Acta Tecnologia, Volume 12, Issue 2, Pages 107-110, 2026 ISSN 2453-675X

Received: 05 Apr. 2026; Revised: 21 May 2026; Accepted: 25 June 2026
https://doi.org/10.22306/atec.v12i2.333

Applications of VSELP coding in US patents

Muhanned AL-Rawi
Bandung Institute of Technology, Jl. Ganesa No.10, Lb. Siliwangi, Kecamatan Coblong, Kota Bandung, Jawa Barat 40132, Indonesia, muhrawi@yahoo.com

Keywords: US patents, VSELP coding, applications.
Abstract: At data rates as low as 4800 bps, Code Excited Linear Prediction (CELP) speech coders perform well. The primary disadvantage of CELP type coders is their high processing demands. A structured codebook is used by the Vector Sum Excited Linear Prediction (VSELP) speech coding, enabling a highly effective search process. Additional benefits of the VSELP codebook structure will be covered in this paper. VSELP codec(coder/decoder) uses a single VSELP excitation codebook, a novel gain quantizer that is resistant to channel errors, an adaptive pre/postfilter arrangement, and a subsample resolution single tap long term predictor. This paper provides a brief explanation of the uses of VSELP in US patents in addition to its structure.

1 Introduction

Known as Code Excited Linear Prediction (CELP) (also known as Vector Excited or Stochastically Excited), Vector Sum Excited Linear Prediction is a type of CELP speech coding [1]. The VSELP speech codec was created with three objectives in mind: i) maximum speech quality; ii) acceptable computational complexity; and iii) resilience to channel errors. These three objectives are necessary for low data rate (4.8–8 kbps) speech coding for telecommunications applications to be widely accepted.

The VSELP speech codec uses a structured excitation codebook effectively to accomplish these objectives. In addition to lowering computational complexity, the structured codebook improves resilience to channel errors [1]. To achieve high speech quality with reasonable complexity, a single optimized VSELP excitation codebook is utilized. For high-pitched speakers, a subsample resolution single tap long-term predictor significantly enhances performance. The use of a novel gain quantizer results in high coding efficiency and robustness to channel errors. Lastly, the quality of the reconstructed speech is improved by using an adaptive pre/post filter arrangement.

1.1 VSELP codec structure

Figure 1 shows a block diagram of the VSELP speech codec. The 4.8 kbps VSELP codec uses two excitation sources. The first source is the adaptive codebook, sometimes referred to as the long term ("pitch") predictor state [2]. The second is the VSELP excitation codebook. For the 4.8 kbps coder, the VSELP codebook contains 1024 codevectors.

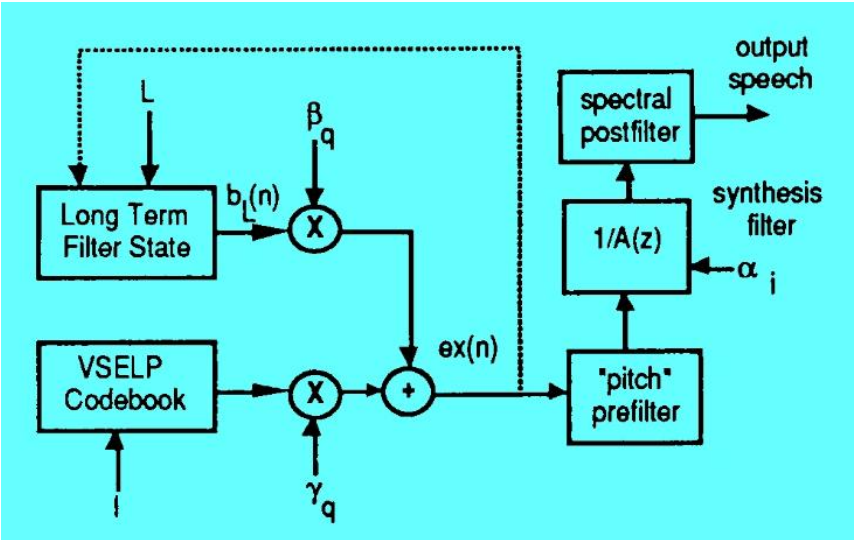

Figure 1 VSELP codec

The excitation vectors selected from the two excitation sources are multiplied by the corresponding gain terms and then added together to form the combined excitation sequence ex(n). After every subframe, the long-term filter state (adaptive codebook) is updated using Ex(n). The synthesis filter has a direct form 10th order linear prediction code (LPC) and is an all-pole filter. The LPC coefficients are interpolated every 7.5 msec subframe after being coded once every 30 msec frame. The excitation parameters are also updated every 7.5 msec subframe. A sub contains 60 samples at a sampling rate of 8 kHz.

2 VSELP Applications in US patents

This section briefly elaborates the applications of VSELP coding in US patents for the last past decade of 21th-century.

The embodiments discussed in [3] pertain to obtaining one or more call quality-indicating measurements from a VoIP communication, altering one or more of the communication's settings, such as the type of codec (e.g., VSELP codec) or the ascribed bit rate of a VBR codec, and using at least one of the one or more call quality measurements to assess the call quality of the VoIP communication following the changes.

The energy in each of several frequency bands of an audio signal's frames is determined by embodiments of an acoustic feedback suppressor using the VSELP code in [4]. A present frame is identified as having feedback characteristics and content that is unlikely to be human voice by comparing the energy in each of the multiple frequency bands to human voice characteristics. The band where feedback is suspected of occurring is subjected to an adaptive gain reduction once it has been established that feedback is taking place.

In [5] a characteristic determining unit that determines whether a sound signal is an audio or speech signal; an encoder (VSELP coding) that converts the sound signal into an encoded signal based on a determination by the characteristic determining unit; a transmitting unit that transmits the encoded signal; a receiving unit that receives the encoded signal; a decoder that decodes the encoded signal; and a packet loss detecting unit that detects a loss of data of the encoded signal and sends a notification to the characteristic determining unit that the data has been lost. The characteristic determining unit instructs the encoder to convert the sound signal portion into a signal portion made up of independently decodable frames after receiving the notification that the data has been lost.,

A technique for regulating an electronic device's average encoding rate is explained in [6]. One step in the process is getting a speech signal. Determining a first average rate is another step in the process. The process also involves using the first average rate to determine a first threshold. By establishing at least one additional threshold based on the first threshold, the technique also controls the average encoding rate. The technique also uses VSELP coding to transmit an encoded speech signal.

A technique for employing adaptive tilt compensation by a speech decoder (using VSELP) is presented in [7]. Receiving a bit stream with multiple parameters representative of a speech signal, identifying an adaptive and a fixed code vector using the plurality of parameters, scaling the adaptive and fixed code vectors to create a scaled adaptive code vector and a scaled fixed code vector, adding the scaled adaptive and fixed code vectors to create a synthesized output, calculating a first reflection coefficient based on the plurality of parameters representative of the speech signal, multiplying the first reflection coefficient by a factor to create a tilt factor, and applying the tilt factor to the synthesized output based on an encoding bit rate are the steps in the method.

The method, device, and non-transitory computer-readable storage medium disclosed in [8] are set up to analyze an input audio signal to determine an estimated pitch period associated with a detection criterion, encode (using VSELP) the filtered audio signal to create a compressed audio bit stream, decode the compressed audio bit stream to create a decoded audio signal, and adaptively filter the decoded audio signal to create an output audio signal. Adaptively filtering the decoded audio signal entails filtering each of the decoded audio signal in a way that is dependent on the estimated pitch period associated with it, and a pitch-based post-filter works to shape noise components situated between harmonics of the input audio signal.

An audio decoder for decoding an encoded audio bitstream is disclosed in one embodiment in [9]. There are at least three distinct decoding modes that can be used with the audio decoder (VSELP or CELP). To extract audio data and control information from the encoded audio bitstream, the audio decoder has a demultiplexer. Additionally, the audio decoder comes with two audio decoders: one set up to function in a first decoding mode using a first decoding technique, and the other set up to function in a second decoding mode using a second decoding technique. Additionally, the second audio decoder has a pitch predictor built into it. Both a short-term and long-term prediction filter are part of the pitch predictor. Additionally, the audio decoder has a selector that allows users to choose from at least three distinct decoding modes using at least some of the control information.

A pitch filter for filtering a preliminary audio signal produced from an audio bitstream is disclosed in certain embodiments in [10]. The pitch filter can be operated in one of two ways: (i) in an active mode, which filters the preliminary audio signal using filtering information to produce a filtered audio signal; or (ii) in an inactive mode, which disables the pitch filter. Based on control information, the pitch filter can be selectively operated in either the active mode or the inactive mode while operating in the coding mode. The preliminary audio signal is generated in an audio encoder/audio decoder (VSELP codec/CELP codec) with a coding mode chosen from at least two different coding modes.

The current disclosure in [11] focuses on a communications system and technique for mitigating tandeming effects. The technique might involve receiving an input bitstream linked to an incoming initial speech signal from a speech encoder (VSELP codec) at a speech decoder (like VSELP codec). The process may also involve figuring out whether coding is necessary and, if so, altering an excitation signal connected to the bitstream. Giving an adaptive encoder the altered excitation signal is another possible step in the process.

A microphone unit in [12] contains a transducer, which converts an acoustic signal into an electrical audio signal; a speech coder, which uses the VSELP codec to extract compressed speech data from the audio signal; and a digital output, which provides digital signals that represent the compressed speech data. A bank of filters with irregularly spaced center frequencies, such as mel frequencies, may be included in the speech coder, which may be a lossy speech coder.

In order to allow the central processing unit to help the digital signal processor create and maintain a communication channel, a communications device in [13] has a magnetic memory that is accessible by both the central processing unit and the digital signal processor. In the event of a power outage or an interruption in the communication channel, the communication device is set up to resume communications.

VSELP was utilized as an application in [14]. The quality of decoded data can be measured at the level of sub-units of a unit of data, such as sub-blocks of a video frame, using the automatic video comparison system for measuring decoded data quality described here. As a result, the system can identify flaws that an automated system that evaluates quality at the frame level might miss otherwise. Because processing encoded media requires a lot of computing power, the automatic video comparison system makes use of a distributed computing system to split up the calculations among numerous parallel-capable computing resources.

VSELP application was used in [15]. A pitch filter is disclosed in certain embodiments for the purpose of filtering an audio bitstream's preliminary audio signal. The pitch filter can be operated in one of two ways: (i) in an active mode, which filters the preliminary audio signal using filtering information to produce a filtered audio signal; or (ii) in an inactive mode, which disables the pitch filter. Based on control information, the pitch filter can be selectively operated in either the active or inactive mode while operating in the coding mode. The preliminary audio signal is generated in an audio encoder or audio decoder with a coding mode chosen from at least two different coding modes.

3 Summary and conclusion

To demonstrate the significance and range of services that VSELP offers, the US patents from the last ten years were displayed. Additionally, the VSELP codes' general structure was illustrated.

References

[1] KEMP, D.P., SUEDA, R.A., TREMAIN, T.E.: *An evaluation of 4800 bps voice coders*, IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 1989, Glasgow, UK, pp. 200-203, 1989. https://doi.org/10.1109/ICASSP.1989.266399

[2] GERSON, I., JASIUK, M.A.: *Vector Sum Excited Linear Prediction (VSELP) speech coding at 8 kbps,* Proceedings of the 1990 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-90), pp. 461-464, 1990. https://doi.org/10.1109/ICASSP.1990.115749

[3] YUAN, N., SUN, Q.Q., COHEN, A., MICHAEL CAMPBELL JUSTIA PATENTS - TEKTRONIX, INC.: *Systems and methods to handle codec changes in call quality calculations*, US 9148351 B2, Tektronix, Inc. Justia Patents - Tektronix, Inc., pp. 1-22, 2015.

[4] GAO, Y., CHEN, L.-M., KUSHNER, W.M., LIU, Y., XIAO, L.: *Method and apparatus for mitigating feedback in a digital radio receiver*, US 9749021 B2, Motorola Solutions, Inc., pp. 1-18, 2015.

[5] ISHIKAWA, T., NORIMATSU, T.: *Encoding and decoding system, decoding apparatus, encoding apparatus, encoding and decoding method*, US 9236053 B2, Panasonic Corporation, pp. 1-36, 2016.

[6] SUBASINGHA, S., ATTI, V.: *Systems and methods for controlling an average encoding rate for speech signal encoding*, US 9263054 B2, Qualcomm Incorporated, pp. 1-31, 2016.

[7] SU, H.-Y., GAO, Y.: *Adaptive tilt compensation for synthesized speech*, US 9401156 B2, Mindspeed Technologies, Inc., pp. 1-26, 2016.
[8] RESCH, B., KJÖRLING, K., VILLEMOES, L.: *Post filter*, US 9552824 B2, Dolby International AB, pp. 1-24, 2017.
[9] RESCH, B., KJÖRLING, K., VILLEMOES, L.: *Audio encoder and decoder with pitch prediction*, US 9558754 B2, Dolby International AB, pp. 1-44, 2017.
[10] RESCH, B., KJÖRLING, K., VILLEMOES, L.: *Selective bass post filter*, US 9830923 B2, Dolby International AB, pp. 1-25, 2017.
[11] TANG, Q.-Y., GAO, Y.: *System and method for reducing tandeming effects in a communication system*, US 9953660 B2, Huawei Technologies Co., Ltd., pp. 1-25, 2018.
[12] LESSO, J.P.: *Microphone unit comprising integrated speech analysis*, US 10297258 B2, Cirrus Logic International Semiconductor Ltd., pp. 1-32, 2019.
[13] ASGHAR, S.: *Multiplexed memory in a communication processing system*, US 10649942 B2, Everspin Technologies, Inc., pp. 1-28, 2020.
[14] BABBAR, G., BRICE, P., HILLEGASS, D., MALYSHEVA, O.: *Automatic video comparison of the output of a video decoder*, US 11064204 B2, Arris Enterprises LLC, pp. 1-20, 2021.
[15] RESCH, B., KJÖRLING, K., VILLEMOES, L.: *Post filter for audio signals*, US 11996111 B2, Dolby International AB, pp. 1-32, 2024.

Review process
Single-blind peer review process.