XV. DIGITAL SIGNAL PROCESSING

Academic and Research Staff

Prof. Alan V. Oppenheim
Prof. James H. McClellan
Dr. Wolfgang F. G. Mecklenbräuker*

Graduate Students

David S. K. Chan
David B. Harris
Anthony P. Holt
Gary E. Kopec
Michael R. Portnoff
Thomas F. Quatieri, Jr.
Elliot Singer
José M. Tribolet

1. SPEED TRANSFORMATIONS OF SPEECH

U. S. Navy Office of Naval Research (Contract N00014-75-C-0951)

Alan V. Oppenheim, Michael R. Portnoff

This research is directed toward transforming a speech signal into a new speechlike signal that is perceived as identical to the original signal in all respects except that it is articulated at a rate different from that of the original speech. Such a technique is of interest in a variety of applications including learning a second language, developing talking books for the blind, generation of stimuli for psychological experiments, enhancement of intelligibility, and interactively searching through large quantities of speech data.

Our current effort is directed toward three problems. The first is the formulation of an appropriate quasi-stationary representation for sampled speech signals. This representation must have the property that time-scaling the parameters of the representation corresponds to implementing the desired speed transformation. The second problem is to develop a high-quality speech analysis-synthesis system based on this representation. As part of this effort, we have developed an efficient analysis-synthesis system based on the discrete short-time Fourier transform. The third problem is to develop a set of rules governing the local scale factor for the speed transformation. Specifically, it is well known that a change in the rate at which speech is articulated does not correspond to a simple time scaling of the speech parameters independent of the uttered phoneme. We expect the scaling rules to take account of the fact, for example, that the formant trajectories of a vowel should be scaled, whereas the duration of a plosive burst should be left unchanged.

2. ENHANCEMENT OF LOWPASS FILTERED SPEECH

U.S. Navy Office of Naval Research (Contract N00014-75-C-0951)

Alan V. Oppenheim, Elliot Singer

This research is part of a more general investigation concerned with enhancing the quality of degraded speech. When the degradation is caused by lowpass filtering, it is conceivable that the high-frequency components originally present could be determined from an examination of the available low-frequency spectrum and the original speech thereby reconstructed. This is especially true for voiced speech where the overall

*Visiting Scientist, Philips Research Laboratories, Eindhoven, The Netherlands.

shapes of the spectral envelopes are well established. In a broad sense, therefore, an enhancement system for lowpass filtered speech would incorporate algorithms for generating the missing high-frequency structure and processing schemes for synthesizing the corresponding signals.

In the initial work on this project we avoided the need for adaptive filtering by keeping the characteristics of the high-frequency spectral shaping filter fixed in time. Also, the excitation function used to drive this filter was derived from the lowpass filtered speech itself rather than through a pitch detection system. This technique, known as spectral flattening, seeks to smooth or flatten the spectral envelope while preserving the identity of the individual harmonic peaks.

Some preliminary experiments with a speech sample lowpass-filtered to 2 kHz produced a result that sounded somewhat unnatural. The system is now being improved by the addition of more sophisticated spectral flattening techniques, as well as more careful structuring of the high-frequency synthesizer. Plans for the future call for the implementation of an adaptive high-frequency filter and the possible extension of the high-frequency reinsertion technique to speech containing both voiced and unvoiced segments.

3. DESIGN AND IMPLEMENTATION OF VARIABLE CUTOFF DIGITAL FILTERS

U. S. Navy Office of Naval Research (Contract N00014-75-C-0951)

National Science Foundation (Grant ENG71-02319-A02)

Anthony P. Holt, Wolfgang F. G. Mecklenbräuker, Alan V. Oppenheim

We are constructing a digital filter with an easily variable cutoff frequency, using the technique of digital frequency warping. In brief, each unit delay of a digital filter is replaced by an all-pass filter. This permits the cutoff frequency of the digital filter to be varied over a wide range by modifying a single parameter. The particular filter that is being constructed is lowpass in design with a cutoff frequency adjustable from 1 kHz to 10 kHz. The cutoff frequency may be controlled from the front panel or by a computer through an appropriate interface. To the user, the filter will behave like an analog filter because of internal analog-to-digital and digital-to-analog conversion. The variable cutoff filter will be incorporated as a presampling and postsampling filter for digital speech processing. With an easily controlled sharp cutoff filter of this kind, the sampling rate applied to the speech input and output in a speech-processing system can easily be varied. The filter now being implemented in general has a nonlinear phase even when the prototype filter is linear phase. We have also developed a method for implementing in a similar manner a variable cutoff digital filter that always maintains exactly linear phase. While, at present, there are no plans for the actual construction of the linear-phase filter, its application to speech-processing systems is being investigated.

4. SPEECH ANALYSIS BY HOMOMORPHIC PREDICTION

U. S. Navy Office of Naval Research (Contract N00014-75-C-0951)

Gary E. Kopec, Alan V. Oppenheim

Linear prediction is a generally accepted method for obtaining all-pole speech representations. In many situations (e. g. , nasalization studies) spectral zeros are important and a more general modeling procedure is required. Unfortunately, the need for pitch synchronization has limited the success of available techniques. In this project we have developed and explored a new approach to pole-zero analysis, called homomorphic prediction, which seems to avoid the synchronization problem. A minimum-phase estimate of the vocal-tract impulse response is obtained by homomorphic filtering of the speech waveform. Such a signal, by definition, has a known time registration. Linear prediction is applied to this waveform to identify its poles. The LPC "residual" (error signal) is computed by inverse filtering. This signal contains the information about the zeros. Its z-transform is then approximated by a polynomial either through a weighted least-squares procedure (homomorphic prediction, using Shanks' method of finding zeros) or by spectral inversion followed by a second pass of LPC (homomorphic prediction involving "inverse LPC"). A preliminary evaluation on real and synthetic speech has been carried out.

In further work we intend to refine the general technique of homomorphic prediction and evaluate the pole-zero modeling within the context of a speech analysis-synthesis system. Toward this end we intend to implement an analysis-synthesis system based on homomorphic prediction and evaluate the potential improvement afforded by the inclusion of the zeros.

5. APPLICATIONS OF HOMOMORPHIC FILTERING TO SEISMIC DATA PROCESSING

U.S. Navy Office of Naval Research (Contract N00014-75-C-0951)

National Science Foundation (Grant ENG71-02319-A02)

Alan V. Oppenheim, José M. Tribolet

Homomorphic filtering is a nonlinear technique that may be applied to deconvolution problems. Two homomorphic filtering approaches to deconvolution are available. The first, homomorphic deconvolution, has been known for a while and has been used with success in speech processing; here deconvolution is performed by linear filtering on the cepstral domain. Another approach, called homomorphic prediction, has been developed recently by us and performs deconvolution by linear filtering in the time domain, with a filter designed by combining homomorphic filtering with predictive filtering. This technique has been applied with success to the compression of air gun signatures encountered in marine seismic data. An exploratory use on lunar data has yielded promising results.

The aim of this research project is to study the application of homomorphic deconvolution and prediction to the deconvolution of seismic data. Toward this goal a systematic study of deconvolution of synthetic data will be made, to be followed by an evaluation of the performance of these techniques on real data.

6. STRUCTURES FOR THE IMPLEMENTATION OF TWO-DIMENSIONAL DIGITAL FILTERS

National Science Foundation (Grant ENG71-02319-A02)

David S. K. Chan, James H. McClellan

A new structure for efficient implementation of two-dimensional filters has been derived from the recursion for Chebyshev polynomials. This new structure represents an improvement over the structures proposed previously by Mecklenbräuker and Mersereau because the coefficients of the filter are well behaved. A comparative study of the two structures is under way in order to quantify these effects.

Additional work is proceeding in an investigation of the computational structure for these transformation filters. Such a study reveals the logical constraints on the order of arithmetic operations and required storage organization for special-purpose hardware implementation.

7. McCLELLAN TRANSFORMATIONS FOR TWO-DIMENSIONAL DIGITAL FILTERS

National Science Foundation (Grant ENG71-02319-A02)

David B. Harris, Wolfgang F. G. Mecklenbräuker, Thomas F. Quatieri, Jr.

We have shown that a generalized form of the original McClellan transformation is a viable technique for the design of two-dimensional filters with quadrilateral symmetry. The essence of the technique involves mapping a one-dimensional finite impulse response (FIR) filter into a two-dimensional FIR filter through a substitution of variables. Several straightforward algorithms for estimating the transformation parameters have been developed.

We have also shown that two-dimensional filters designed by using a transformation possess a series of highly efficient implementations that are more efficient than direct implementations and are even more efficient than a fast Fourier transform (FFT) implementation for filters of order up to approximately 50×50. The implementation exploits the structure of the one-dimensional prototype filter. We have also explored techniques for overcoming potential difficulties inherent in these realizations, such as storage and effects of finite register length.

Finally, we have considered the limitations of McClellan transformations in producing optimum two-dimensional filters when the desired pass- and stop-band edges for ideal two-dimensional filters coincide with contours of the transformations. Necessary and sufficient conditions on the transformation guaranteeing that every optimum one-dimensional prototype maps to an optimum two-dimensional approximation have been found for the case of first-order transformations.