



INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

G.711

Appendix II
(02/2000)

SERIES G: TRANSMISSION SYSTEMS AND MEDIA,
DIGITAL SYSTEMS AND NETWORKS

Digital transmission systems – Terminal equipments –
Coding of analogue signals by pulse code modulation

Pulse code modulation (PCM) of voice frequencies

**Appendix II: A comfort noise payload definition
for ITU-T G.711 use in packet-based multimedia
communication systems**

ITU-T Recommendation G.711 – Appendix II

(Formerly CCITT Recommendation)

ITU-T G-SERIES RECOMMENDATIONS
TRANSMISSION SYSTEMS AND MEDIA, DIGITAL SYSTEMS AND NETWORKS

| | |
|--|--------------------|
| INTERNATIONAL TELEPHONE CONNECTIONS AND CIRCUITS | G.100–G.199 |
| INTERNATIONAL ANALOGUE CARRIER SYSTEM | |
| GENERAL CHARACTERISTICS COMMON TO ALL ANALOGUE CARRIER-TRANSMISSION SYSTEMS | G.200–G.299 |
| INDIVIDUAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON METALLIC LINES | G.300–G.399 |
| GENERAL CHARACTERISTICS OF INTERNATIONAL CARRIER TELEPHONE SYSTEMS ON RADIO-RELAY OR SATELLITE LINKS AND INTERCONNECTION WITH METALLIC LINES | G.400–G.449 |
| COORDINATION OF RADIOTELEPHONY AND LINE TELEPHONY | G.450–G.499 |
| TESTING EQUIPMENTS | |
| TRANSMISSION MEDIA CHARACTERISTICS | G.600–G.699 |
| DIGITAL TRANSMISSION SYSTEMS | |
| TERMINAL EQUIPMENTS | G.700–G.799 |
| General | G.700–G.709 |
| Coding of analogue signals by pulse code modulation | G.710–G.719 |
| Coding of analogue signals by methods other than PCM | G.720–G.729 |
| Principal characteristics of primary multiplex equipment | G.730–G.739 |
| Principal characteristics of second order multiplex equipment | G.740–G.749 |
| Principal characteristics of higher order multiplex equipment | G.750–G.759 |
| Principal characteristics of transcoder and digital multiplication equipment | G.760–G.769 |
| Operations, administration and maintenance features of transmission equipment | G.770–G.779 |
| Principal characteristics of multiplexing equipment for the synchronous digital hierarchy | G.780–G.789 |
| Other terminal equipment | G.790–G.799 |
| DIGITAL NETWORKS | G.800–G.899 |
| DIGITAL SECTIONS AND DIGITAL LINE SYSTEM | G.900–G.999 |

For further details, please refer to ITU-T List of Recommendations.

ITU-T RECOMMENDATION G.711

PULSE CODE MODULATION (PCM) OF VOICE FREQUENCIES

APPENDIX II

A comfort noise payload definition for ITU-T G.711 use in packet-based multimedia communication systems

Summary

This appendix defines a comfort noise payload format (or bit-stream) for ITU-T G.711 use in packet-based multimedia communication systems.

The use of the payload format is intended for packet-based systems with a large header overhead where the packet transmission rate plays a significant role in the overall system bit-rate. In this situation, the use of VAD/DTX/CNG can significantly reduce the packet transmission rate and hence improve the bandwidth efficiency.

Source

Appendix II to ITU-T Recommendation G.711 was prepared by ITU-T Study Group 16 (1997-2000) and was approved under the WTSC Resolution No. 1 procedure on 28 February 2000.

FOREWORD

ITU (International Telecommunication Union) is the United Nations Specialized Agency in the field of telecommunications. The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of the ITU. The ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Conference (WTSC), which meets every four years, establishes the topics for study by the ITU-T Study Groups which, in their turn, produce Recommendations on these topics.

The approval of Recommendations by the Members of the ITU-T is covered by the procedure laid down in WTSC Resolution No. 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

INTELLECTUAL PROPERTY RIGHTS

The ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. The ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, the ITU had not received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementors are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database.

© ITU 2000

All rights reserved. No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the ITU.

CONTENTS

| | Page |
|--|-------------|
| Appendix II – A comfort noise payload definition for ITU-T G.711 use in packet-based multimedia communication systems..... | 1 |
| II.1 Scope..... | 1 |
| II.2 Comfort noise payload definition | 1 |
| II.2.1 Noise level | 1 |
| II.2.2 Reflection coefficients..... | 2 |
| II.2.3 Payload packing..... | 2 |
| II.3 Guidelines for use | 2 |
| II.3.1 Factors affecting system performance | 3 |
| II.3.2 Illustration of bandwidth savings in packet-based network applications | 4 |
| II.4 Performance results..... | 4 |
| II.5 Example solution | 7 |
| II.5.1 Algorithm description..... | 7 |
| II.5.2 Tested configuration | 10 |

Recommendation G.711

PULSE CODE MODULATION (PCM) OF VOICE FREQUENCIES

APPENDIX II

A comfort noise payload definition for ITU-T G.711 use in packet-based multimedia communication systems

(Geneva, 2000)

II.1 Scope

This appendix defines a comfort noise payload format (or bit-stream) for ITU-T G.711 use in packet-based multimedia communication systems. The payload format is generic and may also be used with other speech codecs without built-in Discontinuous Transmission (DTX) capability such as ITU-T Recommendations G.726 [1], G.727 [2], G.728 [3], and G.722 [4]. The payload format provides a minimum interoperability specification for communication of comfort noise parameters. The comfort noise analysis and synthesis as well as the Voice Activity Detection (VAD) and DTX algorithms are unspecified and left implementation-specific. However, an example solution has been tested and is described. It uses the VAD and DTX of G.729 Annex B [5] and a comfort noise generation algorithm (CNG) which is provided for information.

The use of the payload format is intended for packet-based systems with a large header overhead where the packet transmission rate plays a significant role in the overall system bit-rate. In this situation, the use of VAD/DTX/CNG can significantly reduce the packet transmission rate and hence improve the bandwidth efficiency.

II.2 Comfort noise payload definition

The comfort noise payload consists of a description of the noise level and spectral information in the form of reflection coefficients. The use of spectral information is optional and the all-pole model order is left unspecified. The encoder can determine the appropriate model order based on such considerations as quality, complexity, expected environmental noise and signal bandwidth. The model order is not explicitly transmitted since it can be derived from the length of the payload at the receiver. For complexity or other reasons, the decoder may reduce the model order by setting higher order reflection coefficients to zero.

II.2.1 Noise level

The noise level is expressed in $-dBov$, with values from 0 to 127 representing 0 to -127 $dBov$. $dBov$ is the level relative to the overload of the system. The noise level is packed with the Most Significant Bit (MSB) first with the unused bit always set to 0 according to Figure II.1.

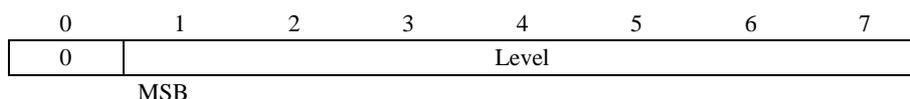


Figure II.1/G.711 – Noise level bit packing

II.2.2 Reflection coefficients

The spectral information is transmitted using reflection coefficients [6]. From the polynomial:

$$A(z) = 1 - \sum_{j=1}^M \alpha_j z^{-j}$$

obtained by linear prediction analysis, the set of reflection coefficients may be obtained from the set of LPC coefficients using a backward recursion of the form:

$$k_i = -a_i^{(i)}$$

$$a_j^{(i-1)} = \frac{a_j^{(i)} + a_i^{(i)} a_{i-j}^{(i)}}{1 - k_1^2} \quad 1 \leq j \leq i-1$$

where i goes from M , to $M-1$, down to 1 with the initial condition:

$$a_j^{(M)} = \alpha_j \quad 1 \leq j \leq M$$

Note that the above formulation results in the solution to k_1 given by:

$$k_1 = -\frac{r_1}{r_0}$$

where r_i is the i th autocorrelation coefficient of the input signal.

Each reflection coefficient can have values between -1 and 1 and is quantized uniformly using 8 bits. The quantized value is represented by the 8 bit index N , where $N = 0, \dots, 254$, and index $N = 255$ is reserved for future use. Each index N is packed into a separate byte with the MSB first. The quantized value of each reflection coefficient can be obtained from its corresponding index using:

$$\hat{k}_i(N_i) = \frac{258}{32768} \cdot (N_i - 127) \quad \text{for } N_i = 0, \dots, 254; -1 < \hat{k}_i(N_i) < 1$$

II.2.3 Payload packing

The first byte of the payload must contain the noise level as shown in Figure II.1. Quantized reflection coefficients are packed in subsequent bytes in ascending order as in Figure II.2 where M is the model order.

| | | | | | |
|------|-------|-------|-------|-----|-------|
| Byte | 1 | 2 | 3 | ... | M+1 |
| | Level | N_1 | N_2 | ... | N_M |

Figure II.2/G.711 – CN payload packing format

The total length of the payload is $M + 1$ bytes. Note that a 0th order model (i.e. no spectral envelope information) reduces to transmitting only the energy level.

II.3 Guidelines for use

A block diagram of a speech communication system with VAD/DTX/CNG capabilities is shown in Figure II.3. The job of the VAD is to discriminate between active and inactive voice segments in the input signal. During inactive voice segments, the role of the CNG is to sufficiently describe the ambient noise while minimising the transmission rate. A Silence Insertion Descriptor (SID) frame containing a description of the noise is packed into the CN payload and sent to the receiver. The

DTX algorithm determines when a SID frame is transmitted. The SID frame may be sent periodically or only when there is a significant change in the background noise characteristics. The CNG algorithm at the receiver uses the information in the SID to update its noise generation model and then produce an appropriate amount of comfort noise.

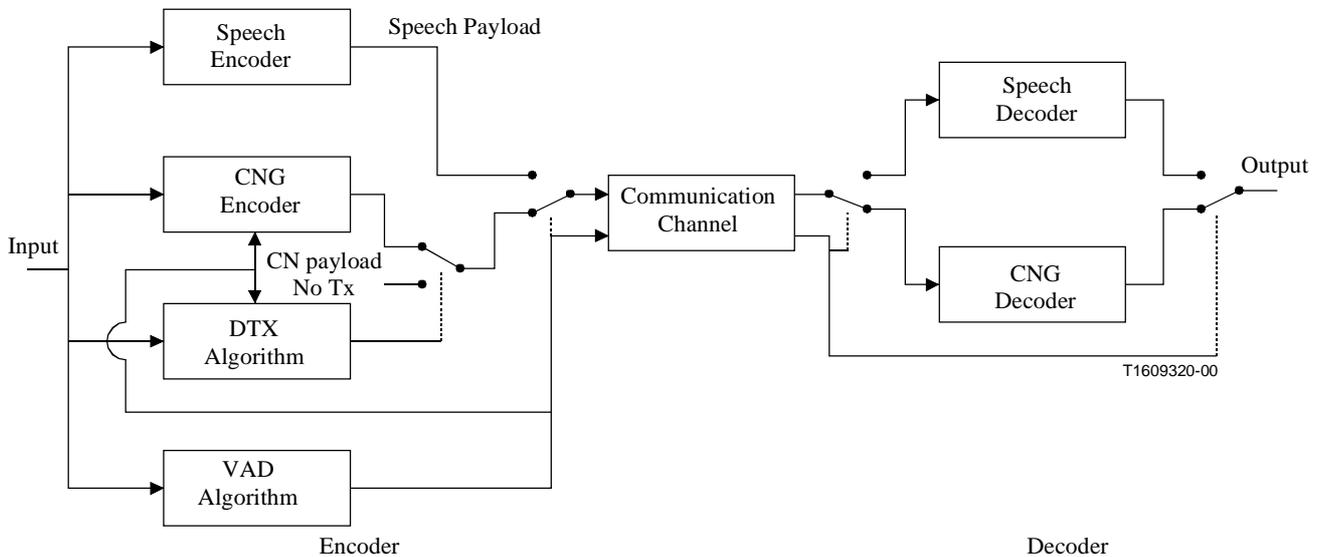


Figure II.3/G.711 – Speech communication system with DTX

II.3.1 Factors affecting system performance

The purpose of the VAD/DTX/CNG components is to reduce the transmission rate during inactive speech periods while maintaining an acceptable level of output quality. Both the quality and efficiency are affected by the performance of each of the components. Care must be taken to jointly consider the characteristics of the VAD, DTX, and CNG algorithms. Otherwise the resulting system may achieve poor performance.

II.3.1.1 VAD

The role of the VAD algorithm is to classify the input signal into active speech and inactive speech or background noise. Misclassifying inactive speech as active speech has an adverse affect on system efficiency by unnecessarily increasing the transmission rate. In this case, the speech quality is unaffected. However, when active speech is misclassified as inactive, the speech signal is clipped and the speech quality degrades. Most DTX algorithms employ a hangover period when transitioning from active to inactive speech in order to avoid clipping the tail end of speech. During the hangover period, inactive speech frames are reclassified as active speech. The hangover period is also important in order for the CNG encoder to obtain an accurate estimate of the ambient noise.

II.3.1.2 DTX

The DTX algorithm determines the frequency of SID frame transmission during periods of inactive speech. Simple DTX schemes update periodically (e.g. 5 Hz to 30 Hz). More complex DTX algorithms analyse the input signal and transmit only when a significant change in ambient noise character is detected [5].

II.3.1.3 CNG

The role of the CNG is to describe and reproduce the ambient noise. The noise may be adequately described by its energy and spectral content. In order to avoid abrupt changes in comfort noise character, it is important to average the parameter estimation over a period of time. The appropriate

amount of averaging is dependent on the ambient noise, the performance and hangover of the VAD, as well as the update rate of the DTX.

The model order used is a factor in the accuracy of the spectral estimation. The optimal order is dependent upon the ambient noise present and the signal bandwidth. It is also important to match the spectral character of the noise produced by the CNG with that of the speech codec. Accordingly, it is suggested that any pre-processing of the input signal before analysis within the speech encoder also be done within the comfort noise encoder.

II.3.2 Illustration of bandwidth savings in packet-based network applications

Table II.1 illustrates how the use of discontinuous transmission in a packet-based communication system can significantly reduce the transmission rate and hence improve the bandwidth efficiency. The example assumes a 40-byte packet overhead, 60% speech activity, and a DTX update rate of 10 Hz.

Table II.1/G.711 – Bandwidth Savings

| Codec | Bit rate (bit/s) | Packet size (ms) | IP bit rate (bit/s) | 1 byte CN payload | | 11 byte CN payload | |
|-------|------------------|------------------|---------------------|--------------------------|-------------|--------------------------|-------------|
| | | | | IP bit rate (Ave. bit/s) | Savings (%) | IP bit rate (Ave. bit/s) | Savings (%) |
| G.711 | 64 000 | 5 ms | 128 000 | 78 112 | 39.0 | 78 432 | 38.7 |
| G.711 | 64 000 | 10 ms | 96 000 | 58 912 | 38.6 | 59 232 | 38.3 |
| G.711 | 64 000 | 20 ms | 80 000 | 49 312 | 38.4 | 49 632 | 38.0 |
| G.726 | 32 000 | 5 ms | 96 000 | 58 912 | 38.6 | 59 232 | 38.3 |
| G.726 | 32 000 | 10 ms | 64 000 | 39 712 | 38.0 | 40 032 | 37.5 |
| G.726 | 32 000 | 20 ms | 48 000 | 30 112 | 37.3 | 30 432 | 36.6 |
| G.728 | 16 000 | 5 ms | 80 000 | 49 312 | 38.4 | 49 632 | 38.0 |
| G.728 | 16 000 | 10 ms | 48 000 | 30 112 | 37.3 | 30 432 | 36.6 |
| G.728 | 16 000 | 20 ms | 32 000 | 20 512 | 35.9 | 20 832 | 34.9 |

E.g. Assuming an RTP/UDP/IP header of 40 bytes, 60% speech activity, and a DTX update rate of 10 Hz, the average IP Bit Rate with G.711 and an 11-byte CN payload is given by: $((64\ 000\ \text{bit/s}) + (40\ \text{bytes} \times 8\ \text{bit/byte} \times (1.0/0.005\ \text{s}))) \times (0.6) + ((40+11)\ \text{bytes} \times 8\ \text{bit/byte} \times 10/\text{s}) \times (0.4) = 78\ 432\ \text{bit/s}$.

II.4 Performance results

A subjective evaluation of an example CNG implementation employing the CN payload was performed. The method of assessment used the Absolute Category Rating (ACR) method as defined in ITU-T Recommendation P.800. The speech material used in the experiment consisted of simple, meaningful, short sentences in North-American English. The source material was Modified IRS filtered (ITU-T Recommendation P.830 Annex D) and arranged in pairs. Each sentence-pair lasted approximately 7 to 8 seconds, with a time interval between sentences of approximately 1 second. The evaluation contained both clean and noisy input conditions, including babble, street, office and car noise.

The speech codec used in the experiment was G.711, processed according to the procedure in Figure II.4. In this experiment, the implementation consisted only of the CNG algorithm. The G.729 Annex B VAD and DTX algorithms were used [5]. Trace files containing the VAD and DTX decisions were obtained according to the procedure in Figure II.5 with the "SYNC" flag enabled in order to align the output with the input file.

The G.711 with comfort noise was obtained using the procedure in Figure II.6. The Source File was down-sampled and level adjusted by a gain G and then encoded by the combination of G.711 and the CNG. The input data was buffered into 10 ms frames. The frame encoding of the CNG algorithm was aligned to the beginning of the speech file in order to "synchronise" with the framing corresponding to the VAD and DTX trace files. On a 10 ms frame basis, the VAD and DTX trace files were used to control the operation of the CNG algorithm. For active speech frames, G.711 was used to process the input data frame. For inactive frames, the CNG algorithm was employed. The DTX flag controlled the update of the CNG parameters. At the decoder, the VAD flag was used to indicate if the current frame is active speech or inactive. A complementary gain $1/G$ (to produce a constant listening level) was then applied and the result was up-sampled and stored as a "processed file".

The results of this noisy ACR experiment showed that, for all cases of interest, G.711 with the test CNG algorithm performs equivalently to G.711 without VAD/CNG. This includes the clean background case as well as the noisy background cases (babble, car, office and street noise).

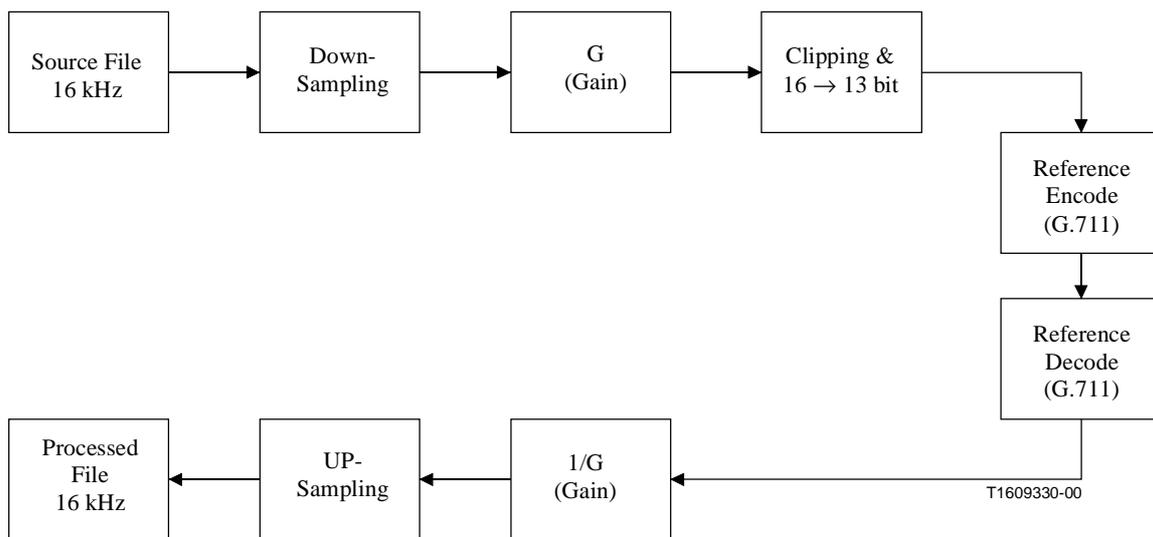
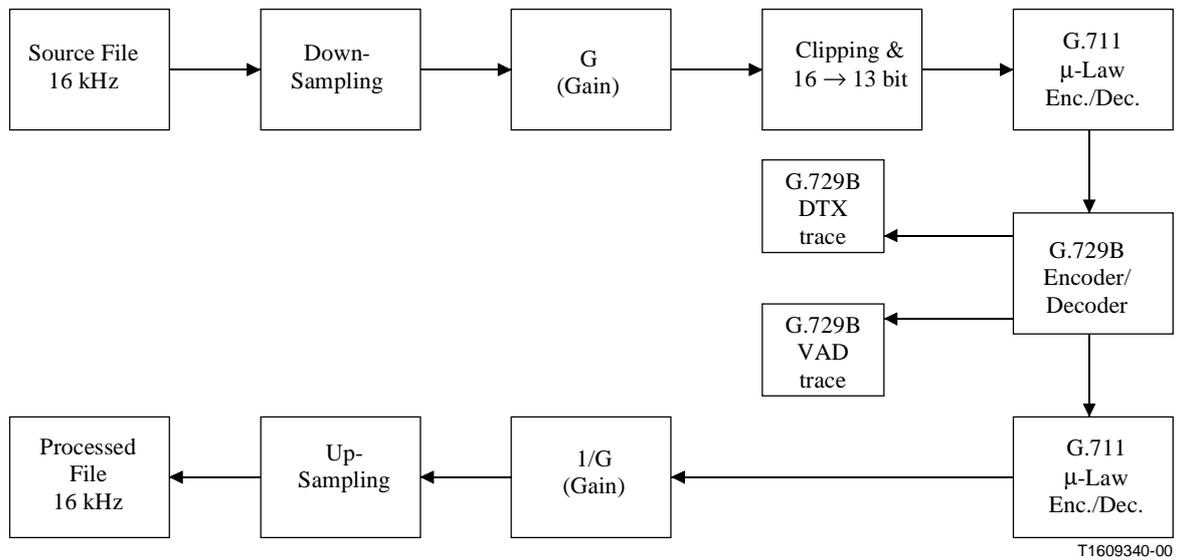
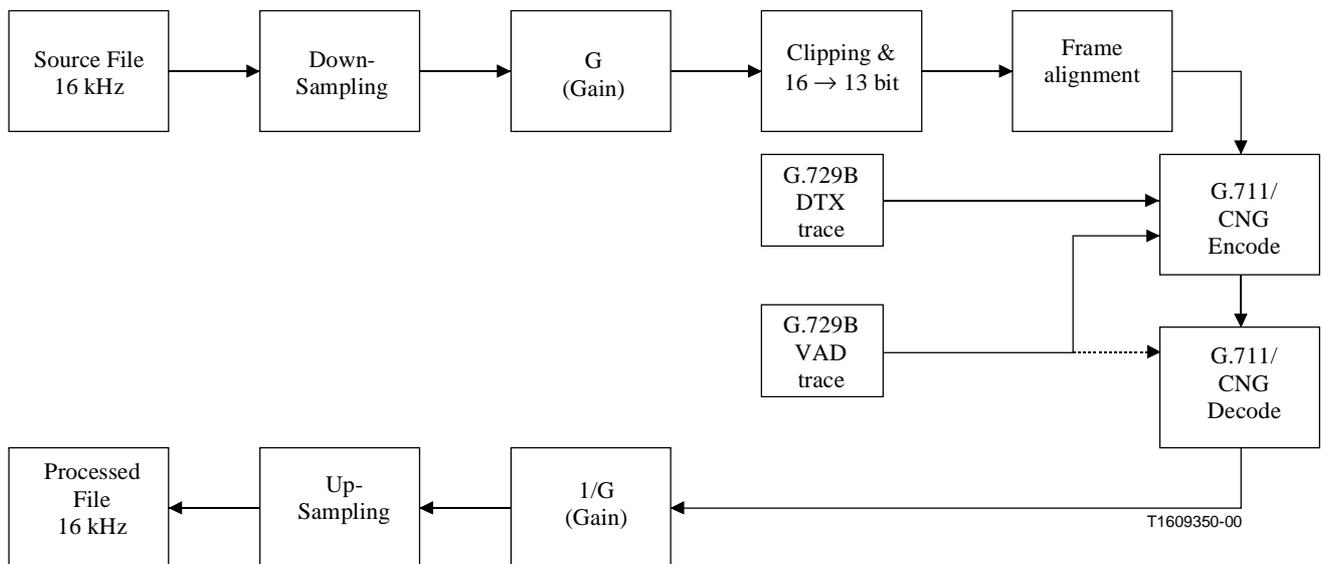


Figure II.4/G.711 – Processing G.711 without CNG



G.729B Is Annex B to Recommendation G.729

Figure II.5/G.711 – G.729B processing to obtain VAD/DTX trace files



G.729B Is Annex B to Recommendation G.729

Figure II.6/G.711 – Processing G.711 with CNG

II.5 Example solution

This subclause describes a comfort noise generation scheme using the comfort noise payload format described in this appendix, which was used in the assessment described in II.4.

II.5.1 Algorithm description

II.5.1.1 Encoder

The encoder must be called every frame by the calling program. For active voice frames, the input signal is pre-processed and the internal buffers are updated before returning. For inactive frames, the estimates of the background noise energy and spectral content are updated. In the case of an SID frame, the estimated parameters are quantized and packed into the channel buffer for transmission to the decoder. The SID update rate was determined by the DTX from G.729 Annex B [5]. The details of the CNG encoder are contained in the following subclauses.

II.5.1.1.1 Pre-processing

The input signal is pre-processed by a 1st order high-pass IIR filter to remove any undesired low-frequency component. The high-pass filter is given by:

$$H(z) = \frac{1 - z^{-1}}{1 - (127/128)z^{-1}}$$

II.5.1.1.2 Autocorrelation analysis

The normalised autocorrelation coefficients r_m and frame energy E are computed based on the pre-processed signal windowed with a 25 ms asymmetric window. For 8.0 kHz sampling rate, the window is given by:

$$w(n) = \begin{cases} 0.54 - 0.46 \cos\left(\frac{2\pi n}{339}\right) & n = 0, 1, \dots, 169 \\ \cos\left(\frac{2\pi(n-170)}{119}\right) & n = 170, 171, \dots, 199 \end{cases}$$

Running averages of both the normalised autocorrelation coefficients and the frame energy are then computed for the i th frame according to the following:

$$\begin{aligned} \bar{r}_m(i) &= \bar{r}_m(i-1) \cdot \beta_1 + \bar{r}_m(i) \cdot (1.0 - \beta_1) & m = 1, 2, \dots, M \\ \overline{LE}(i) &= \overline{LE}(i-1) \cdot \beta_2 + \overline{LE}(i) \cdot (1.0 - \beta_2) \end{aligned}$$

where LE is the base-2 logarithm of the frame energy, and M is the model order. β_1 and β_2 are frame size dependant constants. If the frame size is less than or equal to 7.5 ms, β_1 and β_2 are set to 0.8, otherwise they are set to 0.6. The averages are reset to the current frame values if the previous frame was active speech.

II.5.1.1.3 Reflection coefficient computation

The mean square error between the instantaneous and averaged normalised autocorrelation coefficients is computed according to the equation:

$$d = \frac{1}{M} \sum_{m=1}^M (\bar{r}_m(i) - r_m(i))^2$$

If d is less than an adaptive threshold Th and the last frame was inactive, the averaged coefficients $\bar{r}_m(i)$ are used for reflection coefficient computation; else the instantaneous coefficients $r_m(i)$ are used. The threshold Th is determined every frame according to the following algorithm:

```

if (PrevVad == 1)
    Th = 0.0
else
    Th += 0.2857 * (FRAME_SIZE / SAMPLING_RATE)
    if (Th > 0.06)
        Th = 0.06
    end
end
end

```

The reflection coefficients $k_m(i)$ are computed from the selected autocorrelation coefficients using the Levinson-Durbin algorithm.

II.5.1.1.4 Quantization

For Silence Insertion Descriptor (SID) frames, the energy $\overline{LE}(i)$ and the reflection coefficients $k_m(i)$ are quantized and packed according to the specified payload format.

II.5.1.2 Decoder

The decoder produces comfort noise by passing a scaled white noise excitation through a linear prediction synthesis filter. The details follow in the following subclauses.

II.5.1.2.1 Parameter update

The reflection coefficients from the last received SID frame are used in the current frame. Let the last received comfort noise parameters be denoted LE_{SID} where the energy has been converted from dBov to base-2 logarithm. The energy used in the current frame is given by:

$$LE(i) = LE(i-1) \cdot \alpha + LE_{SID} \cdot (1.0 - \alpha)$$

where $\alpha = 0.9$. This smoothing procedure is done to avoid any abrupt changes in signal energy in the comfort noise.

II.5.1.2.2 Excitation generation

A random number generator with a Gaussian distribution is used to produce the sequence Rn that is scaled by the factor η to the correct energy according to the equation:

$$\eta = \sqrt{\frac{E(i) \cdot \prod_{m=1}^M (1.0 - \hat{k}(N_m)^2)}{\frac{1}{L} \cdot \sum_{j=0}^{L-1} Rn(j)^2}}$$

where L is the length of the excitation, and $E(i)$ is the frame energy.

A constant approximation for the denominator of the above equation is used to avoid the dot product operation and reduce complexity.

II.5.1.2.3 LP synthesis

The reflection coefficients are converted to linear prediction coefficients for use in the linear prediction (LP) synthesis filter according to the following recursion [6]:

$$a_i^{(i)} = -\hat{k}_i(N_i)$$

$$a_j^{(i)} = a_j^{(i-1)} + \hat{k}_i(N_i)a_{i-j}^{(i-1)} \quad 1 \leq j \leq i-1$$

being solved for $i = 1, 2, \dots, M$ and with the final set defined as:

$$\alpha_j = a_j^{(p)} \quad 1 \leq j \leq M$$

The linear prediction synthesis filter is defined as:

$$\frac{1}{A(z)} = \frac{1}{1 - \sum_{j=1}^M \alpha_j z^{-j}}$$

The scaled excitation is passed through the filter to produce the final comfort noise. The length of the excitation L is, in general, equal to the frame length. However, for the first inactive frame following an active frame, L is equal to the frame length plus the model order (M). In this case, the first M output samples from the synthesis filter are ignored.

II.5.1.3 Delay

There is no delay inherent in the comfort noise algorithm.

II.5.1.4 Complexity

The algorithm has been implemented in 16-bit fixed-point using the ITU-T software Tool Library. The memory and resource usage at different frame sizes operating at 8.0 kHz sampling rate and a 10th order all-pole model is summarized in Table II.2. The WMOPS are obtained using the operations counter within the library and represent the worst case. The ROM is the estimated size on a typical fixed-point DSP.

Table II.2/G.711 – CNG Resource Requirements for a 10th order model

| Frame Size | RAM (words) | ROM (words) | WMOPS |
|------------|-------------|-------------|-------|
| 5 ms | 650 | 1300 | 1.1 |
| 10 ms | 690 | 1300 | 0.66 |
| 20 ms | 760 | 1300 | 0.47 |

II.5.2 Tested configuration

The algorithm as tested is specified in Table II.3.

Table II.3/G.711 – CNG Tested Configuration

| Parameter | As Tested |
|------------------|-----------|
| Sampling Rate | 8.0 kHz |
| Frame Size | 10 ms |
| Model Order | 10 |
| Look-Ahead Delay | 5 ms |

A look-ahead of 5 ms was added during testing by delaying the input to the accompanying speech codec (G.711) as in Figure II.7. The look-ahead was introduced to properly tailor the usage of the VAD of G.729 Annex B to the CNG example solution. The look-ahead delay can be avoided in practice by adding an extra hangover frame to the G.729 Annex B VAD.

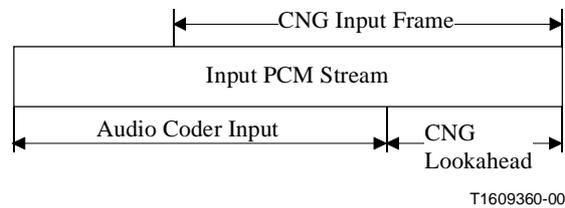


Figure II.7/G.711 – CNG Look-ahead during Testing

References

- [1] CCITT Recommendation G.726 (1990), *40, 32, 24, 16 kbit/s adaptive differential pulse code modulation (ADPCM)*.
- [2] CCITT Recommendation G.727 (1990), *5-, 4-, 3- and 2-bits/sample embedded adaptive differential pulse code modulation (ADPCM)*.
- [3] CCITT Recommendation G.728 (1992), *Coding of speech at 16 kbits/s using low-delay code excited linear prediction*.
- [4] CCITT Recommendation G.722 (1988), *7 kHz audio-coding within 64 kbit/s*.
- [5] ITU-T Recommendation G.729 Annex B (1996), *A silence compression scheme for G.729 optimized for terminals conforming to Recommendation V.70*.
- [6] RABINER (L.R.), SCHAFFER (R.W.): *Digital processing of speech signals*, Prentice-Hall, 1978.
- [7] ITU-T Recommendation G.191 (1996), *Software tools for speech and audio coding standardization*.

ITU-T RECOMMENDATIONS SERIES

- Series A Organization of the work of the ITU-T
- Series B Means of expression: definitions, symbols, classification
- Series C General telecommunication statistics
- Series D General tariff principles
- Series E Overall network operation, telephone service, service operation and human factors
- Series F Non-telephone telecommunication services
- Series G Transmission systems and media, digital systems and networks**
- Series H Audiovisual and multimedia systems
- Series I Integrated services digital network
- Series J Transmission of television, sound programme and other multimedia signals
- Series K Protection against interference
- Series L Construction, installation and protection of cables and other elements of outside plant
- Series M TMN and network maintenance: international transmission systems, telephone circuits, telegraphy, facsimile and leased circuits
- Series N Maintenance: international sound programme and television transmission circuits
- Series O Specifications of measuring equipment
- Series P Telephone transmission quality, telephone installations, local line networks
- Series Q Switching and signalling
- Series R Telegraph transmission
- Series S Telegraph services terminal equipment
- Series T Terminals for telematic services
- Series U Telegraph switching
- Series V Data communication over the telephone network
- Series X Data networks and open system communications
- Series Y Global information infrastructure and Internet protocol aspects
- Series Z Languages and general software aspects for telecommunication systems