Journal of Rehabilitation Research and Development, Vol. 23, No. 1, 1986 pages 25-39

Compression Systems for Hearing Aids and Cochlear Prostheses

M. W. WHITE, Ph. D.

Department of Otolaryngology
University of California, San Francisco
San Francisco, California 94143

Abstract—Audio and speech compression systems have suffered several characteristic deficiencies. Single-channel compression systems cannot compress wideband signals without suffering from either spectral distortion or the inability to respond quickly to fast transients. When the input signal contains noise in addition to the desired speech signal, single-channel systems unnecessarily attenuate speech information. Single-channel compressors cannot compress the input signal differentially as a function of frequency. Multichannel compressors are capable of different levels of compression as a function of frequency. However, standard multichannel compressors unnecessarily attenuate important information about the shape of the short-term speech spectrum. This has resulted in poorer speech perception when using standard multichannel systems as compared with single-channel compression systems. A more general form of multichannel compression can emphasize information about the shape of the short-term speech spectrum. Susceptibility to many forms of noise is also reduced with such multichannel systems. Spectral distortion and undesired rapid overshoots and undershoots of signal level, characteristic of many single- and multi-channel systems, can be substantially reduced with such systems.

INTRODUCTION

In many types of hearing aids, the wide intensity range of auditory stimuli is compressed into the relatively narrow dynamic range of the hearing aid recipient. Also, in cochlear prostheses, the wide dynamic range of our auditory environment must be compressed or "mapped" into the very narrow operating range of electrical stimulation of the nerve. To design effective compression systems, one must understand the strong relationships between the temporal and spectral characteristics of auditory signals and compression systems.

There are three fundamental considerations in the design of compression systems for speech: (i) the statistical characteristics of the speech signal, (ii) the characteristics of the transmission channel at the compressor's output (e.g., the channel's operating range and the distribution of intensity-difference limens across this dynamic range) and (iii) the spectral-temporal fidelity necessary for successful transmission of perceptually significant information. In the design of compressors, most difficulties arise because the input signal has a relatively wide-band spectrum in which the signal level varies rapidly in time. The design problem is most difficult if the input signal's range is very much greater than the output channel's dynamic range.

BACKGROUND

Single-Channel Compressors

In a typical feed-forward compressor such as the one shown in the block diagram in Figure 1, the envelope detector acts as a level estimator and can be implemented in a number of ways. For example, a short-term root-mean square (rms) measure could be used. A rectifier followed by a lowpass filter is an example of an envelope-detector that is relatively simple to implement. In most cases, it is preferable to use a full-wave rectifier, but a half-wave rectifier can also be used. The lowpass filter acts as a "leaky integrator" for smoothing the rectifier's output. The impulse response of the lowpass filter acts as the "integration window". The lowpass filter delays the signal in the lower branch. The delay stage in the upper branch generates an equal delay to synchronize the two branches. The instantaneous nonlinearity (INL) is used to set the compression ratio over the compressor's operating range. The INL is also used to limit the gain of the compressor, so that system noise will not be amplified to the point of audibility.

There are a number of inherent limitations to the success of single-channel compressors. One might design a compressor in which the envelope detector's integration window is relatively long in duration (e.g., a duration equal to that of three cycles of the lowest expected frequency component). In so doing, a designer would be hoping to reduce spectral distortion. However, if the input signal is relatively broad-band compared to the bandwidth of the lowpass filter, relatively large and rapid changes in output level can occur. Because the integration window is long in duration, the compressor's gain will be sluggish in its response to relatively rapid level changes. As a consequence, the compressor's output level will rapidly change by approximately the same ratio as the change in ratio at the input: in other words, very little compression will occur during these intervals. This is considered a temporal distortion because the signal level greatly "overshoots" or "undershoots" its steady-state value. The designer has sacrificed temporal fidelity for spectral fidelity. Even brief amplitude overshoots can severely degrade the usefulness of the processor[a].

At the other extreme, the designer could choose to use a very short duration integration window to improve the rate at which the compressor's gain can change, thereby allowing the compressor to compensate quickly for rapid changes at the input. However, this approach generates a large amount of spectral distortion when relatively low-frequency spectral components are present. Harmonic distortion and intermodulation distortion products are both generated.

Single-channel compressors using a feedback gain-control path (Fig. 2) have essentially the same characteristics as the feedforward configuration discussed above. However, in the feedback configuration it is not possible to exactly synchronize the gain-control signal with the signal whose gain is controlled. The lag between the envelope estimate and the input signal will generate additional distortion.

Many standard compression systems use separate "attack" and "release" integration windows. Generally, a relatively short integrating time-constant is used during the attack interval (i.e., during the segments in which the envelope is increasing) compared with the time-constant used during the release of compression. Such compressors generate both spectral and temporal distortions. The spectral distortion is generated primarily during the fast-attack phase and is particularly apparent with complex stimuli such as speech. The long release phase is plagued with drop-outs or undershoots when the input signal abruptly decreases in level. The compressor's output level can drop well below threshold before the compressor's sluggishly responding gain can increase.

Robustness to noise in single-channel compressors — Single-channel compressors perform poorly in many noisy environments. Without compression, noise which has high energy only within relatively narrow spectral regions will mostly mask the speech signal only in and around those narrow spectral regions; the other spectral regions will be relatively free of interference. But when such noise is added to the input of a single-channel compressor, all frequency regions of the speech signal are attenuated

[a]In those subjects (and in those operating regions) where loudness rapidly increases with small changes in level, even very-short-duration overshoots can cause very dramatic changes in the subject's responses. A rapid increase in the loudness-versus-amplitude function causes the amplitude (necessary for a fixed loudness) versus stimulus duration curve to be relatively flat. For example, in many cochlear implant subjects, extremely brief overshoots can cause the stimulus to be perceived as uncomfortably loud. In many of these subjects, the greatest rate of loudness change occurs at the higher levels.

equally. For example, a single high-amplitude "interfering tone" would cause the entire speech spectrum to be severely attenuated. Even spectral components very distant in frequency from the tone would be severely attenuated. Potential information is lost when these components are severely attenuated, since these more distant spectral components would normally (without compression) be relatively unmasked by the tone.

Certain multichannel compressor designs offer a more appropriate response to such interfering signals than do the single-channel types that we have been discussing.

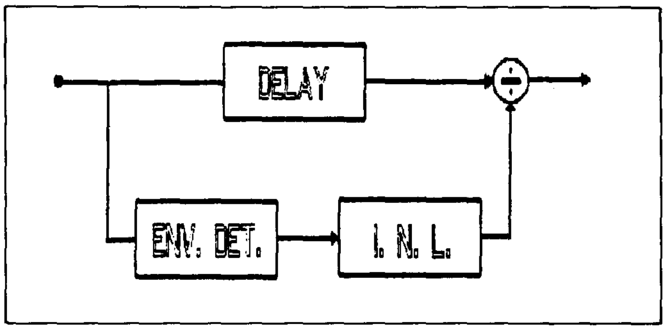


FIGURE 1
A feedforward single-channel compressor (block diagram).

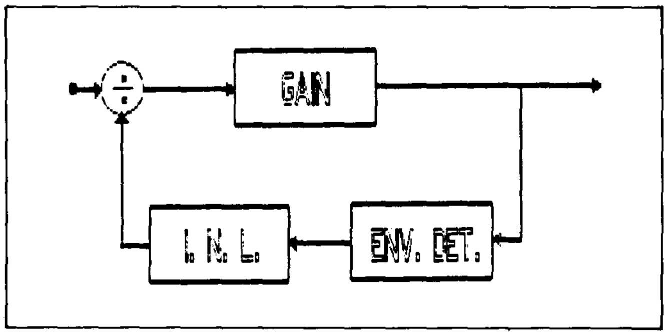


FIGURE 2
A single-channel compressor which uses a feedback path to control the system's gain (block diagram).

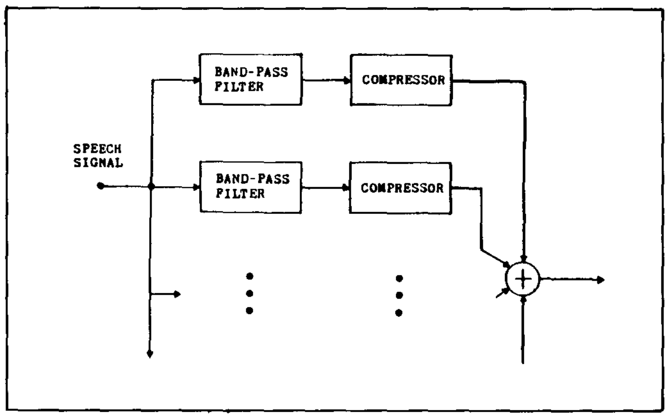


FIGURE 3
A standard multichannel compressor configuration. (Each compressor block represents a single-channel compressor.)

Multichannel Compressors

Multichannel compression systems, shown in a highly simplified form in Figure 3, offer a number of significant advantages.

A subject's dynamic range characteristics are often a strong function of frequency. In addition, the statistics of the input signal can be a significant function of frequency. In a multichannel system, each channel's compressor can be "customized" to map that channel's range of inputs most appropriately into the particular operating range for that channel. Potentially, both the steady-state and the transient characteristics of the input and output could be better matched with the use of a multichannel system.

For the higher frequency channels, channel gains can change at relatively rapid rates without causing spectral distortion. Short-duration integrating windows can be used, because the cycle durations of the lowest-frequency components in the channels are quite short. For most applications, compressor integration intervals on the order of 1.5 to 5 cycles of the lowest frequency within the channel's passband are likely to produce reasonably low levels of spectral distortion and of amplitude undershoots and overshoots.

In Figure 3, the narrow-band filters shown in series with the single-channel compressors can temporally disperse or "smear" rapid amplitude transitions. This is a disadvantage in most applications, where the output signal should accurately reflect the relative time of occurrence of transitions in the spectra. In the systems of Figure 3, the length of the output transition is the sum of (i) the duration of the input transition, (ii) the bandpass filter's impulse response, and (iii) the compressor's impulse response. Temporal dispersion can be reduced by using wider-bandwidth filters; however, if the bandwidths are too wide, the benefits of the multichannel approach will be compromised.

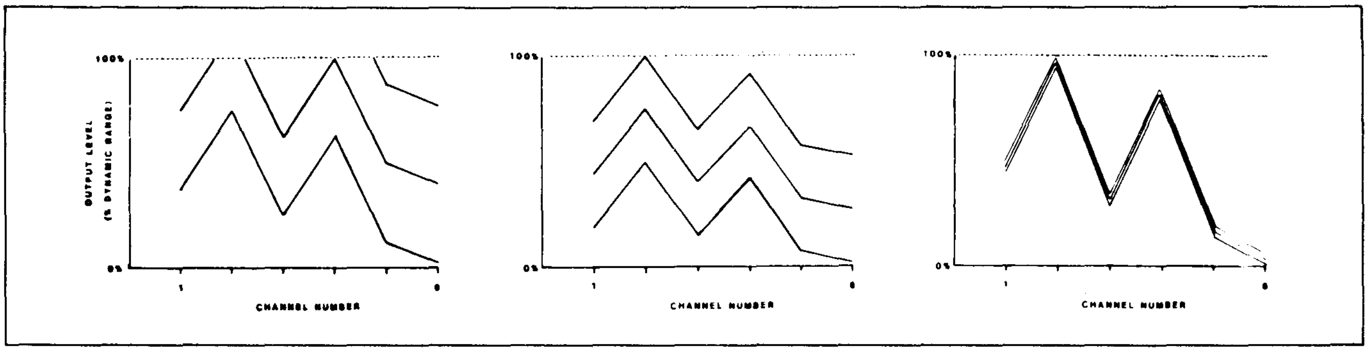


FIGURE 4
Schematic plots of hypothetical responses to a steady-state vowel stimulus. Each graph illustrates the response of a different six-channel compressor to the same vowel presented at three different levels. The output level of each channel is plotted as a percentage of the channel's dynamic range.

There is an additional problem that needs to be solved when using multichannel compression systems. With the multichannel compressor of Figure 3, spectral patterns developed across the channels, and average intensity, are both compressed equally; with this type of compressor it is not possible to compress spectral patterns to a lesser (or greater) degree than the average-intensity information[b]. By examining the responses of three hypothetical multichannel compressors, we can obtain a better understanding of the problem. Each graph in Figure 4 illustrates the steady-state channel output levels of a different six-channel compressor. Each hypothetical compressor is responding to the same "steady-state vowel" stimulus. Each curve represents the compressor's response to that vowel at one of three average signal levels. (Average signal level refers to the signal level derived across the entire speech spectrum.) The average signal level varies because of differences among speakers, difference in speakers' distance from the microphone, such differences as whispered speech levels versus loudly-spoken speech, and differences in the levels within a given speech utterance (e.g., vowels are generally higher in level than are consonants).

In Figure 4, each channel's output level is plotted as a percentage of the channel's dynamic range. Graphs A and B of Figure 4 represent the outputs of a compressor like the one in Figure 3 but the amount of compression in 4B is considerably greater than that in 4A. For the two higher average-input levels in Figure 4A, some of the channels would be perceived as uncomfortably loud (i.e., some channel outputs are over 100 percent on the loudness scale). Because of the greater amount of compression in 4B, none of the channels in 4B are stimulated above their uncomfortable loudness level (UCL) nor do they drop below threshold. However, this large amount of compression on each channel can significantly reduce the amount of across-channel spectral information available to the subject. The decreased spectral information is represented in Graph B of Figure 4 by the reduced differences between the channel output levels[c]. At the highest compression ratios, the output levels of all the channels would remain nearly constant and therefore would convey little or no information (e.g., see Figure 14, Graph A, and the related discussion). To further aggravate the problem, only a relatively small number of intensity-difference limens are available across the restricted dynamic range of many subjects; see Muller, 1983 (1).

Multichannel compression: differential compression of spectral and average-intensity information — The shape of the short-term speech spectrum is considered to be very valuable in the perception of speech: see Stevens, 1983 (2); and Pickett, 1980 (3). In contrast, both the short-term and long-term average intensity of speech appear to play lesser roles in speech perception, as noted by Hendrickson in 1982 (4). As a consequence, it may be advantageous to compress spectral information far less than we compress average-intensity information. We could "decompress" the across-channel spectral representation in Graph B of Figure 4 and thereby utilize nearly the full dynamic range of the transmission channel for the representation of spectral information. Graph C of Figure 4 schematically illustrates

[b]In marked contrast, the single-channel compressor of Figure 1 compresses only the average-intensity information and does not compress the spectral information. In this sense, the single-channel compressor is capable of transferring more spectral information than the multichannel system of Figure 3.

[c]"Across-channel spectral information" is the spectral information that is represented by the differences in the channel-output levels. Information is communicated when these patterns change over time. (The changing shape of the spectrum within the passband of a processor channel is defined as "within-channel spectral information".)

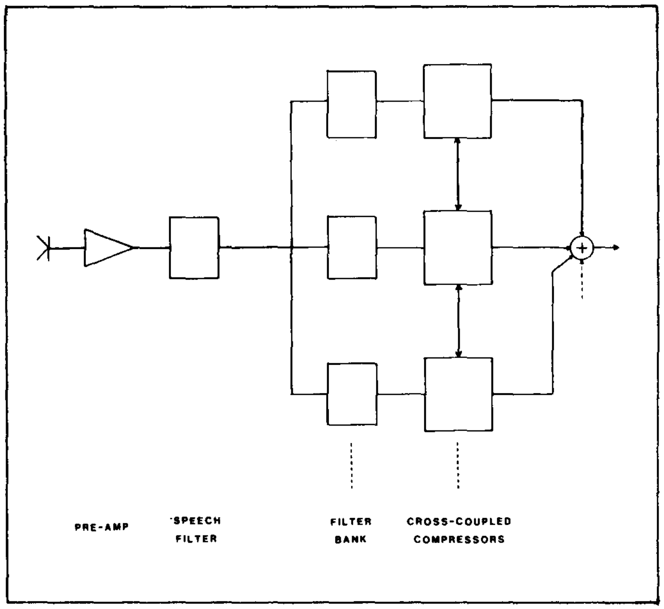


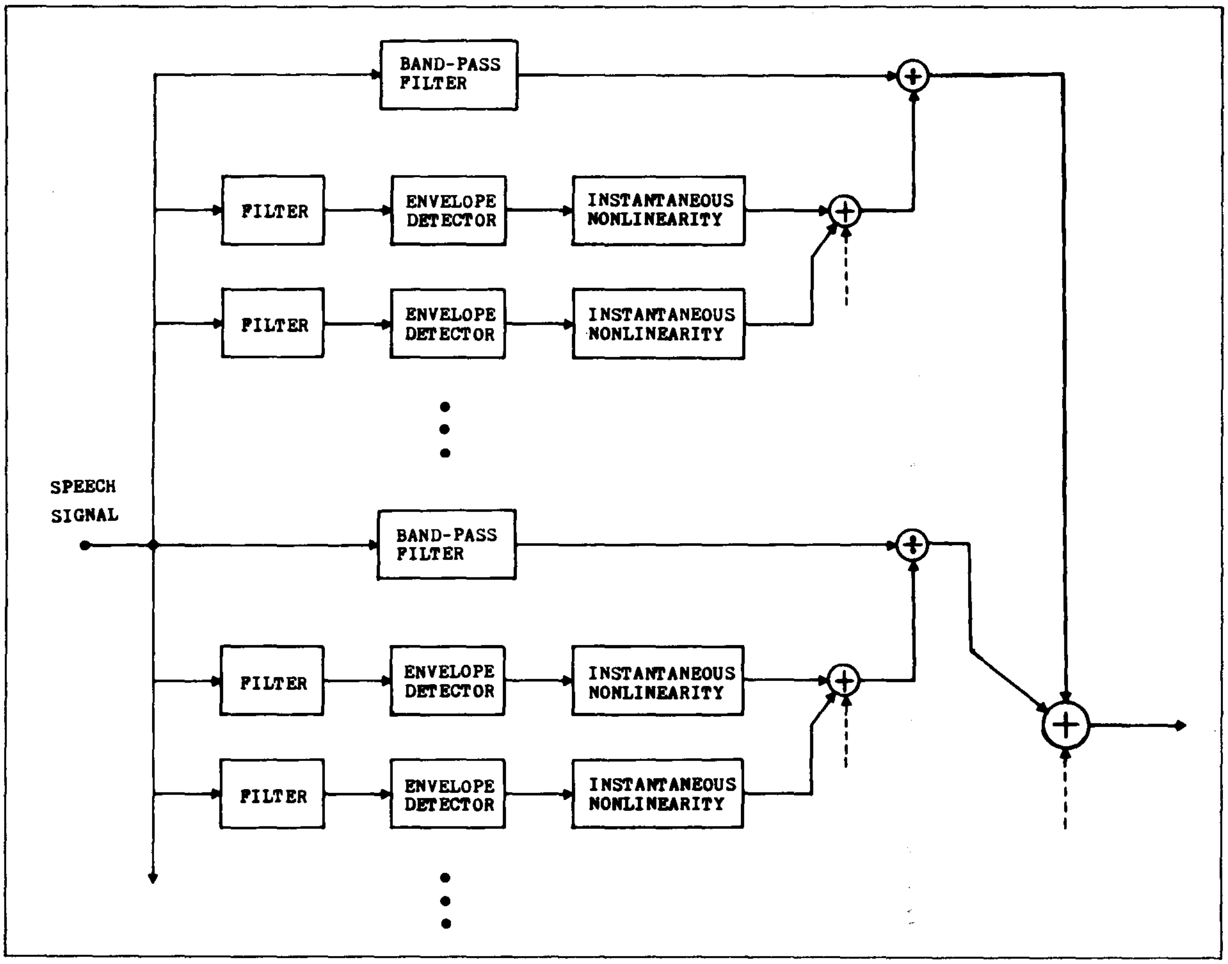
FIGURE 5
Configuration of a cross-coupled multichannel compression system. Included in this block diagram are symbols representing a microphone, a pre-amplifier, and a speech pre-emphasis filter, at left.

the substantial improvement in the across-channel spectral representation that can be obtained in this way.

Figure 5 is a diagram of a "cross-coupled" compressor configuration, in which each compressor's gain is controlled by a weighted combination of the signal levels derived from a set of the channels. A given compressor's gain will be caused to decrease when signal levels increase in those channels that control the channel's gain. A compressor with such a cross-coupled configuration does not generate spectral distortion at any point before the spectral analysis stage (i.e., the filter bank), and it allows the emphasis to be placed on across-channel spectral information relative to average-intensity information.

Figures 6 and 7 are more specific examples of such compressors. In Figures 6 and 7, each chan-

FIGURE 6 (BELOW)
A cross-coupled compressor configuration. Separate filters and envelope estimators are used to control the gain of each channel. The configuration can be made even more versatile if an instantaneous nonlinearity is placed between the summing junction and the divider in each channel.



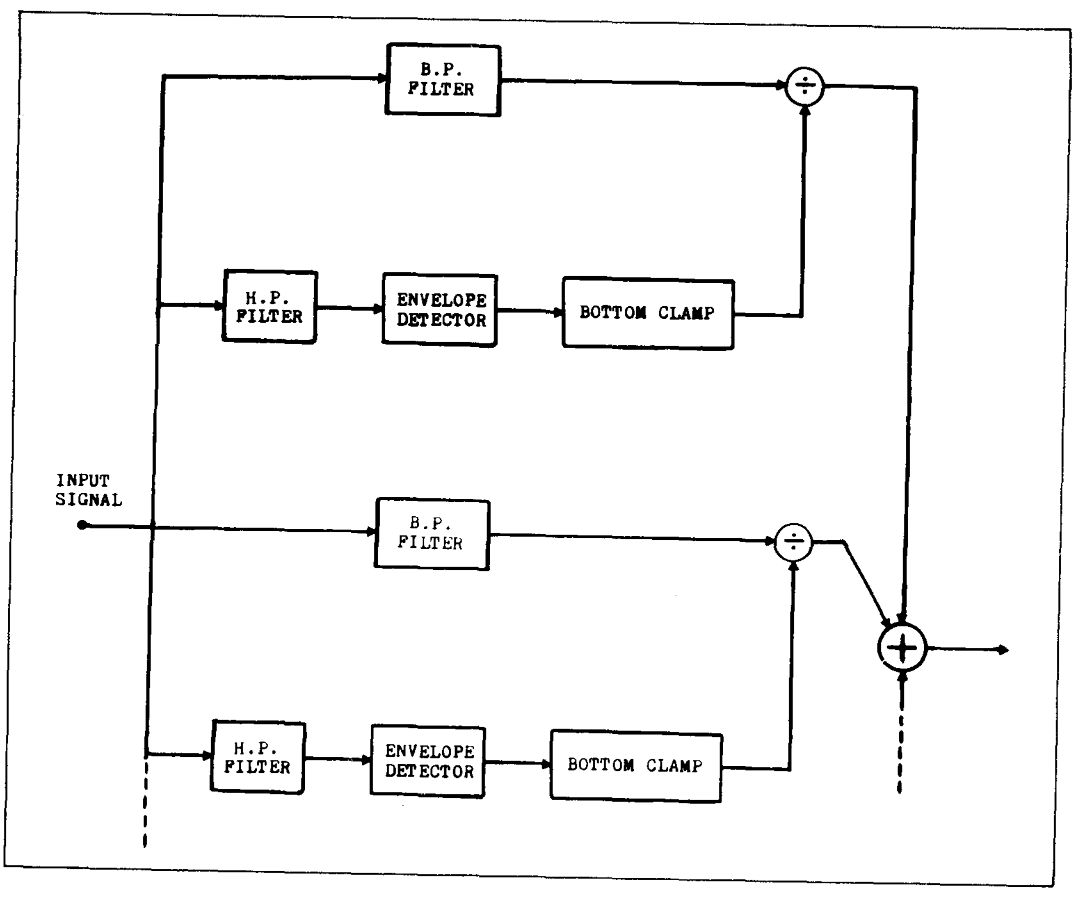
nel's gain is independently derived from a weighted sum of the signal levels derived in one or more frequency ranges. For example, if a filter preceding one of the envelope detectors has a higher gain for a certain band of frequencies, this band will be particularly capable of attenuating that channel's output amplitude.

In all of the compressors illustrated, a "divider" system is used to perform the gain-control function. The divisor should be strongly and positively correlated with the signal feature or features to be compressed. Any changes in the dividend which are correlated with this feature (or features) are reduced (i.e., compressed) by the division. If the divisor is more positively correlated with the input's average intensity than it is with the individual channel's intensity, average-intensity changes will be compressed more than will across-channel level differences.

The transient characteristics as well as the static characteristics of the across-channel interactions must be appropriately controlled[d]. If spectral distortions and temporal overshoots and undershoots are all to be minimized, each envelope estimator's rate of response should be appropriately matched to the frequencies within the channel's passband. To minimize spectral distortion, the gain-controlling signal should not change too rapidly during a cycle of any spectral component within that channel's passband, but to minimize temporal distortions, the gain-control signal should change quickly enough to compensate for rapid changes in level at the system's input. These opposing conditions constrain the design of such compressors. If the compression system were improperly designed, a very fast increase in the amplitude of a high-frequency tone could very quickly reduce the gain of the lower frequency channels and produce a sudden drop in the output waveforms of the low frequency channels, thus generating severe spectral distortion within the low-frequency channels. The rate at which the gain-control signal changes should be constrained

[d]It is possible to emphasize transitions of the speech spectra. An adaptive mechanism (i.e., a mechanism similar to that described by Smith et al. in 1983 (5) could be inserted at the output of each channel. A second method is also available; with "cross-coupled" or "band-coupled" compressors, formant frequency transitions can be emphasized by the appropriate control of the transient features of the across-channel gain-control signals.

FIGURE 7
Configuration of the cross-coupled compressor that was simulated.



by using sufficiently long-duration integrating intervals for controlling the gain of the lower frequency channels. The integrating interval should be from at least 1.5 to several cycles of the lowest frequency components within the channel. Equivalently, in the frequency domain, the lowpass filters of the envelope estimators should pass only frequencies less than those in the passband of the controlled channel. Additional constraints on the lowpass filter design are discussed in the following sections.

COMPARISON OF COMPRESSION SYSTEMS

Simulations

A multichannel processor was simulated in software, incorporating some of the features considered useful for a cochlear prosthesis. Figure 7 is a block diagram of the example (simulation) processor. In this processor, the outputs of the channels are summed to generate one composite output signal. Only one branch per bandpass channel is used to estimate the signal amplitude over a relatively wide band of frequencies as determined by the highpass filter — this is a simplification of the more common configuration (Fig. 6) which uses multiple branches to control the gain of each channel. (The pre-envelope-detector filter will be referred to as the "PED filter". In the compressor of Figure 6, each branch can have a different combination of lowpass cutoff and PED filter frequency response.)

For comparison, two single-channel compressors were also simulated, one being given a relatively long-duration integrating window and the other a much shorter integrating window.

Compressor specifications

In all the simulated compressors, the envelope detector comprises a full-wave rectifier followed by a lowpass filter. In all these compressors, the impulse response of the lowpass filter was a Kaiser window described by Rabiner and Schafer in 1978 (6). Although a flatter frequency response over the passband would have been preferable, the simpler Kaiser lowpass filter was used in the current set of simulations. The instantaneous nonlinearity (INL) immediately following the envelope detector was a "bottom-clamp". With this memoryless nonlinearity, the output signal equals the input signal except when the input signal falls below a specified minimum value; if that happens, the output signal is held at the specified value until the input signal returns to greater values. For signal levels above this minimum or threshold value, the compressor maintains a constant output level in the steady state. This is the special case of "an infinite compression ratio", where no average-intensity information is transmitted. (This is too much compression for many applications; it is also a most demanding requirement for a compression system.) To obtain lower compression ratios, the INL's input-output function should be appropriately contoured for the desired steady-state compression function.

TABLE 1
Specifications of the lowpass filters of two simulated, single-channel compressors. Sampling rate was 20,000 samples/sec.

	Window duration	
Description of Variable	Long	Short
−6 dB Freq. of L.P.	100	694
L.P. stopband ripple (dB)	30	39
Number of L.P. Coef.	207	39
Transition width (Hz)	200	1387

TABLE 2
List of those processing parameters that do not vary across channels in the multichannel compressor of Figure 7.

Description of Variable	All channels
Sampling rate	20000
Maximum gain per channel	1
Bandwidth of B.P. filters	300
B.P. stopband ripple (dB)	40
Number of B.P. Coef.	371
Transition width of B.P. (Hz)	150

The specifications for the lowpass filters of the two single-channel compressor simulations can be found in Table 1. The delay stage (see Figure 1) in each of the two single-channel compressors was set to compensate exactly for the delay introduced by the lowpass filter in the envelope detector. As a consequence, the estimate of the signal level was in synchronization with the input signal.

The multichannel compressor of Figure 7 is composed of 15 contiguous equal-bandwidth channels (six of them are represented in Figure 8). This large number of channels is not essential for many of the benefits derived from such compression systems. In fact, processors using only two to four channels should be useful for many applications. Table 2 is a list of those processing parameters that do not vary across the channels; Table 3 contains the parameters that do vary across the channels. Each channel's gain is controlled by a separate estimate of the input signal's intensity. The rate at which the intensity

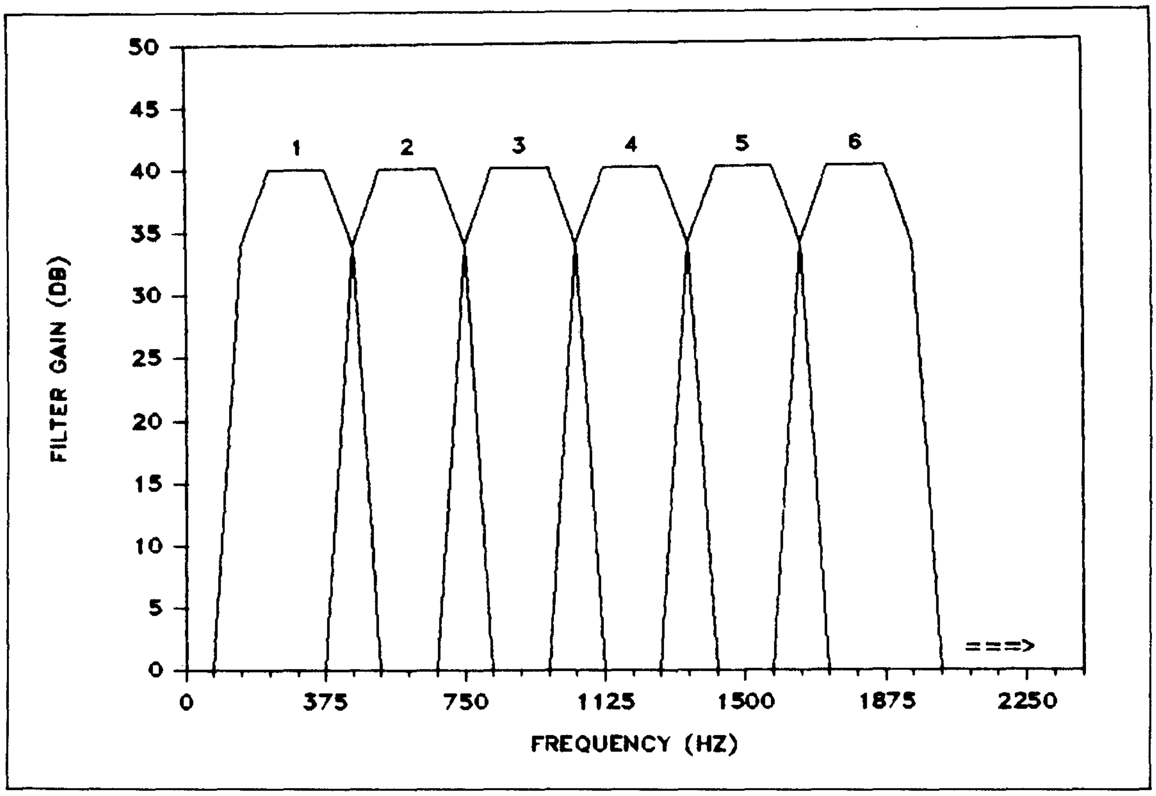


FIGURE 8
Schematic plot of the frequency response of six of the 15 filters used in the multichannel compressor of Figure 7.

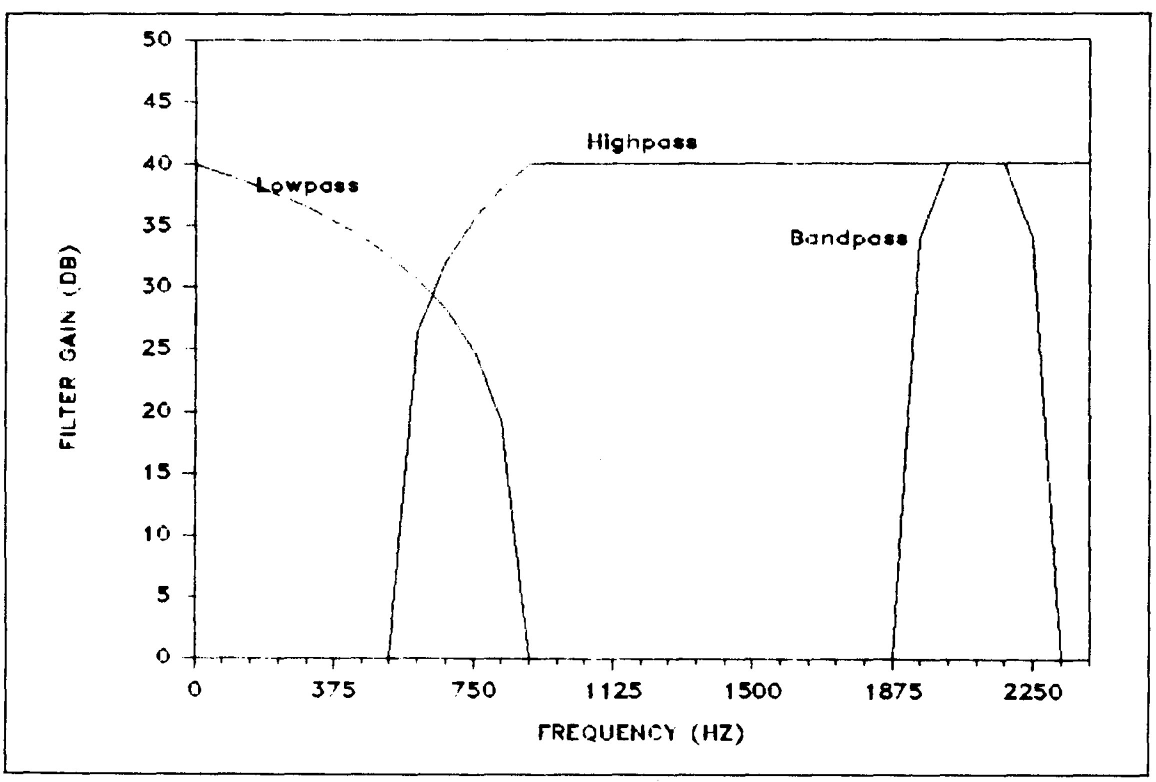


FIGURE 9
Schematic plot of the frequency responses of the three linear filters used in the seventh channel of the example (simulation) multichannel compressor of Figure 7.

TABLE 3
List of those processing parameters that do vary across channels in the multichannel compressor of Figure 7. (The two lowest-frequency channels do not contain highpass filters.)

Description of Variable	Channel 1	Channel 2	Channel 3	Channel 4	Channel 5	Channel 6	Channel 7
Bandpass Filter Section							
Center Freq. of B.P.	300	600	900	1200	1500	1800	2100
Lower −6 dB Freq. of B.P.	150	450	750	1050	1350	1650	1950
Upper −6 dB Freq. of B.P.	450	750	1050	1350	1650	1950	2250
Lowpass Filter Section							
−6 dB Freq. of L.P.	100	188	300	338	363	394	469
Minimum Freq. within L.P. Stopband	200	375	600	675	725	787	937
L.P. Stopband Ripple (dB)	30	30	30	32	35	37	40
Number of L.P. Coef.	207	111	69	67	67	65	59
Transition Width of L.P. (Hz)	200	375	600	675	725	787	937
Highpass Filter Section							
−6 dB Freq. of H.P.	No	No	450	506	543	590	702
Maximum Freq. within H.P. Stopband	H.P.	H.P.	330	371	398	433	515
H.P. Stopband Ripple (dB)			52	59	64	69	84
Number of H.P. Coef.			303	305	305	307	313
Transition Width of H.P. (Hz)			240	270	290	314	374

Description of Variable	Channel 8	Channel 9	Channel 10	Channel 11	Channel 12	Channel 13	Channel 14	Channel 15
Bandpass Filter Section								
Center Freq. of B.P.	2400	2700	3000	3300	3600	3900	4200	4500
Lower −6 dB Freq. of B.P.	2250	2550	2850	3150	3450	3750	4050	4350
Upper −6 dB Freq. of B.P.	2550	2850	3150	3450	3750	4050	4350	4650
Lowpass Filter Section								
−6 dB Freq. of L.P.	544	619	694	769	844	919	994	1069
Minimum Freq. within L.P. Stopband	1087	1237	1387	1537	1687	1837	1987	2137
L.P. Stopband Ripple (dB)	40	40	39	39	40	38	39	38
Number of L.P. Coef.	51	45	39	35	33	29	27	25
Transition Width of L.P. (Hz)	1087	1237	1387	1537	1687	1837	1987	2137
Highpass Filter Section								
−6 dB Freq. of H.P.	815	927	1040	1152	1265	1377	1490	1602
Maximum Freq. within H.P. Stopband	598	680	763	845	928	1010	1093	1175
H.P. Stopband Ripple (dB)	100	116	132	149	164	181	197	213
Number of H.P. Coef.	321	327	333	337	339	343	345	347
Transition Width of H.P. (Hz)	434	494	554	614	674	734	794	854

estimates can change increases with the higher frequency channels.

Figure 9 illustrates the relationships between the lowpass, highpass, and bandpass filters in a typical channel (i.e., channel number 7). So that the compressor can respond quickly to changes in level, a high cutoff frequency for the lowpass filter is often desirable. To minimize distortion, the lowpass filter should pass only frequencies which are components of the envelope signal: therefore, the low-frequency edge of the non-envelope components is an upper limit for the lowpass filter's passband. The low-frequency edge of the non-envelope components is determined by the pre-envelope-detector (PED) filter and the rectifier nonlinearity. With a full-wave rectifier or a "square-law" nonlinearity, non-envelope components are generated at higher frequencies than with a half-wave rectifier (7). The PED filter rejects those low-frequency components which would, after rectification, generate non-envelope components within the passband of the lowpass filter. This is one of the several functions of the PED filters[e].

Simulation Results

Responses to a quickly changing high-frequency signal are illustrated in Figure 10. The input signal was a 3-kHz signal whose amplitude was increased by 10:1 over a 2-msec interval (see waveform A in Figure 10). Waveform B represents the rather poor transient response of the single-channel compressor which utilized a long-duration integration window. The multichannel compressor and the single-channel compressor with a short-duration integration window responded considerably more rapidly (see waveforms C and D in Figure 10). The short-duration integrating window used in this single-channel compressor is identical to the window used in the 3-kHz channel of the multichannel compressor. Therefore, it is not surprising that the latter two compressors exhibit similar responses for this modulated 3-kHz stimulus.

For a 200-Hz input signal, chart A in Figure 11 illustrates the severe harmonic distortion generated by the single-channel compressor with the short-duration integrating window.

The multichannel compressor generated considerably less spectral distortion (see chart B in Figure 11), and the other single-channel compressor (with a long-duration window) also exhibited very little distortion. Those two compressors exhibited nearly identical responses for the 200-Hz stimulus. (The long-duration integrating window used in this single-channel compressor was identical to that used in the 200-Hz channel of the multichannel compressor.)

The single-channel compressor with the short-duration window also generated a considerable amount of intermodulation distortion (see chart A in Figure 12). The input signal consisted of three equal-amplitude components: 200 Hz, 700 Hz, and 1500 Hz. With the same input signal, the multichannel compressor generated considerably less spectral distortion (see chart B of Figure 12).

Another type of multichannel compression system (Fig. 3) was also simulated. The bandpass filter specifications were identical to those in Tables 2 and 3. All the compressors in Figure 3 are identical

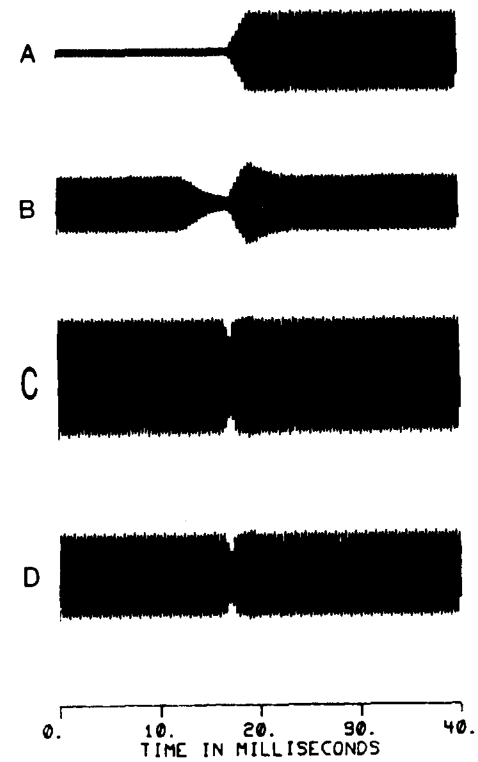


FIGURE 10
Input signal amplitude plot (A) and responses of three (simulated) example compressors. Plot B is the response of the single-channel compressor with the long integrating window. Plot C is the response of the single-channel compressor with the short integrating window. Plot D is the response of the cross-coupled multichannel compressor of Figure 7. (All three response plots in this figure have been shifted to the left in time, relative to the plot (A) of the input signal. The responses were shifted to "compensate for" the fixed delay within the compressors, thus simplifying the visual presentation.)

[e]How the PED filter controls the relative compression of local and more widespread spectral patterns, and how it controls the compression of average-intensity information, will be discussed later.

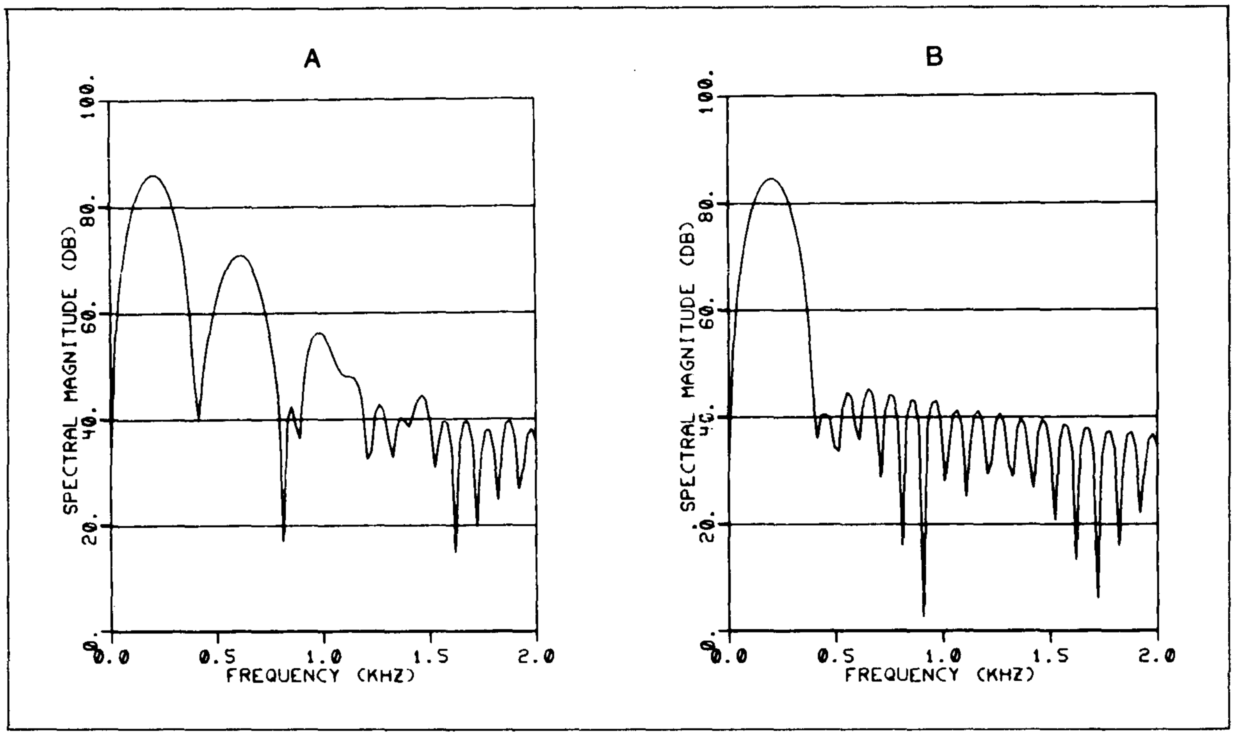


FIGURE 11
Frequency domain responses to a 200-Hz steady-state input signal. A is the response of the single-channel compressor with the short-duration integration window. B is the response of the cross-coupled multichannel compressor. All magnitude spectra were estimated using a 10-msec Hamming window.

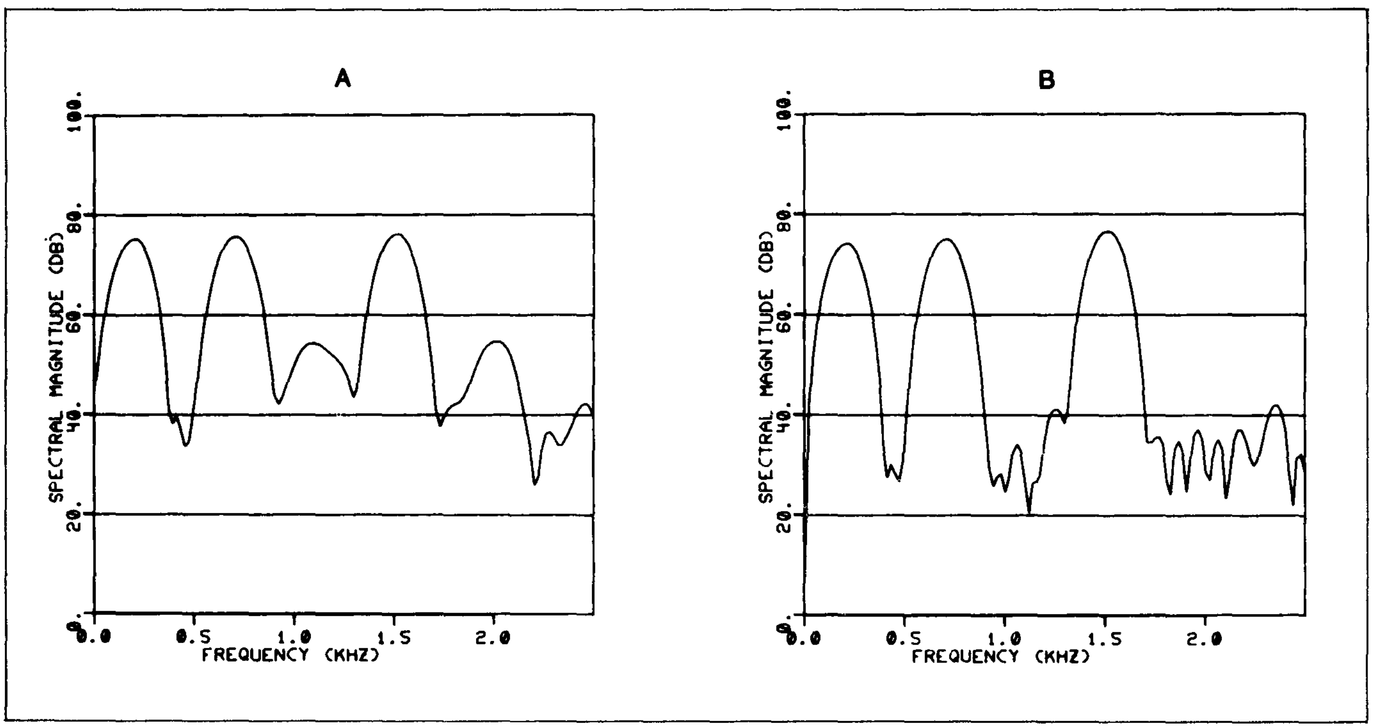


FIGURE 12
Responses to a steady-state complex of three equal-amplitude spectral components at 200, 700, and 1500 Hz. A is the response of the single-channel compressor with the short-duration integration window; B is the response of the cross-coupled multichannel compressor. All magnitude spectra were estimated using a 10-msec Hamming window.

to the single-channel (simulated) compressor which uses the longer duration integration window (Fig. 1 and Table 1). Because these single-channel compressor parameters do not vary across the channels and because the configuration is also "standard" (i.e., Figure 3), this is a good example of what is normally considered a "standard multichannel compressor."

For the next simulation, the input signal was a combination of six spectral components: 500 Hz at 0 dB, 100 Hz at −20 dB, 1500 Hz at 0 dB, 2000 Hz at −20 dB, 2500 Hz at 0 dB, and 3000 Hz at −20 dB. The signal's amplitude was linearly increased by 10:1 over a 10-msec interval. Waveform graph A in Figure 13 is the time-domain representation of the input signal: graph B and graph C in Figure 13 show the responses of the two multichannel compressors. Both strongly compress changes in the average intensity; however, there are three major differences between the responses of the two compressors. First, there is considerably more undershoot and overshoot revealed in graph C. Second, the duration over which the response transition occurs is considerably longer in graph C than in graph B. And the waveform in graph C is quite different from the input waveform. The reason for the difference in waveshape is more clearly demonstrated in the frequency domain.

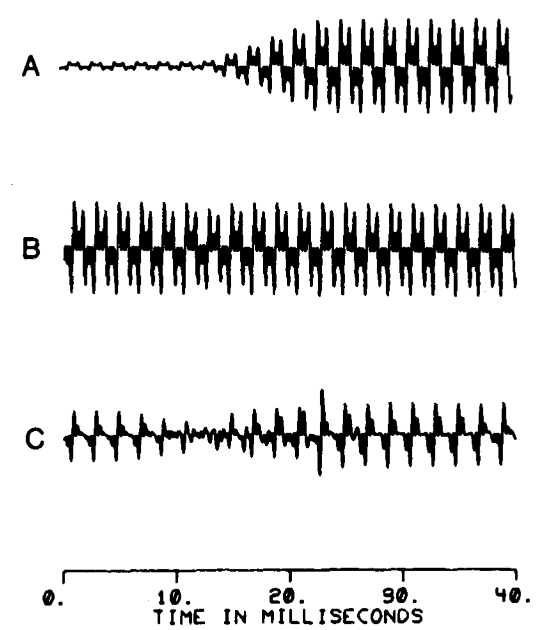


FIGURE 13
Input signal (A) and responses of two example (simulated) compressors. The input signal was composed of six spectral components: 500 Hz at 0 dB, 100 Hz at −20 dB, 1500 Hz at 0 dB, 2000 Hz at −20 dB, 2500 Hz at 0 dB, and 3000 Hz at −20 dB; where 0 dB represents a peak amplitude of 10. The input signal's amplitude was linearly increased by 10:1 over a 10-msec interval. B represents the response of the cross-coupled multichannel compressor illustrated in Figure 7. C represents the response of the standard multichannel system of Figure 3. Both response plots in this figure have been shifted to the left in time, relative to the plot (A) of the input signal.

In Figure 14, magnitude spectra of the steady-state portion of the responses are displayed. Graph A in Figure 14 illustrates the extreme compression of spectral information that is generated with the "standard multichannel compressor". In this example, the gain of those channels receiving the −20 dB components is proportionately larger than the gain for the channels receiving the higher amplitude components. As a consequence, the output spectrum of such compressors is relatively "flat" compared to the input spectrum[f].

Graph B in Figure 14 demonstrates that spectral contrasts can be substantially improved with the multichannel compressor of Figure 7. The response illustrated in Figure 14(B) is a reasonably accurate reproduction of the input signal. This compressor was designed not to compress across-channel spectral information (i.e., at least, not to compress spectral differences that occur over moderately large regions of the speech spectrum). However, this compressor was designed to infinitely compress changes in average intensity. The frequency response of the filter preceding the envelope detector (i.e., the PED filter) in each channel is used to control the relative amount of compression for the two types of information. For example, we could reduce the differences in the channel output levels (i.e., increase the compression of across-channel spectral information) by increasing the gain of each PED filter over the frequency region of the channel's passband.

The spectra (see Figure 15) of the transient portion

[f]In Figure 14A, there are small differences in the magnitudes of the spectral components. Because each channel's compressor has an infinite compression ratio, one might expect to find no amplitude differences between the spectral components. However, the channels not "tuned" to any of the six spectral components also contribute to the composite output of the compression system. The bandpass filters of those channels pass highly attenuated versions of the input signal in which some components are attenuated more than others. The single-channel compressor in each of these channels increases the amplitude of these components. When the outputs from these channels are summed with outputs from the channels "tuned to" the input components, some of the resultant component magnitudes are altered.

Journal of Rehabilitation Research and Development Vol. 23 No. 1 1986

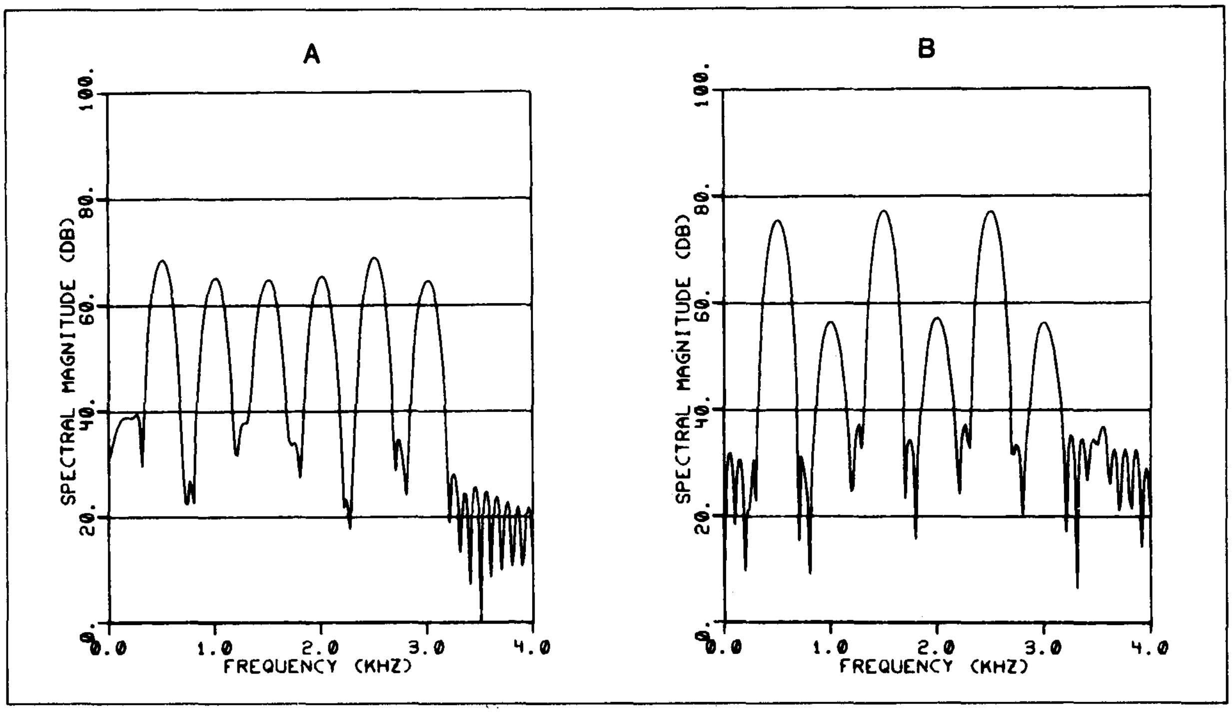


FIGURE 14
Magnitude spectra of the steady-state portion of the time-domain responses illustrated in Figure 13. A represents the response of the standard multichannel compressor of Figure 3; B represents the response of the cross-coupled multichannel compressor. All magnitude spectra were estimated using a 10-msec Hamming window.

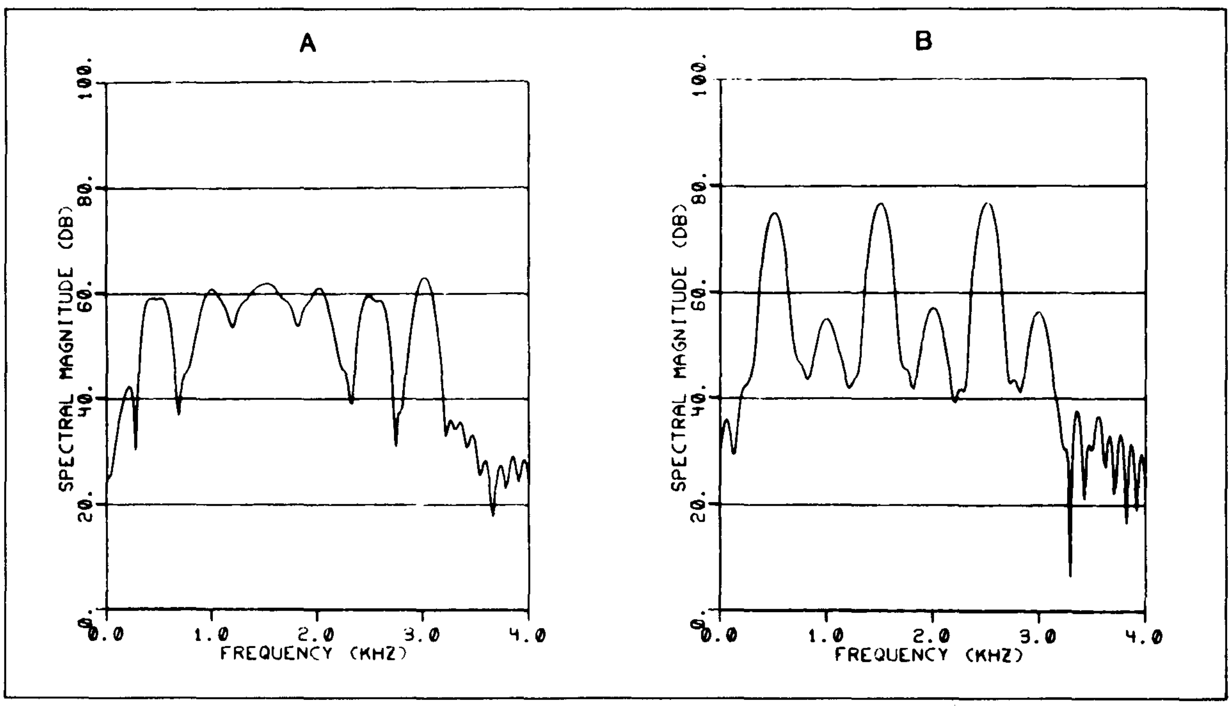


FIGURE 15
Magnitude spectra of the transient portion of the time-domain responses that are illustrated in Figure 13. A is the response of the standard multichannel compressor; B is the response of the cross-coupled multichannel compressor. All magnitude spectra were estimated using a 10-msec Hamming window.

of the responses are similar to the steady-state spectra in Figure 14, except that the spectral lobes in Figure 15(A) are considerably broader than in Figure 14(A). This seems reasonable, since there appears to be considerably more amplitude modulation in Figure 13(C) than in 13(B).

DESIGN CONSIDERATIONS

Robustness to Noise

Multichannel compressors can be designed to be more "robust" to noise than single-channel compressors. Without compression, those forms of noise that have more energy within relatively narrow spectral regions will primarily mask the speech signal in and around those spectral regions, while other spectral regions will be relatively free of interference. But with a single-channel compressor, when noise is added to a speech signal all frequency regions are attenuated equally, without regard to the spectrum of the interfering noise. For example, a single high-amplitude "interfering tone" could cause the entire speech spectrum to be attenuated below audibility; even spectral components of the speech that are very distant from the tone would be severely attenuated. Since these more distant spectral components would normally be relatively unmasked by the tone, it makes little sense to allow a device to attenuate this potentially useful information unnecessarily. In multichannel compressors like the one in Figure 7, a given channel's gain will not be affected by "distant" noise components if the pre-envelope-detector (PED) filter rejects "distant" spectral components. Because each highpass filter (i.e., the PED filter) cutoff frequency increases with the channel's frequency in the example compressor of Figure 7, the higher frequency channels will be relatively unaffected by low-frequency noise. By using bandpass PED filters with center frequencies equal to the channels' center frequencies, we could make the example (simulation) compressor of Figure 7 robust to a wide range of noise spectra as we decreased the PED filter's bandwidth. However, as the bandwidth of the PED filters is decreased, the differences in magnitudes of widely separated spectral components of speech will be compressed considerably more than the magnitude differences of more closely spaced spectral components. If the bandwidth of the PED filters is even further reduced, even local differences in spectral magnitudes can be severely compressed. (Also, quite narrow PED filters can restrict the speed of compression and could thus temporally disperse transition information.)

The relative importance in speech perception of the speech spectrum's "global" (or wide-spread) versus "local" features may be a critical factor in the design of robust compressors.

Temporal Disperson and Filter Bank Design

As previously discussed, the narrow-band bandpass filters in Figure 3 can significantly "smear" or temporally disperse rapid amplitude transitions. This is a limitation in applications where the compressed output should accurately reflect the relative time of occurance of spectral transitions. Fortunately, in the systems of Figures 6 and 7, the long impulse responses of the individual bandpass filters do not have any effect on the system's transient response[9]. If we held all the gain-control signals constant in Figure 7, the system's output would be a nearly perfect rendition of the input signal. Because the filter outputs are summed to form an undistorted wide-band channel, the transient responses of the bandpass filters have no effect on the transient response of the total compression system. Even though individual channels generate responses with severely "dispersed transitions", the composite response will be nearly identical to the input (i.e., if the channel gains are held constant). The temporally-dispersed components of the individual channel responses cancel each other when the filter outputs are summed.

The transient response of the envelope estimators determines the transient performance of the compression system. In Figure 7, both the lowpass filter and the highpass PED filter contribute to the transient response of the envelope estimators. Generally, the envelope detectors in the high-frequency channels will have shorter-duration transitions.

SUMMARY OF MAJOR FEATURES

The multichannel compressors illustrated in Figures 5, 6, and 7 have the following major properties:

1. As is also true of the "standard" multichannel compressor, in these systems the amount of compression can be varied as a function of frequency. One of the major disadvantages of single-channel compressors is the absence of this feature.

2. Across-channel spectral information can be compressed by a factor different from that for the

[9]This is true if the filter-bank is appropriately designed, as discussed by Rabiner and Schafer (6). Each bandpass filter should have the same constant delay at all frequencies. Also, the transfer function of the combined filter-bank outputs should be nearly "flat" in magnitude over the entire speech spectrum.

Journal of Rehabilitation Research and Development Vol. 23 No. 1 1986

compression of average-intensity information. In fact, spectral information can be emphasized with such compression systems. The filter preceding the envelope detector controls the relative amount of compression of the two types of information.

3. The compressor can be designed to be relatively robust to noise. Strong noise components can be prevented from reducing the compressor's gain for speech components distant from the noise components. This is not possible with single-channel compressors; single channel compressors perform quite poorly with noise.

4. Without sacrificing spectral fidelity, the gains of the higher frequency channels can change at a high rate to compensate for rapid amplitude changes at the compressor's input.

5. When summed, the channel outputs generate a composite signal with relatively little temporal dispersion. To obtain this desirable performance, all bandpass filters should have the same constant delay across the entire spectrum. Also, the frequency response of the filter-bank should be nearly flat over the speech spectrum ■

REFERENCES

1. Muller CG: Comparison of percepts found with cochlear implant devices. Annals of the New York Academy of Sciences 405:412-420, 1983.
2. Stevens KN: Acoustic properties used for the identification of speech sounds. Annals of the New York Academy of Sciences 405:2-17, 1983.
3. Pickett JM: The Sounds of Speech Communication. Baltimore, MD, University Park Press, 1980.
4. Henrickson L: Doctoral Dissertation. Stanford University, Stanford, California, 1982.
5. Smith RL, Brachman ML, Goodman DA: Adaption in the auditory periphery. Annals of the New York Academy of Sciences 405:79-93, 1983.
6. Rabiner LR and Schafer RW: Digital Processing of Speech Signals. Prentice-Hall, Englewood Cliffs, NJ, 1978.
7. Clarke KK and Hess DT: Communication Circuits: Analysis and Design. Reading, MA, Addison-Wesley Publishing Co., 1971.