Implementation of an On-Line Speaker Verification Scheme

R. C. Lummis

Citation: The Journal of the Acoustical Society of America **52**, 181 (1972); doi: 10.1121/1.1982135
View online: https://doi.org/10.1121/1.1982135
View Table of Contents: https://asa.scitation.org/toc/jas/52/1A
Published by the Acoustical Society of America

demonstrator, noise cancellation device, noise reduction of a computer card reader, and the effect of noise on plants. Since most noise control projects involve the student in value judgments, an auxiliary educational benefit results.

4:00

OO9. Design and Operation of a Continuous, Areawide Aircraft-Noise Monitoring System. GORDON BRICKEN, *Northrop Corporation, Anaheim, California,* AND JOHN K. HILLIARD, *Ramberg & Lowrey, Santa Ana, California.*—A five-station, permanent, 24-h centralized aircraft noise monitoring system has been installed in the community surrounding Orange County Airport, Santa Ana, California. This paper will discuss the design concept, the results obtained from initial tests involving measurement of over 1000 flights, the practical experience gained from the installation, the true value to the community, and the potential for effective noise management programs. In addition, the measurement data will be used to demonstrate the value of the California noise index CNEL.

FRIDAY, 21 APRIL 1972 GEORGIAN ROOM, 2:00 P.M.

Session PP. Speech Communication VII: Analysis, Identification, and Recognition of Speech and Speakers

JOHN F. BOEHM, *Chairman*

National Security Agency, Ft. George G. Meade, Maryland 20755

Contributed Papers

PP1. Implementation of an On-Line Speaker Verification Scheme. R. C. LUMMIS, *Bell Telephone Laboratories, Murray Hill, New Jersey 07974.*—The implementation is described of an on-line version of the scheme for speaker verification reported previously. A Honeywell DDP-516 computer is used, with microphone and keyboard input, disk storage, and graphic output. Utterances are converted to pitch and gain contours and compared with similar functions fetched from disk that represent past vocal behavior of the speaker whose identity is claimed. The comparison process includes automatic temporal registration. Special features of the implementation include the ability to update stored reference patterns at will so as to incorporate the features of new utterances, and a graphic display that effectively illustrates which portions of a test utterance are within the programmed limits of acceptability.

PP2. Text-Independent Speaker Recognition. B. S. ATAL, *Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey 07974.*—In automatic speaker recognition methods, the speaker to be recognized is usually required to speak the same utterance which was used to obtain the reference pattern for that speaker. However, such a restriction is not generally necessary for speaker recognition by humans. Is reliable automatic speaker recognition similarly possible? A speaker identification experiment was performed in a population of 10 female speakers. The acoustical information was represented by sets of 12 predictor coefficients obtained by minimizing the mean-squared prediction error over speech segments 50 msec in duration. Each set of these predictor coefficients was represented by a 12-dimensional vector. New sets of coordinates which minimized the intraspeaker variance were determined by linear transformation of the original vector space. The segment of the utterance used to identify an individual was different from the segments used to form the reference pattern for that individual. For each segment, the unknown vector was correlated with the reference vectors and the correlations were averaged over a number of segments—the speaker with the largest correlation was identified as the unknown speaker. The over-all identification accuracy was 93% for 40 speech segments. These results suggest that successful automatic speaker identification is possible independently of the spoken text.

PP3. On Evaluating Formant Synthesizer Control Signals. II. W. WATHEN-DUNN, S. B. MICHAELS, L. V. KRIGER, AND H. I. SORON, *Air Force Cambridge Research Laboratories, Bedford, Massachusetts 01730.*—An attempt to evaluate the effectiveness of three sets of formant synthesizer control signals, derived by different means, has utilized VCV utterances in which the C's were one of the six stop consonants and the V's were one of the six vowels, /i, e, a, o, u, ʌ/. The stops are the most quickly articulated phonemes and the first five of the vowels are tense and are thus articulated at positions most removed from the central articulatory configuration, which is represented by the last vowel, /ʌ/. Use of these utterances is predicated on the idea that the stops and the tense vowels make the greatest demands on synthesizer performance and are hence most suited for evaluation. Two tests have been run: (1) an identification test for each of the consonants and (2) comparisons of the same consonant produced by different means. Results for the labial stops, /b, p/, were presented at the 82nd Meeting of the Society in Denver, and this paper will present the results for the remaining stops, /d, t, g, k/.

PP4. Orientation to Speech Sound. T. F. MYERS, *Department of Psychology, University of Edinburgh, Great Britain,* R. ZHUKOVA AND L. A. CHISTOVICH, *Pavlov Pyhsiological Institute, Leningrad, U.S.S.R.*—A series of experiments are described in support of a model of speech recognition. The vowel /a/ from a complex signal generator was presented binaurally, each ear channel being independently controlled for signal intensity and duration of a medial silent interval (0–90 msec), such that a fused percent /a/, /ara/, /ala/, or /a–a/ resulted in the center, or to one or other side of auditory space. Analysis of categorical response probability and lateralization of a percept as functions of duration of the interval and interaural intensity difference enabled the functional levels and operations represented by the model to be substantiated.

PP5. Spoken Digit Recognition Using Partial Autocorrelation Coefficients (PAC). Y. NAKANO, A. ICHIKAWA, AND K. NAKATA, *Central Research Laboratory of Hitachi Ltd., Kokubunji, Tokyo, Japan.*—Spoken digit recognition using autocorrelation coefficients (PAC or k parameters) is investigated with a digital computer. Speech samples are 10 spoken