Inverse filtering. Instrumentation and techniques

Lindqvist-Gauffin, J.

journal: STL-QPSR
volume: 5
number: 4
year: 1964
pages: 001-004

http://www.speech.kth.se/qpsr
I. SPEECH PRODUCTION

A. INVERSE FILTERING. INSTRUMENTATION AND TECHNIQUES

J. Lindqvist

The inverse-filtering equipment previously reported on (4)(8) is now completed and has been used for some preliminary experiments. In addition to the derivation of the glottal waveforms, the anti-resonance units have also been used for pole-zero matchings in the studies of the vocal tract transfer characteristics (6).

A block diagram of the complete system is shown in Fig. I-A-1. The FM-modulated speech signal is recorded on one track of the tape and a pulse-train as a time marker on the other track. From this tape a loop for a signal duration of about 1 sec is made. (This process of making a tape loop is rather time consuming, and some improvement may be desirable in this respect in order to devise a practical scheme. The most promising perhaps, is an application of the new repetitive playback system now under development in the laboratory.) The tape speed of 15 inches/sec gives a frequency response up to 2400 c/s only, although the rest of the equipment is designed for frequencies up to 4000 c/s.

From the tape-recorder the signal is fed to an FM-demodulator and then through a gate. The gate and the oscilloscope are triggered by a counter at a preset number of counted pulses. A reset signal (5000 c/s) for the counter is inserted in the loop splice at the joint of splicing.

The units marked C in Fig. I-A-1 are active low-pass filters designed to compensate for the higher poles. The F1 circuit includes an integration of the signal for compensating the radiation characteristics. This integration has a time constant of 6 c/s, or alternatively 0.6 c/s, by selecting an optional circuit which is built in in the last low-pass filter. The filters are connected in the order shown in Fig. I-A-1 which is considered optimal with respect to both noise and non-linear distortion.

The recorded material for inverse filtering must be correct in phase characteristic down to very low frequencies.
Fig. I-A-1. Block diagram of the inverse-filtering equipment.
This makes the recording sensitive to air-pressure fluctuations and to slight movements of the subject. Special care must be taken during recording with respect to this problem. A frequency analysis was made of the low frequency noise in our anechoic chamber where the recording takes place. This analysis showed a slope of -12 dB/oct in the range 2-10 c/s.

A phase-corrected high-pass filter has been used in some of our recordings to reduce the effect of this noise. The cutoff frequency of this filter is 5 c/s. The remaining low frequency noise has the appearance of a relatively slowly varying random addition to the slope of the zero line of a short sample gated out for inspection. This slope is compensated by adding a suitable DC voltage to the input signal before the integration in the F1-circuit. A Bruel & Kjær condenser microphone (Type 4131) and amplifier (Type 2603) are used for the recording. The cathode follower in the microphone has been redesigned for a low frequency cutoff at about 2 c/s. A simple RC-circuit is available for correcting the phase response of the amplifier and of the microphone when necessary (viz. for a very low pitch). The FM-modulator and demodulator were designed and built in the laboratory.

For normal non-nasalized voice samples, we do not encounter any difficulty in finding an optimum setting of the anti-resonance filters. The frequencies and bandwidths of the zeros are simply set to produce minimum formant frequency ripples in the output waveform. Fig. I-A-2 shows a typical glottal wave together with the input speech wave and waveforms representing the first, second, and third formant oscillations in isolation. A formant may be isolated by removing the particular zero and adding a broad-band band-pass filter. No integration is used in this case because we would rather suppress the fundamental frequency.

Unfortunately, the closure of the nasal cavities by the soft palate during the production of normal "non-nasal" vowels is often incomplete. This could introduce an error in the resulting waveform. From the sweep-tone study of the vocal tract response we know that it is possible to make a rather good match up to about 1500 c/s, for nasalized vowels, by using an additional pole-zero pair in the lowest frequency range. Consequently, it is possible to inverse filter even highly nasalized vowels by
Speaker J. L. 
vowel [a] 
F₀ = 134 c/s

Fig. I-A-2. Inverse filtering of vowel [a], male subject, F₀ = 134 c/s, 
together with input speech wave and isolated first, second 
and third formant oscillation.
adding an extra zero-pole pair to the inverse filter and limiting the frequency range appropriately.

Another condition which was not considered in the design is the coupling to the subglottal system. Under normal conditions this coupling is supposed to be very small. In pathological cases, however, where the subject has an incomplete closure of the glottis, the coupling to the subglottal air may cause additional pole-zero pairs which may not be apparently different from those caused by nasalization at least when only the frequency range below 1500 c/s is considered. The main difference is the time variation within one glottal cycle which is the case only for a subglottal coupling.

The effect of the pole of this additional pole-zero pair upon the resulting waveform is generally not very significant since the residue amplitude of this subglottal pole is anyway small. Fig. I-A-3 shows a case where the inverse functions were approximated without any pole. Three zeros were set at 400, 775 and 1100 c/s and two low-pass filters at 1200 c/s.

Fig. I-A-4 pertains to a sample of a hoarse voice where the subglottal coupling appears to be negligible. Note the particular excitation of the second formant on medium and high voice efforts. The low amplitude of the formants in the hoarse voice compared with the normal voice in the same figure is due to the lack of energy in higher frequencies (in the voice source).

Experiments on normal male voices with various fundamental frequencies and intensities confirm the generally accepted observation that a prolonged closed interval and a better defined closure moment characterize high intensity conditions.

Our experience with the inverse-filtering techniques is on the whole quite good. A requirement for successful performance is that great care be devoted to the recording of the speech material.

Future studies of voice source characteristics will incorporate:

1. Determination of the voice source spectrum and the waveform at various pitch values and voice intensities in sustained samples uttered by normal subjects.
Fig. I-A-1. Inverse-filtered output of a speaker with an assumedly leaking glottis.
Fig. 1-A-4. Inverse-filtered output of a hoarse voice compared with a normal voice uttered by the same speaker.
2. Time variation in the waveform and the spectrum of the voice source in connected speech.

3. An estimation of the diagnostic value of the inverse filtering. Comparisons with glottographic (5)(9) and stroboscopic techniques(2).

References:

(3) Björk, L.: "Velopharyngeal function in connected speech. Studies using tomography and cineradiography synchronized with speech spectrography" (Thesis work for the degree of M.D.), Acta Radiologica, Suppl. 202 (Stockholm 1961).