as well as with measuring discriminability of simple speech segments through the use of tactile devices by themselves. Thus far, tactile devices employed in the study include Tactaid VII (which all subjects in the field evaluation have received), Tactaid II (tested only on three of the subjects who had used the Tactaid II prior to Tactaid VII), and a high-performance single-channel vibrator (Alpha-M AV-6 Minishaker) employed only in laboratory testing.

Results of segmental discrimination tests conducted with Tactaid VII indicate that, averaged across subjects and across consonant and vowel pairs, performance is roughly 70 percent correct.[42] Additional data collected on one subject with Tactaid II showed that performance for both devices was approximately similar. Results of tests conducted for speechreading alone and speechreading in combination with a tactile device (using sentence and continuous-discourse materials) showed a range of performance across subjects. For example, for the reception of words in CUNY sentences, scores from all subjects ranged from 32 to 86 percent correct (averaging 53 percent) for speechreading alone and 29 to 93 percent (averaging 58 percent) for speechreading plus Tactaid VII. Of the six subjects for whom data are currently available on this task, improvements to speechreading with the use of Tactaid VII were observed for four subjects (ranging from 5 to 17 percentage points improvement), while two subjects have shown no improvements to speechreading of CUNY sentences with Tactaid VII thus far. Data from three subjects tested with Tactaid II (as well as with Tactaid VII) on this task indicate greater improvements to speechreading (by approximately 8 percentage points) when using Tactaid II compared with Tactaid VII. Some preliminary data obtained on three subjects for speechreading of CUNY sentences with the Minishaker indicate that the benefits provided by this device are comparable to those observed with Tactaid II. The results of a questionnaire assessing subjects' use of tactile devices indicate that most prefer Tactaid VII to Tactaid II. This is primarily because of the greater utility of Tactaid VII in detecting and recognizing environmental sounds in addition to the benefits it provides for speechreading.[43]

1.7.4 Continuation and Completion of Previous Work on Natural Methods of Tactual Communication

During the past year, we have published summaries of our work on the communication abilities of deaf-blind subjects who use various natural methods of tactual communication, including the Tadoma method of speech reception and tactual reception of fingerspelling and sign language.[44] A manuscript describing a series of experiments conducted to demonstrate improvements to speech reception through Tadoma using supplementary tactual information has also been published.[45]

1.8 Super-Auditory Localization for Improved Human-Machine Interfaces

Sponsor

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

---

[42] C.M. Reed, L.A. Delhorne, and N.I. Durlach, "Results Obtained with Tactaid II and Tactaid VII," *Proceedings of the Second International Conference on Tactile Aids, Hearing Aids, and Cochlear Implants,* eds. A. Risberg and K.-E. Spens (Stockholm, Sweden: Royal Institute of Technology, forthcoming.

[43] C.M. Reed and L.A. Delhorne, "Field Study of Deaf Adult Users of Tactaid II and Tactaid VII," Presentation made at the Annual Convention of the Association for Late Deafened Adults, Boston, Massachusetts, September 10-13, 1992.

[44] C.M. Reed, N.I. Durlach, and L.A. Delhorne, "Natural Methods of Tactual Communication," chapter in *Tactile Aids for the Hearing Impaired,* ed. Ian R. Summers, (Whurr Publishers Limited, 1992), pages 218-230; C.M. Reed, N.I. Durlach, and L.A. Delhorne, "The Tactual Reception of Speech, Fingerspelling, and Sign Language by the Deaf-Blind," *SID Digest* 102-105 (1992).

[45] C.M. Reed, W.M. Rabinowitz, N.I. Durlach, L.A. Delhorne, L.D. Braida, J.C. Pemberton, B.D. Mulcahey, and D.L. Washington, "Analytic Study of the Tadoma Method: Improving Performance through the Use of Supplementary Tactual Displays," *J. Speech Hear. Res.* 35: 450-465 (1992).

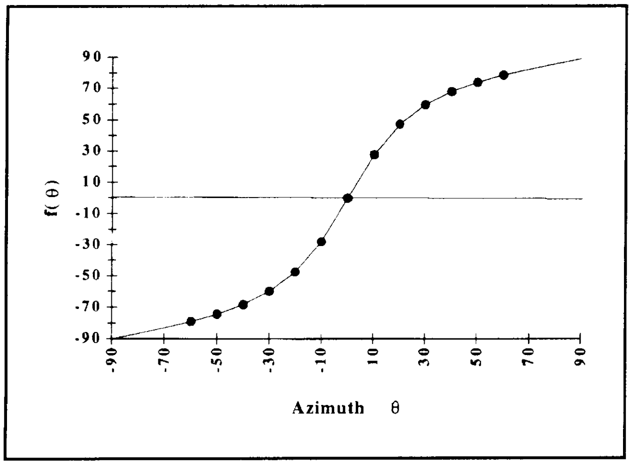

Figure 1.

Project Staff

Nathaniel I. Durlach, Eric M. Fuchs, Dr. Richard M. Held, Dr. William M. Rabinowitz, Yun Shao, Barbara G. Shinn-Cunningham, Min Wei

General background on this project was presented in pages 312-313 of the annual *RLE Progress Report Number 134*. During the past year, work has advanced along the following fronts.

1. Further localization identification experiments have been conducted using a transformation of the function relating head-related transfer function to azimuth such that azimuthal space is magnified in front and minified off to the side (see figure 1). As expected, results show increased resolution in the center and decreased resolution to the side. In addition, in contrast to our initial impression that no sensorimotor adaptation took place, i.e., that response bias failed to decay with an increase in exposure time, our results now clearly show consistent and substantial adaptation (in terms of both the direct effect and the negative after effect). However, results also indicate that the decrease in response bias over time is accompanied by a decrease in resolution over time. We are now studying these data as well as performing further experiments to determine the underlying causes of this phenomenon. Illustrative data showing the changes in both resolution and bias are presented in figure 2.

2. Further work on facilities development has included four projects. First, we have acquired, integrated, and tested a new analog hardware processor to cross-check the results already obtained with the convolvotron. Second, we have constructed a pseudophone (a head-worn microphone/earphone/amplification system with controllable geometry and acoustic characteristics) to provide us with a system in which localization cues are transformed acoustically rather than by signal processing and in which no head-tracking is required. Third, we have begun to develop a new head tracker based on inertial sensors for use with acoustical simulation systems. Fourth, and finally, we have developed an additional experimental set-up which allows us to simulate a sound source held in the hand so that we can make direct comparisons with classical studies on adaption in the visual sense using optical prisms and a manual pointing response.

3. Further work on the dissemination of our results has included talks at the Society for Information Displays, the First International Conference on Auditory Displays, and the Acoustical Society of America, as well as the publication of an article in *Presence* (Durlach et al., 1992).

1.9 Research on Reduced-Capability Human Hands

Sponsor

U.S. Navy - Office of Naval Research
Grant N00014-90-J-1935

Project Staff

Lorraine A. Delhorne, Nathaniel I. Durlach, Dr. Mandayam A. Srinivasan

General background on this project was presented in *RLE Progress Report Number 134* (pages 313-314). During the past year, attention has been focused on further data collection and data analysis for constrained-hand performance in a subset of the tasks previously used by the Navy in the TOPS teleoperator study (Smith and Shimamoto, 1991).

Although direct comparisons between the results of our tests and those contained in the TOPS project are difficult to interpret because of the many differences in the two testing situations (e.g., direct vision versus vision through a helmet mounted display, working in air versus working under water, etc.), our results suggest a number of important conclusions.

First, the results indicate that all tasks considered were essentially two-finger tasks. We say this because (1) the results obtained with two fingers were nearly as good as those obtained with more fingers (both for the real hands and for the teleoperator hands) and (2) we know from casual observa-

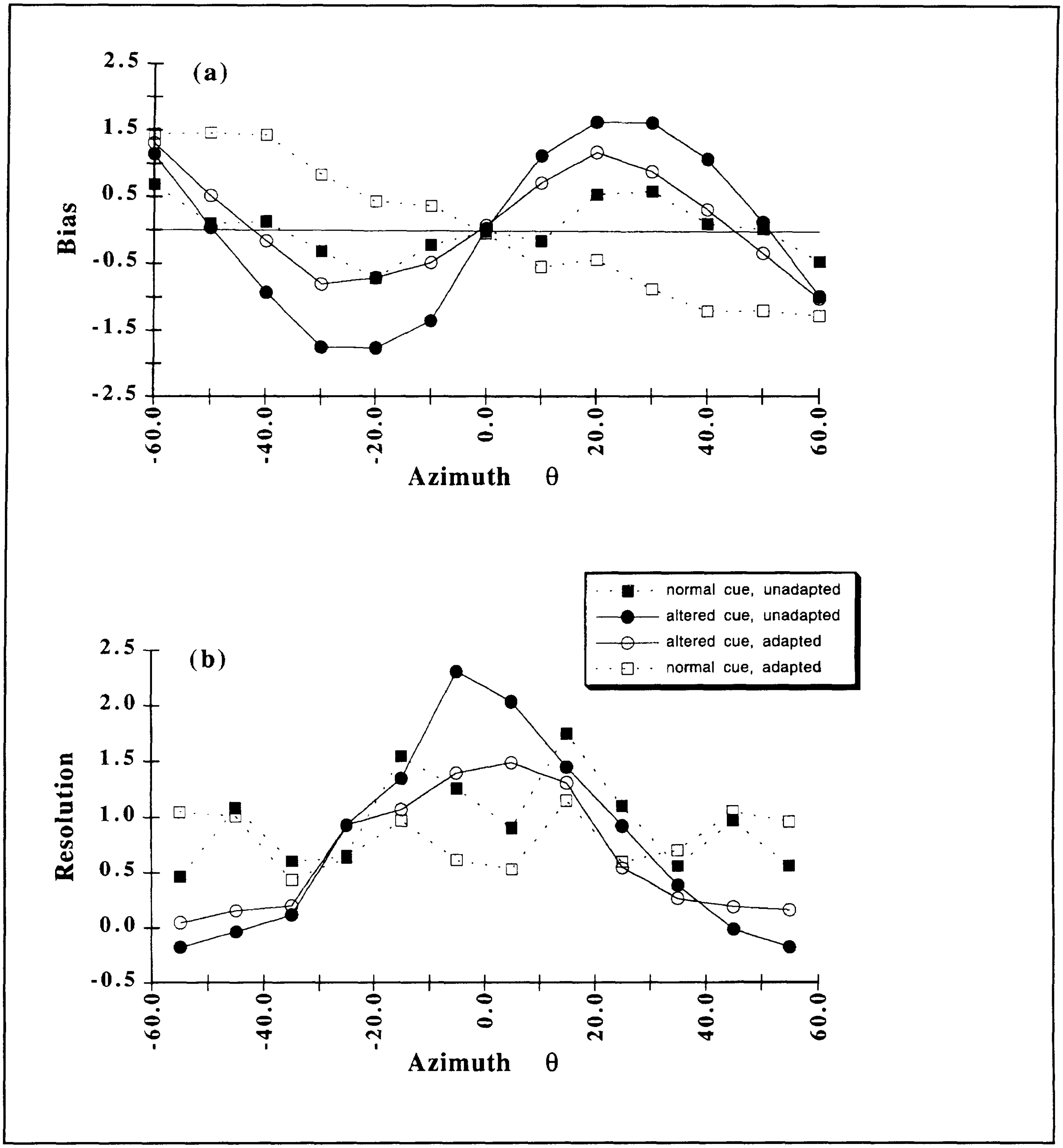

Figure 2.

tion that there exist many tasks for which two fingers are grossly inadequate.

Second, the results indicate, quite surprisingly, that performing the task with a normal human hand using a pair of pliers is (1) much superior to performing the task with either of the teleoperator systems and (2) nearly as good as performing the task with the system consisting of the thumb and index finger of the normal human hand. Result (1) suggests that the two teleoperator systems had a great deal of mechanical "slop," i.e., that information on position and force (in one or both directions) exhibited only a low level of fidelity. Item (2) suggests that for the given tasks both the compliant nature of the human finger pads and the tactile sensing in the human finger pads were relatively unimportant. (This result has been confirmed by recent experiments with anesthetized fingers.) In our opinion, this feature of the results would have been radically different if the operator were deprived of visual information. Overall, we believe our results are consistent with the idea that visual information is more redundant with tactile information than with force information. Stated differently, for non-compliant objects, vision can tell you a lot about shape and texture, but not about applied force.

As discussed in our previous report (*RLE Progress Report Number 134*), we believe that the problem of defining an adequate set of basis tasks is the most crucial research problem now facing the study of hand function and its applications to the design and evaluation of artificial hands.

Current work on this project is being devoted to preparation of a final report for this grant. It is anticipated that further research in this area will be carried out under our other grants concerned with manual sensing and manipulation.

1.10 Mechanistic Modeling of Primate Fingerpad

Sponsor

National Institutes of Health
Grant 5 R29 DC00625

Project Staff

Kiran B. Dandekar, Dr. Mandayam A. Srinivasan, Rogeve J. Gulati

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

1.10.1 Determination of Geometric and Material Properties 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 the primate fingertips was determined from accurate epoxy replicas of human and monkey fingertips. Using a videomicroscopy setup, we obtained images of orthographic projections of the epoxy replicas at various known orientations. The images were then digitized and processed to determine the boundary of the finger at each orientation. By combining the boundary data for all the different orientations, we were able to reconstruct the 3D external geometry of the fingertip.[46] We have reconstructed several human and monkey fingertips by this method.

For mechanical modeling of the human fingerpad, the Poisson's ratio, which is a measure of its compressibility, is required as an input to mathematical models. The Poisson's ratio for the human fingerpad *in vivo* is unknown at present. Available *in vitro* data for skin cannot be used as the mechanical properties are known to be different for *in vivo* and *in vitro* conditions. Also, the fingerpad as a heterogeneous structure behaves differently from the skin by itself. We have established an experimental setup where non-invasive measurements of the *in vivo* compressibility of the fingertip can be done on human subjects.[47] We have indented the fingerpads of four subjects with three different indentors and measured the change in volume of the fingerpad. Our results show that the compressibility of the fingertip increases with increases in both depth of indentation and contact area with the indentor. The highest change in

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

[47] M.A. Srinivasan, R.J. Gulati, and K. Dandekar, "*In Vivo* Compressibility of the Human Fingertip," *Proceedings of the Bioengineering Division, American Society of Mechanical Engineers Winter Annual Meeting,* Anaheim, California, November 8-13, 1992.

fingertip volume was about 5 percent, indicating that the Poisson's ratio of the fingerpad is close to, but not equal to 0.5.[48]

1.10.2 Fingertip Models and Finite Element Analyses

In past experiments, we have indented the fingerpads of humans and monkeys *in vivo* using a line load delivered by a sharp wedge and then photographed the resulting skin surface deflections. We have shown that the homogeneous elastic model of the fingertip only roughly approximates the experimental data, while a simple alternative model, which views the fingertip as an elastic membrane filled with an incompressible fluid (like a "waterbed") predicted the observed profiles very well.[49]

More recently, we have performed linear and nonlinear finite element analyses of a series of mechanistic models of the fingerpad under a variety of mechanical stimuli.[50] The models range from a semi-infinite medium to a three-dimensional model based on the actual finger geometry, composed of either a homogeneous elastic material or a thick elastic shell containing a fluid. Simulations of the mechanistic aspects of neurophysiological experiments involving mapping of receptive fields with single point loads, determination of spatial resolution of two-point stimuli, and indentations by single bars as well as periodic and aperiodic gratings have been carried out for 2D homogeneous models. We have also solved the nonlinear contact problem of indentations by cylindrical objects. The large number of numerical calculations needed even for the linear two-dimensional models necessitated the use of the MIT supercomputer. The results show, for example, that the strain energy density at the receptor site is probably the relevant stimulus that causes the responses recorded from slowly adapting afferent fibers.

After analyzing the 2D models, we found that a 3D model based on the actual geometry of the fingerpad was necessary to generate testable hypotheses about the transduction of mechanical stimuli in primate tactile sense. Using the fingertip geometry data obtained by employing the videomicroscopy system, we constructed a realistic computer model of the primate fingerpad. We then generated a finite element mesh in the model using the solid modeler software PATRAN. After having completed some preliminary checks, we are currently simulating neurophysiological experiments using the 3D model.

1.10.3 Tactile Sensing of Microtexture, Shape, and Softness

We have been collaborating with Professor R.H. LaMotte of Yale University School of Medicine in conducting psychophysical and neurophysiological studies on how humans sense by touch alone the microtextures, shapes and softness of objects.[51] By using methods of photolithography, we etched submicron-high bar gratings on glass plates. We have shown that humans can detect extremly fine textures composed of parallel bars only 50 nanometers high.[52] Our earlier neurophysiological recordings with bigger texture elements indicate that when such fine textures are stroked, skin vibrates and causes Pacinian Corpuscles to respond, thus enabling detection of the microtexure.[53]

Based on a theoretical analysis of the mechanics of contact, we have demonstrated that the receptors

---

48 M.A. Srinivasan, R.J. Gulati, and K. Dandekar, "Effect of Indentor Shape on the *in vivo* Compressibility of the Human Fingertip," submitted to *J. Biomech.* (1992).

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

50 M.A. Srinivasan and K. Dandekar. "Role of Mechanics in Cutaneous Mechanoreceptor Response," *Soc. Neurosc. Abstr.* (1991); M.A. Srinivasan and K. Dandekar, "Role of Fingertip Geometry in the Transmission of Tactile Mechanical Signals," *Proceedings of the Bioengineering Division, American Society of Mechanical Engineers Winter Annual Meeting,* Anaheim, California, November 8-13, 1992.

51 M.A. Srinivasan, "Tactual Interfaces: The Human Perceiver," in *Human-Machine Interfaces for Teleoperators and Virtual Environments,* eds. N.I. Durlach, T.B. Sheridan, and S.R. Ellis, NASA Conference Publication 10071, 1991; M.A. Srinivasan and R.H. LaMotte, "Tactile Discrimination and Representation of Texture, Shape, and Softness," *Human-Machine Interfaces for Teleoperators and Virtual Environments,* eds. N.I. Durlach, T.B. Sheridan, and S.R. Ellis, NASA Conference Publication 10071, 1991.

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 (London: 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).

respond to the low-pass filtered versions of surface pressures.[54] Thus, 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 by exploiting its relation to surface pressure rather than differentiating (which is a noise enhancing process). For the next set of experiments, we have designed and fabricated a high performance tactile stimulator controlled by a PC through transputers. Our ongoing experiments on shape sensing involve use of precisely manufactured cylindrical, ellipsoidal, and spherical surfaces as stimuli.[55]

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

1.10.4 Development of a Computational Theory of Touch

Although the "hardware" of the tactile apparatus in humans and robots is different, they have the common feature of mechanosensors embedded in a deformable medium. Thus, the computational problem of coding (predicting sensor response for a given mechanical stimulus at the surface) and decoding (inferring the mechanical stimulus at the surface by suitably processing the sensor response) need similar mechanistic analyses for their solution. We have developed a computational theory for an idealized medium subjected to arbitrary pressure or displacement loading conditions and have provided explicit formulae for the coding and decoding problems.[56]

In collaboration with Dr. Annaswamy of the Department of Mechanical Engineering at MIT, we have investigated some of the identification and control problems that occur in the context of manipulation when compliance is present in the end-effectors as well as in the object.[57] 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. Assuming that the finger and the compliant object are constrained to deform along a single spatial dimension, we have carried out parameter identification using either force or displacement inputs to the rigid backing of the end-effector. Based on this analysis, control strategies are developed to achieve a desired manipulation of the object in the workspace. Animated graphical renderings have been developed to visually illustrate the presence or absence of slipping and crushing during an active manipulation task. The theoretical results can be used to generate testable hypotheses for human or robot experiments on tactual sense.

1.11 Biomechanics of Skin-Object Contact

Sponsor

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

Project Staff

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

Although physical contact is ubiquitous in our interactions with objects in the environment, we do not yet understand the phenomena occurring at the

---

[54] 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, (London: Macmillan Press, 1991).

[55] R.H. LaMotte, M.A. Srinivasan, and A. Klusch-Petersen, "Tactile Discrimination and Identification of the Shapes and Orientations of Ellipsoidal Objects," *Soc. Neurosci. Abstr.* (1992).

[56] M.A. Srinivasan, *Tactile Sensing in Humans and Robots: Computational Theory and Algorithms,* Newman Laboratory Technical Report, Dept. of Mech. Eng., MIT, 1988; M.A. Srinivasan, "Computations in Tactile Sensing," *Proceedings of Computation and Neural Systems (CNS '92) Meeting,* San Francisco, California, 1992.

[57] A.M. Annaswamy and M.A. Srinivasan, "Adaptive Control for Grasping and Manipulation of Compliant Objects with Compliant Fingerpads," *Proceedings of the American Control Conference,* Boston, Massachusetts, 1991; A.M. Annaswamy, M.A. Srinivasan, and S. Weibel, "Identification for Grasping and Manipulation of Compliant Objects with Compliant Fingerpads," submitted to *IEEE Trans. Automatic Control.*

skin-object interface. As mentioned before, the spatio-temporal distribution of mechanical loads on skin at the contact interface is the source of all tactile information. These loads, specified as pressure, displacements, etc., depend on the geometrical and material properties of both the contacting entities as well as the overall forces of interaction. The goal of this project is to determine the growth and motion of contact regions and 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. These results are being used to gain a deeper understanding of the data we have already obtained for the same test objects, such as electrophysiologically recorded responses of cutaneous mechanoreceptive afferent fibers, and psychophysically determined human discriminabilities.

Our initial goal was to record the forces and videoimages of contact under active touch when subjects press or stroke various transparent specimens mounted on a force transducer. Accordingly, we designed a videomicroscopy system and a force sensing system, both interfaced to a computer. The videomicroscopy system consists of a set of video zoom lenses attached to a high-resolution CCD camera whose output can either be digitized directly at about 5 frames/s or stored on a laserdisk at real-time frame rates (30 frames/s) for off-line digitization. The zoom lenses enable continuous variation of magnification with the field of view covering the entire fingerpad or only a few fingerprint ridges. High contrast images are achieved with coaxial 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 comparable range of forces (5 Newtons). The transparent test objects can be attached to these sensors for both biomechanical and psychophysical experiments.

Using the test facility described above, we have performed a set of experiments with human subjects to investigate the relationship between the contact force, contact area, and compliance of the object. The experiments involved active indentation of transparent compliant rubber specimens and a glass plate with the subjects' fingerpads. Static video images of the contact regions were captured at various force levels and magnifications. In order to minimize the effects of non-uniform illumination, we implemented homomorphic image processing algorithms. This essentially consisted of high-pass filtering to remove the low frequency (spatial) illumination intensity variations so that only the higher frequency variations due to the fingerprint patterns remained in the image. For high magnification images, since the spatial frequency of non-uniform lighting was of the same order as the finger ridge frequency, image decimation techniques had to be employed in addition to homomorphic processing. The processed images showed that contact regions consisted of discontinuous "islands" along each finger ridge with clear distinction between contact and non-contact regions over the entire field of view.

As expected, the results show that the nominal overall contact area increases as the contact force increases for all the specimens. More interestingly, the percentage of actual contact area relative to the nominal (overall) contact area also increases and is a function of the object compliance. For the plane glass plate, the percentage contact area is only around 18 percent for 0.1 N force and about 22 percent for 1.5 N force. In contrast, the percentage contact area for a compliant rubber specimen is much higher than that for the plane glass plate, about 43 percent for 0.1 N force and more than 80 percent for 1.5 N force. The implication to tactual discrimination of softness is that for objects whose compliances can be discriminated 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 different 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 R.H. LaMotte of the Yale University School of Medicine.

In addition to helping us develop an understanding of the processes involved in the peripheral neural encoding of object compliance information, the results of these experiments provide testable hypotheses for investigating cortical representation of this information and the processing algorithms used by the central nervous system for decoding the peripheral neural signals. These results are also useful in the design of tactile displays that convey object compliance information to human operators in teleoperation and virtual environment systems.[58] The results of the experiments also help in the design of tactile sensors for robots and development of algorithms to decode the sensor information on object compliance.

1.12 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, Louise Jandura, Kiran B. Dandekar

This is a new project funded by the University Research Initiative Program of the Office of Naval Research (ONR) which involves several investigators from MIT, Harvard University, and Yale University. The premise is that humans and robots can be viewed as two implementations of systems that perform tactual sensorimotor tasks under a common set of constraints. By devising a research plan for the study of human and robot hands that focuses on the commonalities rather than the differences, we propose to achieve mutually beneficial advances in both fields. Specifically, we propose to carry out a program of research with the following related thrusts. We will:

1. Construct a high precision "glass smooth" robot which will serve as the core of both human and robotic experiments. This robot will have an integrated design of tactile sensing and mechanical subsystems. Initially, the sensors will include force and motion sensors; subsequently, we will incorporate a newly developed tactile array using micromechanics techniques.

2. Carry out a suite of experiments aimed at (1) determining the abilities and limitations of the sensorimotor functions of human and robot hands and (2) understanding the mechanisms and information representations they employ. We propose to design and fabricate several experimental stations based on robotic technology for conducting a wide range of biomechanical, neurophysiological, and psychophysical experiments on human and animal tactual systems as well as measurements on robotic prototypes.

3. Develop methods applicable to both humans and robots for integrating tactual data from multiple types of sensors, multiple spatially separated sensors, and sequences of multiple contacts to achieve a consistent perception of the external environment. We will explore methods for using this integrated information for carrying out exploratory and manipulation tasks as well as inferring object structure and function.

4. Articulate design criteria for robot systems to be employed in tactual tasks. We will test these criteria in a number of robotic testbeds.

During the past few months, a two degree of freedom robot capable of motion in a plane has been fabricated and assembled and is currently being tested for use as a tactile stimulator. The original version of the robot was designed by Professor R. Howe of Harvard University and minor modifications were made to enhance its use as a tactile stimulator. The stimulator can be operated under both position and force control. It will be used to conduct a variety of biomechanical and psychophysical experiments to investigate human tactile sense.

In collaboration with Dr. J.K. Salisbury of MIT and his students, the first generation design of an Instrumented Screwdriver (ISD) has been completed, and its fabrication is underway. This device will simulate the task of using a screwdriver in order to test the force control capability of humans. The "target screw" is simulated by a computer-controlled magnetic particle brake. For example, during the performance of force tracking tasks by a subject, the force vector under each of two or three fingerpads can be measured by custom-designed six-axis force sensors. The subject will perform the tasks using active touch with his or her fingerpads under both normal and locally anesthetized conditions to isolate the sources of tactual information. We have completed a set of similar force tracking experiments using a stationary six-axis force sensor and the data is being analyzed. Preliminary results show that local anesthesia, which blocks tactile information from cutaneous mechanoreceptors,

[58] M.A. Srinivasan, sections on haptic perception and haptic interfaces, in *Research Directions in Virtual Environments: Report of an NSF Invitational Workshop, Computer Graphics,* ed. G. Bishop et al. 26(3): 153-177 (1992); M.A. Srinivasan and J.K. Salisbury, sections on haptics, in *Virtual Environment Technology for Training,* BBN Report No. 7661, Prepared by the Virtual Environment and Teleoperator Research Consortium (VETREC) affiliated with MIT, 1992.

does not seriously impair the performance of the subjects. This implies that among the tactile and kinesthetic sources of force information, humans may be relying on the kinesthetic channel for force control, even though the tactile channel has better sensory resolution.

1.13 Publications

Durlach, N.I., A. Rigopulos, X.D. Pang, W.S. Woods, A. Kulkarni, H.S. Colburn, and E.M. Wenzel. "On the Externalization of Auditory Images." *Presence* 1: 251-257 (1992).

Durlach, N.I. "Virtual Environments: Introductory Remarks." *J. Acoust. Soc. Am.* 4: 2331-2332 (1992).

Durlach, N.I., R.M. Held, and B.G. Shinn-Cunnigham. "Super Auditory Localization Displays." *Society for Information Display International Symposium, Digest of Technical Papers* 23: 98-101 (1992).

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

LaMotte, R.H., Srinivasan M.A., and A. Klusch-Petersen. "Tactile Discrimination and Identification of the Shapes and Orientations of Ellipsoidal Objects." *Soc. Neurosci. Abstr.* (1992).

Rabinowitz, W.M., 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).

Rankovic, C.M., R.L. Freyman, and P.M. Zurek. "Potential Benefits of Adaptive Frequency-Gain Characteristics for Speech Reception in Noise." *J. Acoust. Soc. Am.* 91: 354-362 (1992).

Reed, C.M., N.I. Durlach, and L.A. Delhorne. "Historical Overview of Tactile Aid Research." In *Proceedings of the Second International Conference on Tactile Aids, Hearing Aids, and Cochlear Implants.* Eds. A. Risberg and K.-E. Spens. Stockholm, Sweden: Royal Institute of Technology, 1992. Forthcoming.

Reed, C.M., L.A. Delhorne, and N.I. Durlach. "Results Obtained with Tactaid II and Tactaid VII." *Proceedings of the Second International Conference on Tactile Aids, Hearing Aids, and Cochlear Implants.* Eds. A. Risberg and K.-E. Spens. Stockholm, Sweden: Royal Institute of Technology, 1992. Forthcoming.

Reed, C.M., and L.A. Delhorne. "Field Study of Deaf Adult Users of Tactaid II and Tactaid VII." Presentation made at the Annual Convention of the Association for Late Deafened Adults, Boston, Massachusetts, September 10-13, 1992.

Reed, C.M., N.I. Durlach, and L.A. Delhorne. "Natural Methods of Tactual Communication." In *Tactile Aids for the Hearing Impaired.* Ed. Ian R. Summers. Whurr Publishers Limited, 1992, pages 218-230.

Reed, C.M., N.I. Durlach, and L.A. Delhorne. "The Tactual Reception of Speech, Fingerspelling, and Sign Language by the Deaf-Blind." *SID 92 Digest* pp. 102-105 (1992).

Reed, C.M., W.M. Rabinowitz, N.I. Durlach, L.A. Delhorne, L.D. Braida, J.C. Pemberton, B.D. Mulcahey, and D.L. Washington. "Analytic Study of the Tadoma Method: Improving Performance through the Use of Supplementary Tactual Displays." *J. Speech Hear. Res.* 35: 450-465 (1992).

Shinn-Cunningham, B.G., N.I. Durlach, and R.M. Held. "Adaptation to Transformed Auditory Localization Cues in a Hybrid Real/Virtual Environment." *J. Acoust. Soc. Am.* 4: 2334 (1992).

Shinn-Cunningham, B.G., P.M. Zurek, and N.I. Durlach. "Adjustment and Discrimnation Measurements of the Precedence Effect." *J. Acoust. Soc. Am.* Forthcoming.

Srinivasan, M.A. "Computations in Tactile Sensing." *Proceedings of Computation and Neural Systems (CNS '92) Meeting,* San Francisco, California, 1992.

Srinivasan, M.A. Sections on haptic perception and haptic interfaces. In *Research Directions in Virtual Environments: Report of an NSF Invitational Workshop.* Ed. G. Bishop et al. *Comput. Graphics* 26 (3): 153-177 (1992).

Srinivasan, M.A., and K. Dandekar. "Role of Fingertip Geometry in the Transmission of Tactile Mechanical Signals." *Proceedings of the Bioengineering Division, American Society of Mechanical Engineers Winter Annual Meeting,* Anaheim, California, November 8-13, 1992.

Srinivasan, M.A., R.J. Gulati, and K. Dandekar. "In vivo Compressibility of the Human Fingertip."

Proceedings of the Bioengineering Division, American Society of Mechanical Engineers Winter Annual Meeting, Anaheim, California, November 8-13, 1992.

Srinivasan, M.A., and J.K. Salisbury. Sections on haptics, In *Virtual Environment Technology for Training,* BBN Report No. 7661, The Virtual Environment and Teleoperator Research Consortium (VETREC) affiliated with MIT, 1992.

Tan, H.Z., X.D. Pang, and N.I. Durlach. "Manual Resolution of Length, Force, and Compliance." *Proceedings of the American Society of Mechanical Engineers, Winter Annual Meeting,* Anaheim, California, November 8-13, 1992.

Uchanski, R.M., K.M. Millier, C.M. Reed, and L.D. Braida. "Effects of Token Variability on Vowel Identification." In *The Processing of Speech: From the Auditory Periphery to Word Recognition.* Berlin: Mouton de Gruyter. Forthcoming.

Zurek, P.M. "Binaural Advantages and Directional Effects in Speech Intelligibility." In *Acoustical Factors Affecting Hearing Aid Performance II,* Eds. G.A. Studebaker and I. Hochberg. Boston: Allyn and Bacon, 1992.

Zurek, P.M. "Detectability of Transient and Sinusoidal Otoacoustic Emissions." *Ear Hear.* 13: 307-310 (1992).

Theses

Gulati, R. *In Vitro Compressibility of the Human Fingerpad.* B.S. thesis., Dept. of Mech. Eng., MIT, 1992.

Martin, G.R. *Studies of Real-Time Multiband Adaptive Gain Hearing Aids.* S.M. thesis, Dept. Electr. Eng. and Comput. Sci., MIT, 1992.

Stadler, R.W. *Optimally Directive Microphones for Hearing Aids.* S.M. thesis, Dept. Electr. Eng. and Comput. Sci., MIT, 1992.