Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2364091 A1 2003/05/29

(21) 2 364 091

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(22) Date de dépôt/Filing Date: 2001/11/29

(41) Mise à la disp. pub./Open to Public Insp.: 2003/05/29

(51) Cl.Int.^{7}/Int.Cl.^{7} H04L 12/24, H04L 12/56

(71) Demandeur/Applicant:
CATENA NETWORKS CANADA INC., CA

(72) Inventeur/Inventor:
VERREAULT, ERIC, CA

(74) Agent: OGILVY RENAULT

(54) Titre : SYSTEME ET METHODE DE COMPENSATION DES VARIATIONS DE RETARD DES PAQUETS VOCAUX

(54) Title: SYSTEM AND METHOD FOR COMPENSATING PACKET VOICE DELAY VARIATIONS

System And Method For Compensating Packet Voice Delay Variations

Background of Invention:

The invention applies to any continuous-time signals transported over a packetized medium, such as voice, music, voice-band or any other signals.

In applications where continuous time signals are packetized, packets are typically buffered at the receiving site and their play-out delayed in order to compensate for the variations of the network end-to-end delays. The play-out buffer operates simply by introducing an additional delay that allows the system to hold packets scheduled to be played-out later in time. Therefore, it offers a time window over which the network end-to-end delay can vary. In the case of a non real-time application, like audio or video streaming, the selected delay introduced by the buffer is, by design, set to a very large size; the size being chosen to minimize the probability of receiving late packets. In the case of a real-time application like video conferencing or audio conversation, large delays impair the usability of the system. Therefore, the delay should be minimized but it is an art to determine the exact delay to select such that the probability of late packets is acceptable.

In the case of a real-time system, if the selected delay is observed to be too small or too large, it is desirable to either increase it to minimize the number of late packets or to decrease it to make the system more transparent. A reduction in the delay will create excess packet(s) and an increase in the delay creates gap(s) in the play-out.

The proposed invention proposes a method that eliminates the discarding of packets and the gaps when the play-out delay is adjusted.

Detailed Technical Background:

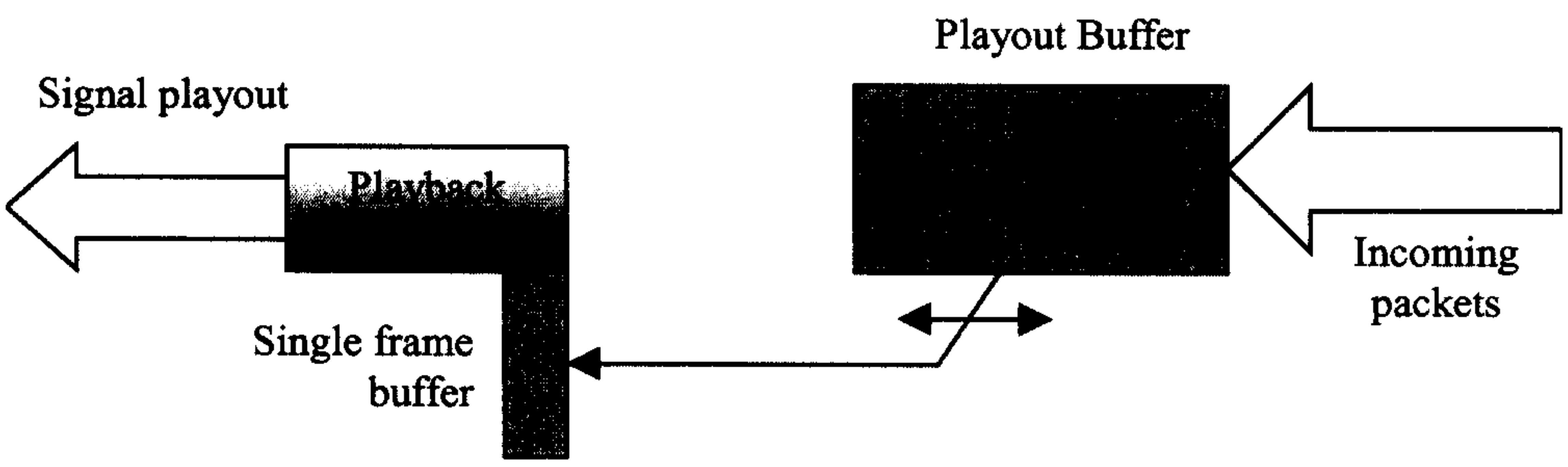
Most packetized voice applications buffer packets at the receiving end to delay the play-out. The inserted delay guarantees a time window over which the network end-to-end delay variations can be hidden. Two approaches are commonly used to set the delay:

either to set the delay once for the whole session or to adjust the delay from one talkspurt to the other. In both cases, everything is done to avoid adjusting the delay during non-silence period. It has been shown that adjusting the delay on a per-packet basis (within the talkspurt) introduces gaps and slips that are damaging to the quality of the audio [1,2].

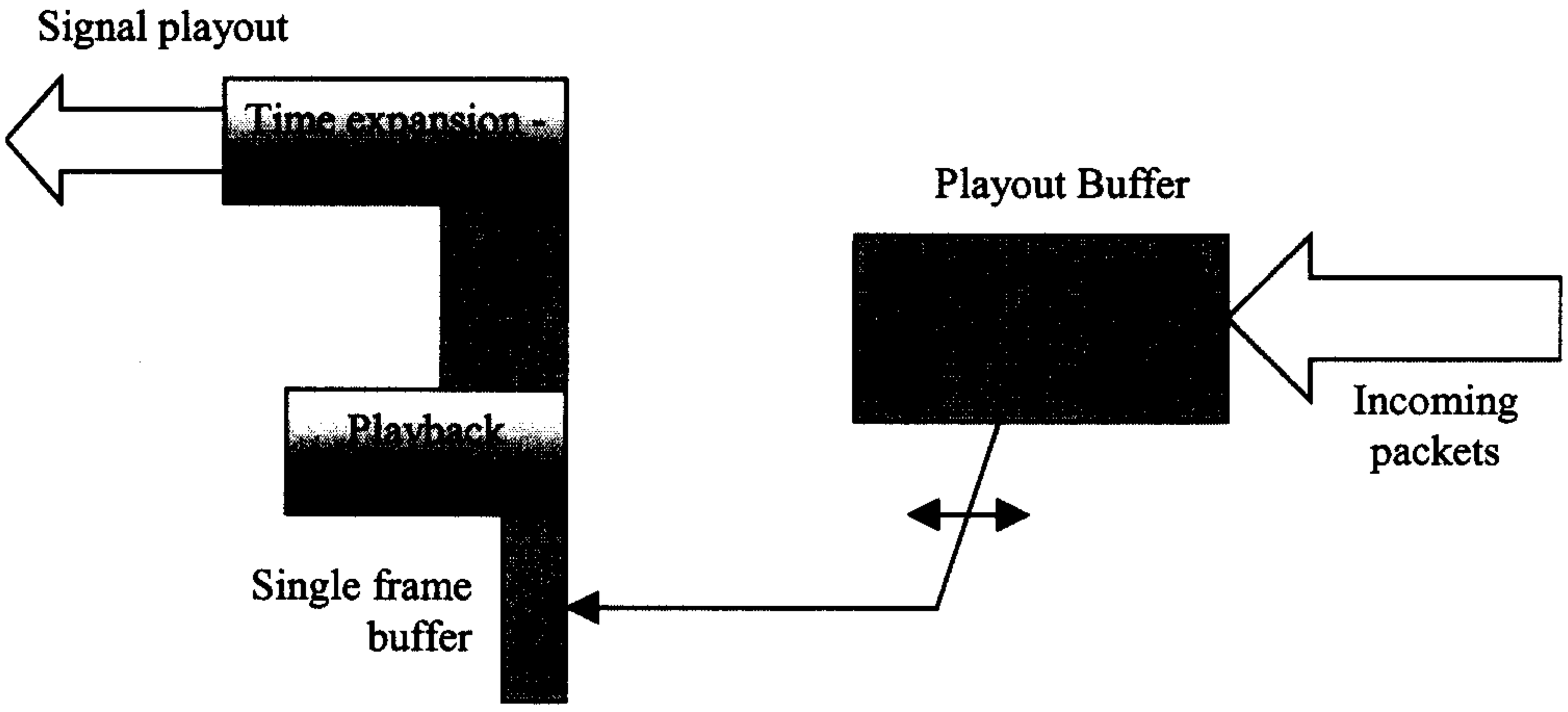
Description of Invention:

It is possible to adjust the play-out delay without introducing gaps or slips by speeding out or reducing the play-out rate. This is in effect compressing or expanding the time scale. If delay changes do not occur very often, then the time scale change is very small and the changes in the audio are barely detectable if not undetectable.

The following diagram shows the different components required in a playback system for packetized voice. On the right side of the diagram, packets are received from the network and their content is inserted in the play-out buffer providing a time delay to eliminate the end-to-end delay variations of the network. A play-out delay is selected and a voice frame is periodically extracted from the delay buffer and inserted in a single frame buffer. The voice frame carried over the network can be under a compressed form or as a direct sample per sample digitized representation. The playback element recreates the waveform.



The proposed delay compensation is an add-on over the standard set of components used in a playback system for packetized voice. The following diagram shows the insertion point of the new device within a standard playback system. At the output of the playback element, a time compression / expansion module is coupled to the playback element through a specialized buffer. In the normal steady state operation, the specialized buffer holds one or more packets of voice signal. If the change of delay is limited to a single packet at a time, then the specialized buffer only needs to hold a single packet of voice samples.

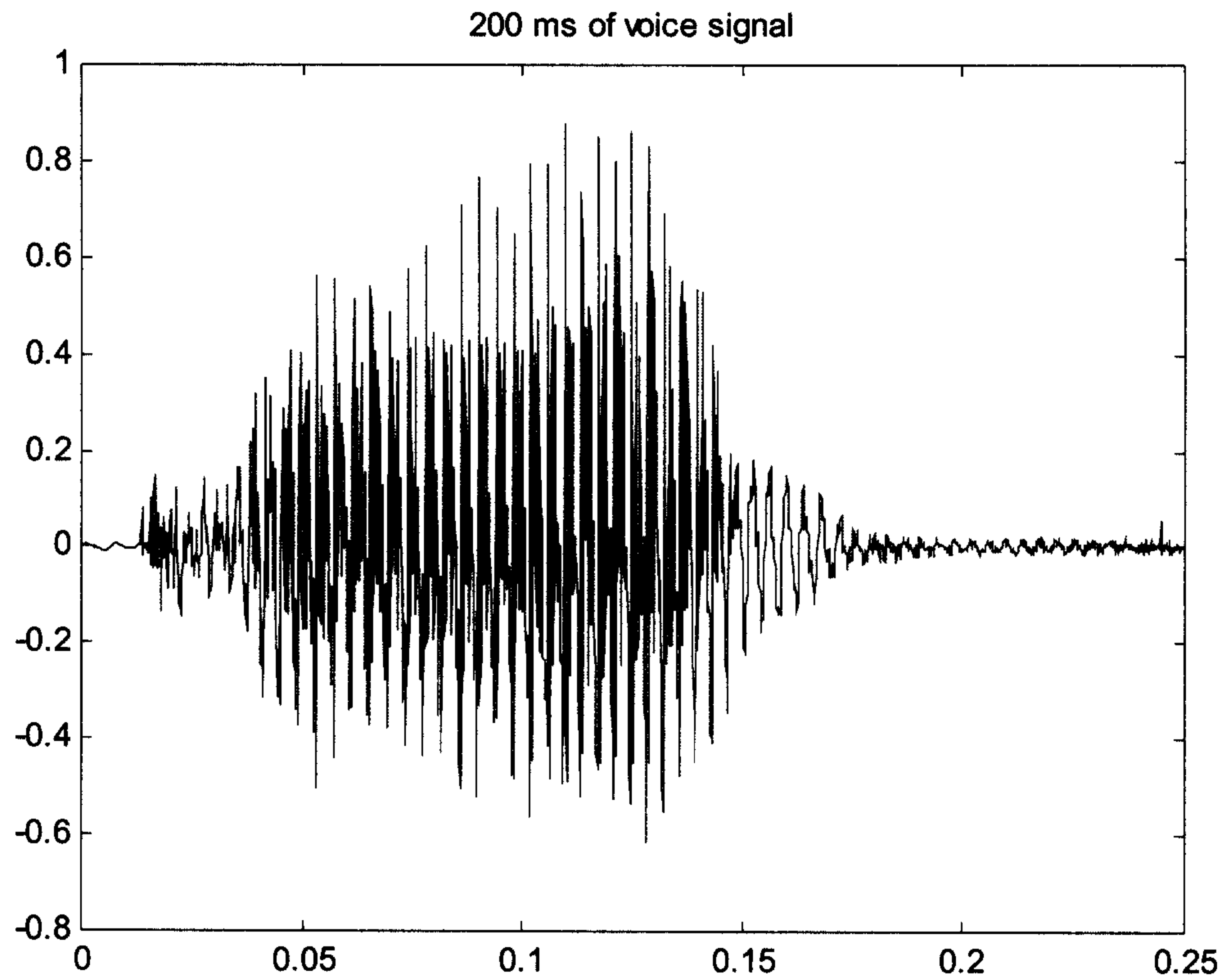


In the case of a reduction of an excessive delay introduced by the play-out buffer, two or more packets are sent to the playback element that creates the waveform for all of them. This requires the specialized buffer to hold two or more packet lengths of the signal to play-out. In reaction to the extra samples inserted in the specialized buffer, the time expansion / compression module can either detect this excess automatically or be programmed to start compressing the time scale. By compressing the time scale, the module plays out more samples per unit of time than in the normal steady state case. After a long enough time, the excess samples will have been played-out and the time expansion/compression module either automatically or via supervisory intervention returns to the normal time scaling.

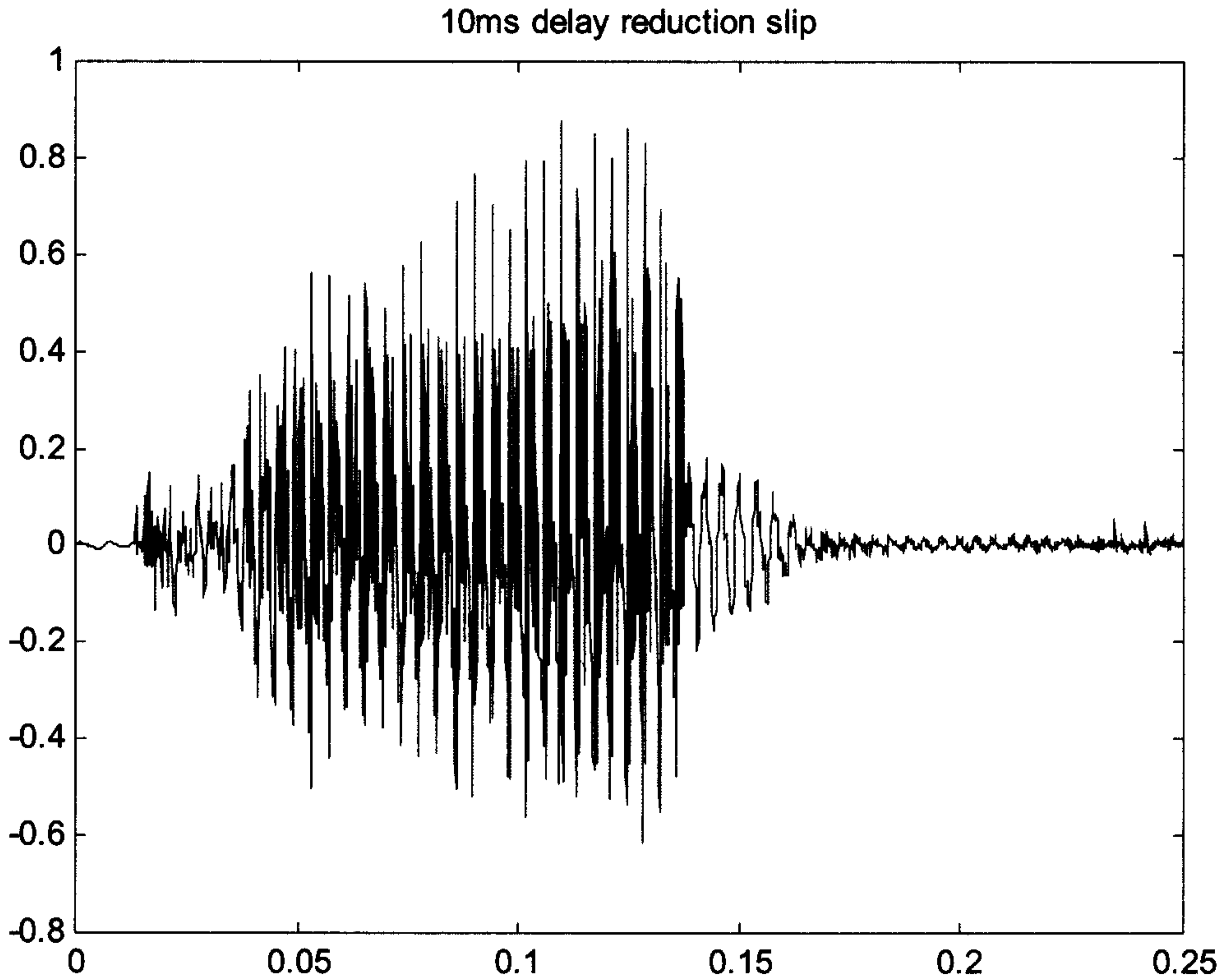
In the case of an increase in the delay of the play-out buffer, the playback element does not receive the sufficient number of packets to cover the delay slip performed. This exhausts the specialized buffer of its normal amount of accumulated packets. In reaction to the exhaustion, the time expansion / compression module can either detect this excess automatically or be programmed to start expanding the time scale. By expanding the time scale, the module plays out fewer samples per unit of time than in the normal steady state case. After a long enough time, the specialized buffer will hold the normal steady state number of packets and the time expansion/compression module either automatically or via supervisory intervention returns to the normal time scaling.

The time scale expansion and compression module can be implemented in various ways. In an all-digital system, this can be achieved through a cascade of digital interpolation and decimation. Another way is to adjust the playback system clock controlling the digital to analog conversion.

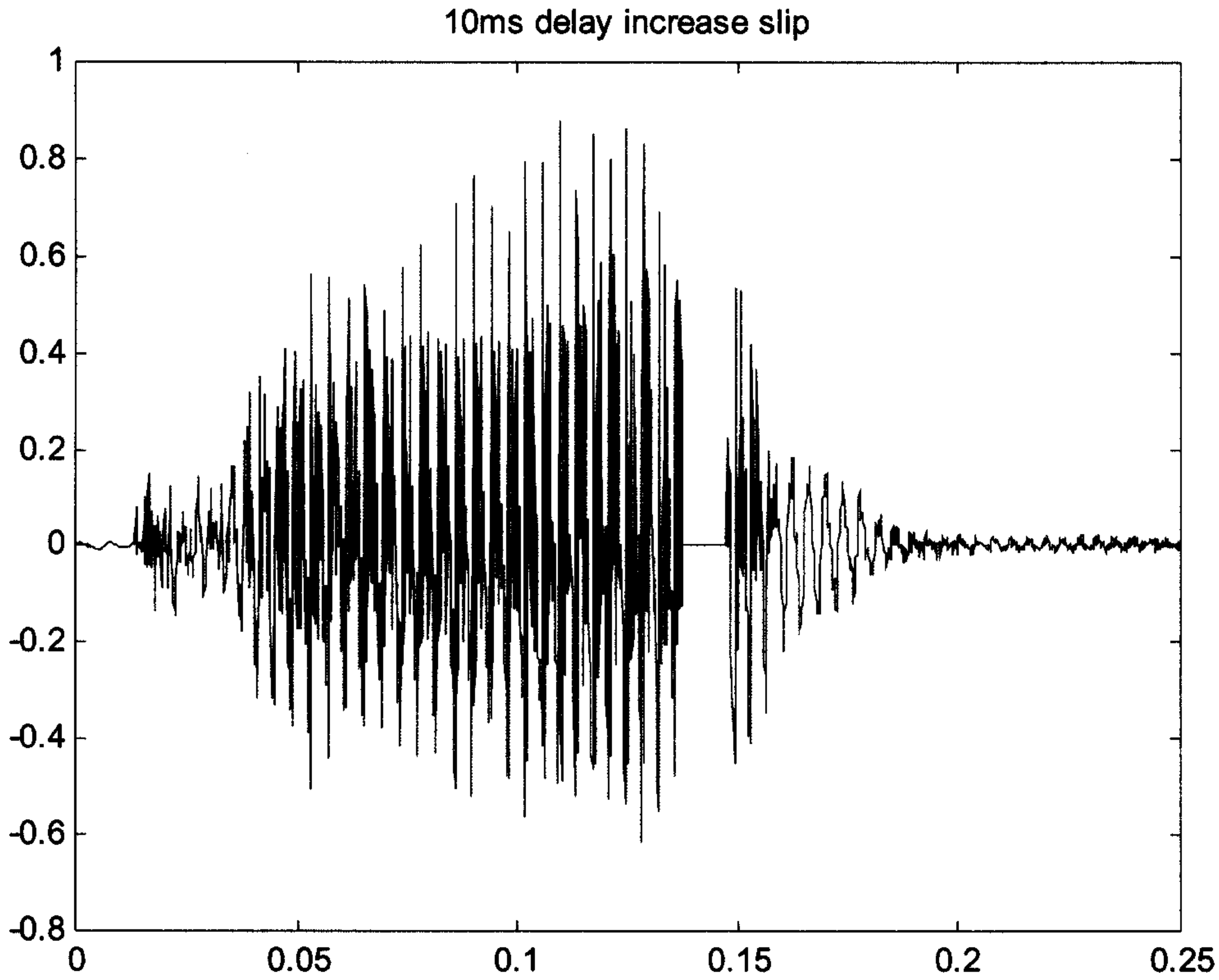
The following figures shows the effect of delay adjustments without and with the time compression / expansion module. All figures show the amplitude evolution over ¼ second of a spoken word. The following figure shows the original signal:



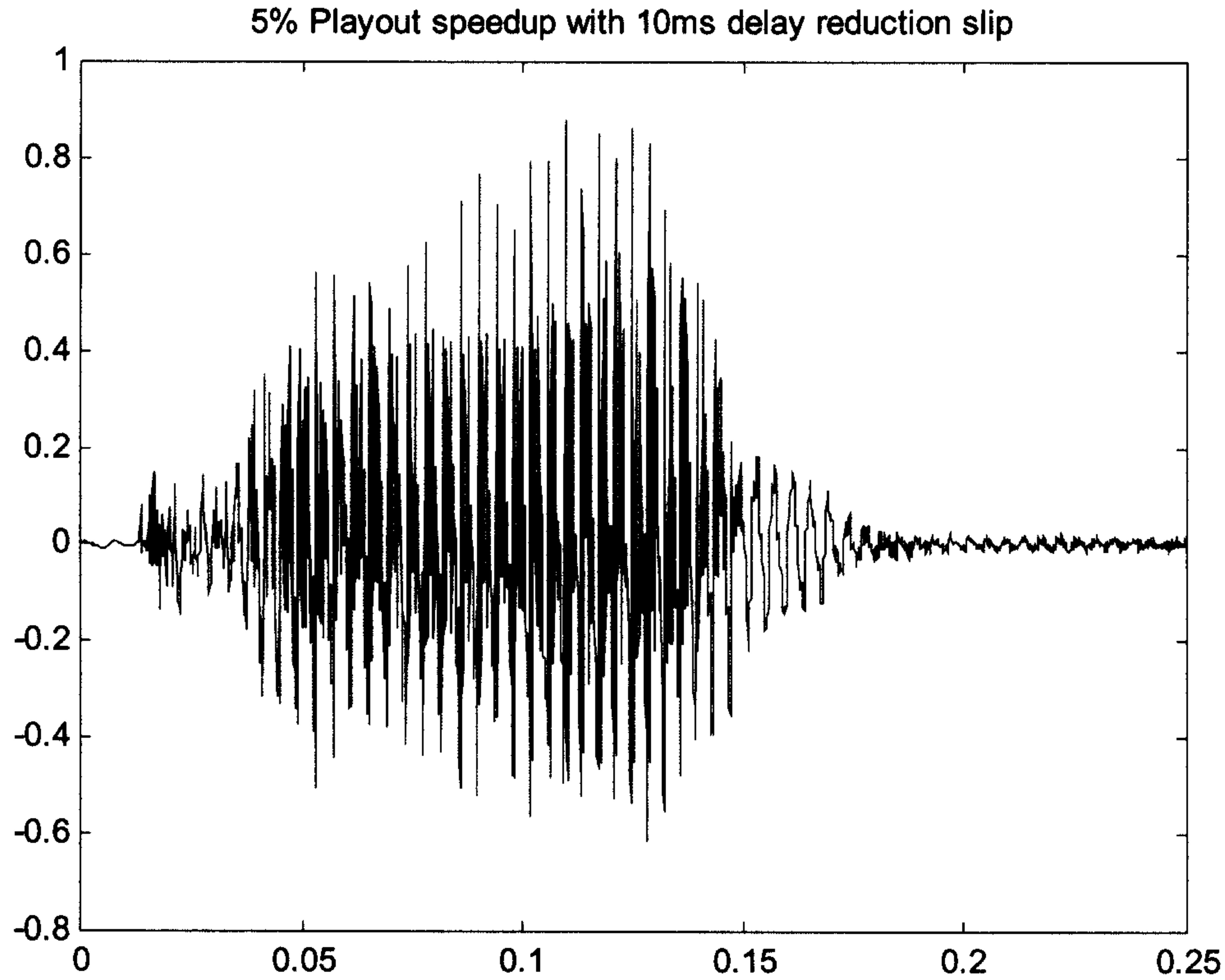
The following figure shows the signal when a 10ms reduction occurs in the delay used by the play-out buffer without the benefit of the time compression technique. The adjustment occurs at around 0.13 s and it can be observed that the waveform suffers from a rapid drop in the decaying overall envelope.



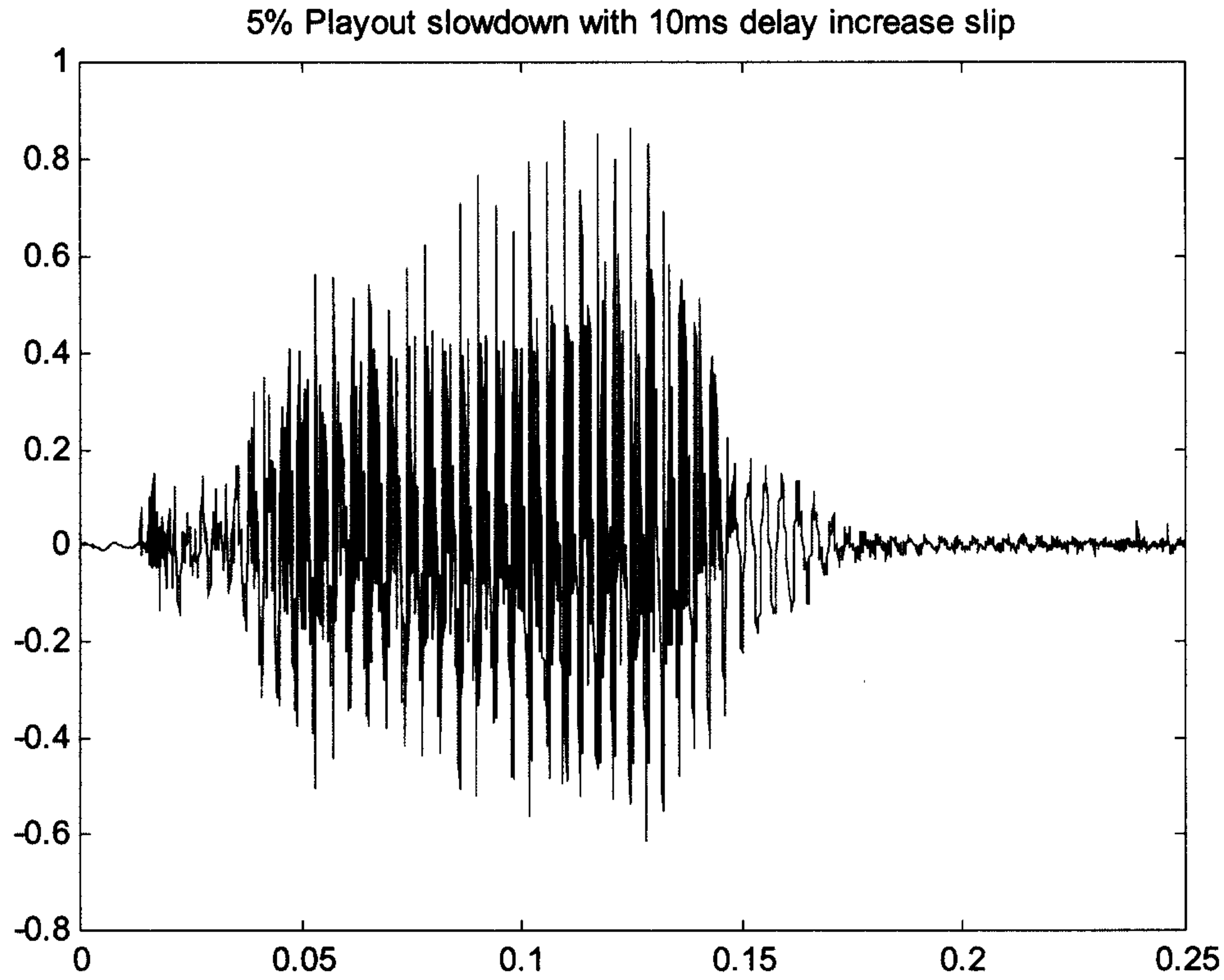
The following figure shows the signal when a 10ms increase occurs in the delay used by the play-out buffer. The adjustment occurs at around 0.13 s and it can be observed that the waveform suffers from the insertion of a gap of silence.



The following figure shows the signal when a 10ms reduction occurs in the delay used by the play-out buffer but using compensation offered by the time expansion / compression module. The delay adjustment occurs at around 0.13 s time and the time scale is compressed by 5% over 200ms duration. There are no visually noticeable artifacts, and listening to the sample reveals no unpleasant discontinuities.



The following figure shows the signal when a 10ms increase occurs in the delay used by the play-out buffer but using compensation offered by the time expansion / compression module. The delay adjustment occurs at around 0.13 s time and the time scale is expanded by 5% over 200ms duration. The silence gap that was observable in the non-compensated case does not show up in this case and there are no visually noticeable artifacts. Listening to this sample reveals no unpleasant side effects.



References:

[1] Danny Cohen, *"Issues in Transnet Packetized Voice Communication"*, Proceedings of the Fifth Data Communications Symposium Snowbird, Utah, September 1977.

[2] F. **Alvarez-Cuevas** et al. *Voice Synchronization in Packet Switching Networks*. IEEE Network, pages 20--25, September 1993

Application number/ Numéro de demande : 2364091

Documents of poor quality scanned
(request original documents in File Prep. Section on the 10th floor)

Documents de piètre qualité numérisés
(Pour obtenir les documents originaux, veuillez vous adresser à la Section de préparation des dossiers, située au 10e étage)