

FOUR BITS PER SAMPLE SPEECH CODER BASED ON ADPCM

Samardzija D., Domazetovic A., Lukac Z.,

Department of Electrical Engineering, Faculty of Technical Sciences Novi Sad, Yugoslavia

I INTRODUCTION

A great number of speech and audio digital coding algorithms is available today. Basically, they are optimized for different applications and implementation platforms (digital signal processor (DSP) platforms). Usually, a specific application and DSP platform require additional optimization of the chosen algorithm or even the development of a new one. Optimization of the algorithms effects the following characteristics [1]:

- Quality of decoded speech and/or audio signals
- Compression rate
- Complexity— processor power and memory requirements
- Delay

This paper describes development, features and real-time implementation of the speech-coding algorithm that is a result of the specific requirements set by application targets and DSP platform.

- Excellent-good and good quality of decoded speech and audio signals(music) respectively, corresponding to relative speech coder quality
- Bit rate must not exceed 4 bits per sample.
- Low complexity that allows coding and decoding of two independent channels with sampling rate up to 16kHz (preferably 32kHz)
- It has to be implemented on the DSP platform: MAS 3507D is the platform that introduces the following limitations[2]:
 - Processing power is 24 MIPS
 - 4000 words are the total memory space for dynamic data, coefficients and program code
 - Fractional arithmetic
- Synchronization system that supports fast forward and rewind options for the decoder. It also has to be supported by synchronization words inserted into the contest of the coded signal

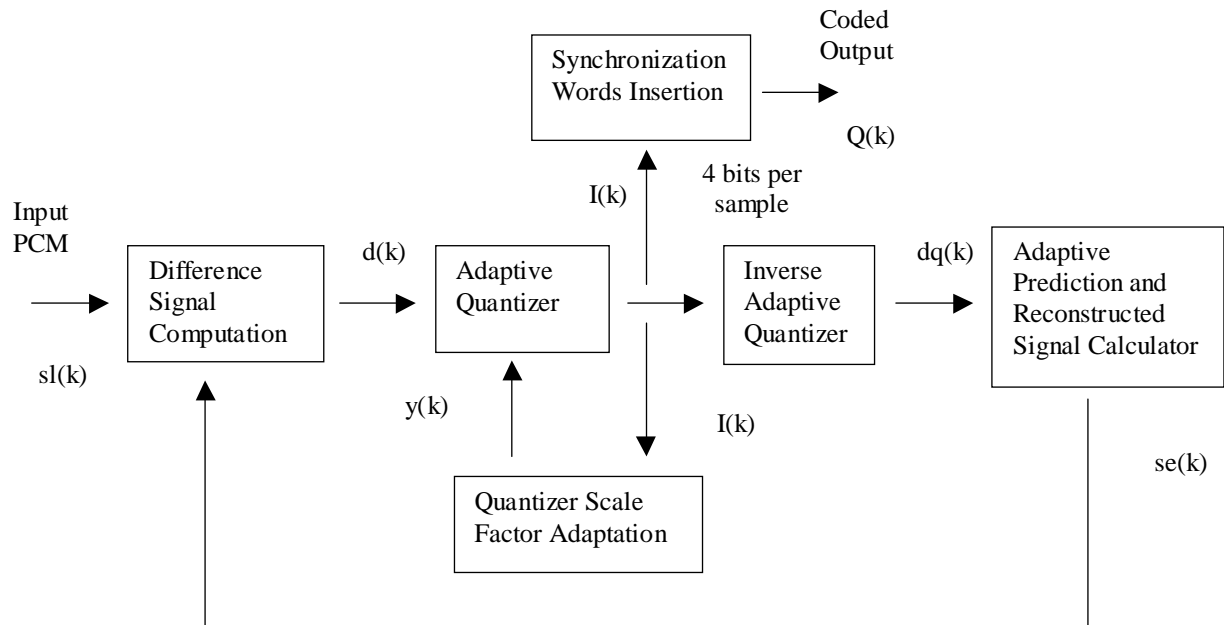


Figure 1. Block diagram of the encoder

II REQUIREMENTS AND SELECTION OF THE INITIAL ALGORITHM

Development of this codec has been initiated by following the need for digital dictating machine (e.g. digital speech recorder and player) implemented on a DSP platform. Here is the list of the requirements that have been imposed:

As an initial part of the design process, a number of well known speech coding algorithms has been considered in order to find the most suitable for the application. Here is the list of the algorithms that have been examined:

1. CELP 4.8kbps FS 1016 [3]
2. CELP 16kbps ITU-T Rec. G.728 [4]
3. CELP 5.3kbps and 6.3kbps ITU-T Rec. G.723.1 [5]
4. ADPCM 32kbps ITU-T Rec. G.726 [6]

The ADPCM 32kbps ITU-T Rec. G.726 has been recognized as an initial algorithm because it is the closest to match the requirements. But it could not be implemented straightforwardly. A number of modifications had to be introduced. These modifications have contributed to definition of the new speech-coding algorithm that has met the requirements.

This new algorithm is not compatible to the ITU-T Rec. G.726

(2) into the non-linear quantization area where the resolution is higher. This adaptation routine implies the fast adaptation due to large fluctuations of the difference signals. The scale factor is:

$$y(k) = (1 - 2^{-5})y(k-1) + 2^{-5}W[I(k)] \quad (3)$$

$W[]$ is a table containing 8 entries. It is taken from the ITU-T Rec. G.726[6].

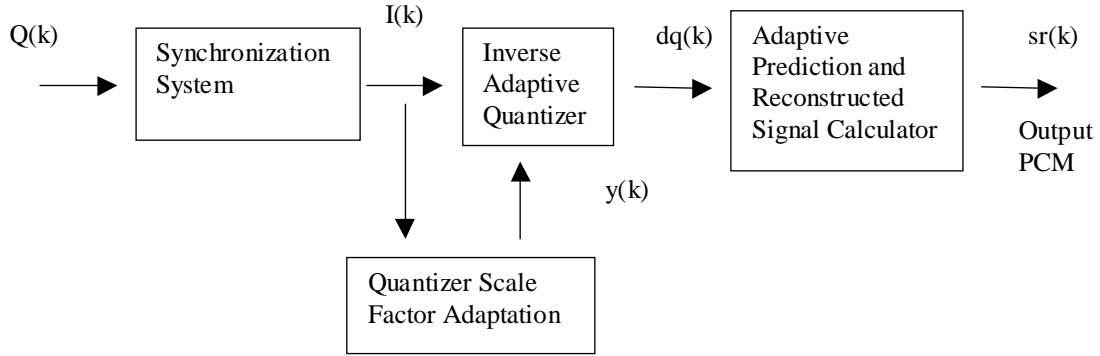


Figure 2. Block diagram of the decoder

III BUILDING BLOCKS

This Chapter presents the building blocks of the encoder and the decoder. The encoder is depicted in Figure 1 and the decoder in Figure 2. The encoder contains the internal decoder equivalent to the remote one.

DIFFERENCE SIGNAL COMPUTATION

It computes the difference signal $d(k)$ from the PCM input signal $sl(k)$ and the signal estimate $se(k)$:

$$d(k) = sl(k) - se(k) \quad (1)$$

ADAPTIVE QUANTIZER

Non-uniform adaptive quantizer is used to quantize the difference signal $d(k)$. The value:

$$\log_2|d(k)| - y \quad (2)$$

is mapped into the non-linear quantization table containing 8 entries (three bits). The sign of $d(k)$ is coded as an additional bit. Finally, the adaptive quantizer results in four bits per sample code word $I(k)$ that is conveyed into the internal decoder and Synchronization Words Insertion routine.

This block is implemented following the bit-exact definition of the quantizer recommended by the ITU-T Rec. G.726[6].

INVERSE ADAPTIVE QUANTIZER

Quantized difference signal $dq(k)$ is obtained by executing the inverse operations corresponding to (2). This block is implemented in the bit-exact manner recommended by the ITU-T Rec. G.726[6].

QUANTIZER SCALE FACTOR ADAPTATION

This block is responsible for the adaptive prediction of the difference signal in logarithmic domain. It results in $y(k)$ that attempts to minimize the value (2). It tries to bring value

Comparing to equivalent routine used in the ITU-T Rec. G.726, this routine is simplified by exclusion of the adaptation to slow signals (e.g. voice-band data, tones...). This modification simplifies the routine, but retains the quality of the decoded speech.

ADAPTIVE PREDICION AND RECONSTRUCTED SIGNAL CALCULATION

This is the second adaptive procedure employed in the algorithm. Unlike the previously described procedure, this one predicts the value of the incoming signal sample in linear domain. The predicted value $se(k)$ attempts to minimize the difference $d(k)$ (1). $se(k)$ is calculated as linear combination of previously decoded sample and quantized differences:

$$se(k) = sr(k-1) + \sum_{i=1}^6 bi(k-1) dq(k-i) \quad (4)$$

where $sr(k)$ is the reconstructed signal:

$$sr(k) = se(k) + dq(k) \quad (5)$$

$bi(k)$ ($i=1..6$) are the coefficients that are updated according to well known LSM algorithm:

$$bi(k) = (1 - 2^{-8}) bi(k-1) + 2^{-7} \text{sgn}(dq(k)) \text{sgn}(dq(k-i)) \quad (6)$$

This routine is highly optimized for the arithmetic used in the DSP platform. Fractional arithmetic is applied. Unlike the ITU-T Rec. G.726 this solution does not require bit-exact implementation. Benefits of this approach are evident in the case of convergence properties of the decoder (e.g. if the initial states of the encoder and decoder are not matched this algorithm will converge, and no perceivable degradation will be noticed due to the convergence being very fast). This

characteristic of the algorithm is crucial for the implementation of fast-forward and rewind options of the decoder.

SYNCHRONIZATION SYSTEM

Self-synchronizing decoder has to accomplish the task of initial synchronization, and later, when synchronization is established, it has to maintain it. The synchronization is based on the synchronization information encoded into the bit-stream and created by the encoder. The encoder periodically inserts words:

- Start Synchronization word that contains information about sampling rate and the number of encoded channels
- Tracking Synchronization word that enables the decoder to maintain the synchronization and to detect the synchronization loss

The decoder analyzes the bit-stream, searching for above words and changes its state corresponding to the received information. If the synchronization is established, and analysis of the bit-stream shows the missing of the Tracking or Start Synchronization words, the synchronization loss is detected, and the decoder returns to initial state.

IV QUALITY OF DECODED SPEECH AND AUDIO SIGNALS

The quality of decoded speech and audio signals has been determined by using perceptual comparison technique. A group of five trained listeners compared the quality of the decoded signals using six different speech-coding algorithms including this one (named SC4). The referent relative speech coder quality scale was taken from [7]. Table 1 presents the results of these tests:

Coder Name	Speech	Music
CELP 4.8kbps FS 1016	3.1	1
CELP 5.3kbps G.723.1	3.6	1.5
CELP 6.3kbps G.723.1	4.0	1.7
CELP 16kbps G.728	4.2	4.0
ADPCM 32kbps G.726	4.4	4.0
ADPCM 32kbps SC4	4.4	4.2

Unacceptable =1, Poor, Acceptable, Good, Excellent =5
Table 1. Quality of decoded speech and audio signals

V DSP REAL-TIME IMPLEMENTATION AND APPLICATION

The real-time implementation of the encoder and the decoder on the DSP MAS 3507D platform was straightforward. Using the DSP resources that are presented in Chapter 2, this implementation offers the following sampling rates and the number of channels that can be coded and decoded:

Sampling rates	Number of Channels
8 kHz	2 (stereo effect)
11.025 kHz	2 (stereo effect)
16 kHz	2 (stereo effect)
32 kHz	1

Table 2. Sampling rates and number of channels

This implementation is applied as digital speech and music recorder and player. The application platform is depicted in Figure 3. During the recording session, the DSP encodes the input signal and creates the bit-stream that is stored in Flash RAM. During the replay session, the bit-stream is conveyed from the Flash RAM via the controller to the DSP where decoding is performed. If required by user, during the replay, fast-forward or rewind options can be set.

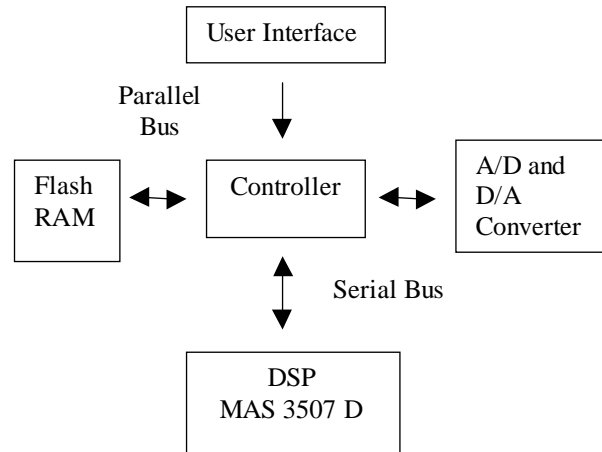


Figure 3. Application platform

VI CONCLUSION

The quality of the coder was determined by the use of the specially designed perceptual testing. It showed that the coder satisfied the speech quality requirements yet well exceeded expected audio quality.

Due to the new solutions applied on the initial algorithm, complexity of this new algorithm is four times lower than that of the initial one referring to the DSP platform.

The most critical procedures are redesigned to accommodate fractional arithmetic of the processor.

The synchronization system introduces the fast forward and rewind options as well as self-synchronizing decoder that is able to determine the sampling rate and the number of encoded channels only by analyzing the bit-stream.

As a result of all these new solutions, the initial algorithm has been changed significantly. Supported by this new algorithm, the DSP MAS 3507D platform is fully exploited, meeting successfully the application requirements.

REFERENCES

- [1] Jayant N., *Signal Compression: Technology Targets and Research Directions*, IEEE Journal on Selected Areas in Communications, Jun 1992.
- [2] *MASC 3500 Preliminary Specification*, Intermetall, Frieberg, Germany, 1995.
- [3] Domazetovic A., Samardzija D., Kovacevic J., *Implementation of 4800 bps CELP on MASC 3500 DSP*, Conference on Telecommunications ETRAN, Vrnjacka Banja, Yugoslavia, Jun 1998.
- [4] Temerinac M., Samardzija D., Lukac Z., *Real-Time Implementation of Low Delay CELP (ITU-T Rec. G.728) on ITT DSP MAS 3503 C*, Conference on

Telecommunications ETRAN, Zlatibor, Yugoslavia, Jun 1997

[5] Lukac Z., Samardzija D., *A Solution of CELP Encoder at 5.3kbit/s and 6.3kbit/s on ITT MASC 3500 Processor*, Conference on Telecommunications ETRAN, Vrnjacka Banja, Yugoslavia, Jun 1998.

[6] *ITU-T Recommendation G.726*

[7] Cox R.V, *Three New Speech Coders from ITU Cover a Range of Applications*, IEEE Communications Magazine, Jun 1998.

Summary: This paper describes development, features and real-time implementation of the speech-coding algorithm. Basic ideas and implementation of the each building block are presented. Analysis is presented following the modifications introduced upon the initial algorithm (ITU-T Rec. G.726-ADPCM). Results of perceptual quality testing and real-time implementation are also shown. Validity of this algorithm is confirmed by the final application in the digital recording and replaying machine.

**FOUR BITS PER SAMPLE SPEECH CODER BASED
ON ADPCM,**

Samardzija D., Domazetovic A., Lukac Z.