TS 101 318 V1.1.1 (1998-08)

Technical Specification

Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON); Using GSM speech codecs within ITU-T Recommendation H.323

Reference

DTS/TIPHON-03001 (c7c00icr.PDF)

Keywords

codec, GSM, speech, internet, interworking, multimedia

ETSI

Postal address

F-06921 Sophia Antipolis Cedex - FRANCE

Office address

650 Route des Lucioles - Sophia Antipolis
Valbonne - FRANCE
Tel.: +33 4 92 94 42 00   Fax: +33 4 93 65 47 16
Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Internet

secretariat@etsi.fr
http://www.etsi.fr
http://www.etsi.org

Contents

Intellectual Property Rights

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available **free of charge** from the ETSI Secretariat. Latest updates are available on the ETSI Web server (http://www.etsi.fr/ipr or http://www.etsi.org/ipr).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Foreword

This Technical Specification (TS) has been produced by ETSI Project Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON).

1 Scope

The present document intends to enable interoperable use of GSM codecs between ITU-T Recommendation H.323 [16] systems from different vendors. It defines the presentation of GSM codec capabilities for ITU-T Recommendation H.245 [14] and new RTP encodings for GSM speech codecs. It includes an informative presentation of canonical RTP encoding defined by IETF for the GSM 06.10 speech codec.

The present document is applicable to all multimedia systems using such versions of ITU-T Recommendation H.245 [14] control protocol versions that do not readily support GSM speech codecs.

2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non-specific.
- For a specific reference, subsequent revisions do not apply.
- For a non-specific reference, the latest version applies.
- A non-specific reference to an ETS shall also be taken to refer to later versions published as an EN with the same number.

[1] EG 200 351 (V2.1): "ETSI object identifier tree, Rules and registration procedures".

[2] ETS 300 961: "Digital cellular telecommunications system (Phase 2+); Full rate speech; Transcoding (GSM 06.10)".

[3] ETS 300 963: "Digital cellular telecommunications system; Full rate speech; Comfort noise aspect for full rate speech traffic channels (GSM 06.12)".

[4] ETS 300 964: "Digital cellular telecommunications system; Full rate speech; Discontinuous Transmission (DTX) for full rate speech traffic channels (GSM 06.31)".

[5] ETS 300 969: "Digital cellular telecommunications system (Phase 2+); Half rate speech; Half rate speech transcoding (GSM 06.20)".

[6] ETS 300 971: "Digital cellular telecommunications system; Half rate speech; Comfort noise aspects for the half rate speech traffic channels (GSM 06.22)".

[7] ETS 300 972: "Digital cellular telecommunications system (Phase 2+); Half rate speech; Discontinuous Transmission (DTX) for half rate speech traffic channels (GSM 06.41)".

[8] ETS 300 726: "Digital cellular telecommunications system; Enhanced Full Rate (EFR) speech transcoding (GSM 06.60)".

[9] ETS 300 728: "Digital cellular telecommunications system; Comfort noise aspects for Enhanced Full Rate (EFR) speech traffic channels (GSM 06.62)".

[10] ETS 300 729: "Digital cellular telecommunications system; Discontinuous Transmission (DTX) for Enhanced Full Rate (EFR) speech traffic channels (GSM 06.81)".

[11] ETS 300 909: "Digital cellular telecommunications system (Phase 2+); Channel coding (GSM 05.03)".

[12] ETS 300 737: "Digital cellular telecommunications system (Phase 2+); In-band control of remote transcoders and rate adaptors for Enhanced Full Rate (EFR) and full rate traffic channels (GSM 08.60)".

[13] ETS 300 598: "European digital cellular telecommunications system (Phase 2); In-band control of remote transcoders and rate adaptors for half rate traffic channels (GSM 08.61)".

[14] ITU-T Recommendation H.245: "Control protocol for multimedia communication".

[15] ITU-T Recommendation H.225.0: "Call signalling protocols and media stream packetization for packet based multimedia communication systems".

[16] ITU-T Recommendation H.323: "Packet based multimedia communications systems".

[17] RFC 1889, H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson: "RTP: A Transport Protocol for Real-Time Applications".

[18] Internet draft ietf-avt-profile-07, H. Schulzrinne: "RTP Profile for Audio and Video Conferences with Minimal Control".

NOTE The above Internet draft is expected to replace the RFC 1890 (same title) as the Internet standards track document for RTP audio/video profiles.

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the following definitions apply:

bit-order: the order of numbering of bits, either big-endian or little-endian.

big-endian: refers to which bytes are most significant in multi-byte data types . In big-endian architectures, the leftmost bytes (those with a lower address) are most significant.

little-endian: refers to which bytes are most significant in multi-byte data types . In little-endian architectures, the rightmost bytes are most significant.

SID codeword: fixed bit pattern for labelling a traffic frame as a SID frame.

SID field: the bit positions of the SID codeword within a SID frame.

SID frame: frame conveying information on the acoustic background noise. A SID frame is characterized by the SID codeword.

speech frame: traffic frame that is not classified as a SID frame.

3.2 Abbreviations

For the purposes of the present document, the following abbreviations apply:

DTX	Discontinuous Transmission
SID	Silence Descriptor
TS	Technical Specification

4 ITU-T Recommendation H.245 and GSM speech codecs

The ITU-T Recommendation H.245 [14] specifies a control protocol, which is used by several multimedia systems specified by ITU-T and others. There is no standard way to negotiate and use the GSM speech codecs in ITU-T Recommendation H.245 [14] versions 1 or 2. However, ITU-T Recommendation H.245 [14] **AudioCapability** has element **nonStandardParameter** which can be used to include any audio codec capability not included by ITU-T. The **nonStandardParameter** is defined as follows:

```
NonStandardParameter    ::=SEQUENCE
{
    nonStandardIdentifier   NonStandardIdentifier,
    data                    OCTET STRING
}

NonStandardIdentifier   ::=CHOICE
{
    object      OBJECT IDENTIFIER,
    h221NonStandard SEQUENCE
    {
        t35CountryCode      INTEGER (0..255),   -- country, per T.35
        t35Extension        INTEGER (0..255),   -- assigned nationally
        manufacturerCode    INTEGER (0..65535)  -- assigned nationally
    }
}
```

The **nonStandardParameters** are identified either by an ASN.1 **OBJECT IDENTIFIER** or by a T.35 manufacturer code.

4.1 A NonStandardParameter for GSM speech codecs

When any of the GSM speech codecs [2], [5], [8] would be included in the **AudioCapability**, the **object** element shall contain **gsmFullRateId**, **gsmHalfRateId** or **gsmEnhancedFullRateId OBJECT IDENTIFIER**. The **data** element shall contain **TIPHONGSMAudioCapability**.

```
-- $Id: gsm.asn,v 1.2 1997/08/27 17:07:23 pessi Exp pessi $ --

TIPHON-ETS-XXX-YYY DEFINITIONS ::=
BEGIN

tiphonGsmId OBJECT IDENTIFIER ::= {
  ccitt(0) identified-organization(4) etsi(0) tiphonGsm(XYYY) version(0)
}

-- The object element of nonStandardIdentifier element in the
-- NonStandardParameter shall contain one of following object
-- identifiers

gsmFullRateId         OBJECT IDENTIFIER ::= { tiphonGsmId fr(0) }
gsmHalfRateId         OBJECT IDENTIFIER ::= { tiphonGsmId hr(1) }
gsmEnhancedFullRateId OBJECT IDENTIFIER ::= { tiphonGsmId efr(2) }

-- The data element in the NonStandardParameter shall contain
-- the following

GSMAudioCapability ::= SEQUENCE
{
   audioUnitSize        INTEGER(1..256),
   comfortNoise         BOOLEAN,  -- codec generates SID frames
   scrambled            BOOLEAN,  -- alternative RTP frame encoding
   ...
}

END
```

The normal ITU-T Recommendation H.245 [14] encoding of ASN.1, i.e., the aligned variant PER encoding, shall be used for **GSMAudioCapability**, too.

If the **comfortNoise** is **TRUE**, the endpoint shall send the SID frames as specified in the comfort noise [3], [6], [9] and DTX [4], [7], [10] specifications of the GSM codec in question.

The use of the **scrambled** field is for further study.

5 RTP encodings

An RTP packet can contain one or more output frames from audio codec [17], [18]. Each frame is contained in an octet string, which contains an integral number of octets. In other words, frames do not overlap over byte boundaries. The octet string containing one frame is referred as *buffer* below. RTP encoding defines how the codec parameters are stored in this buffer.

The bits of the buffer are numbered in big-endian manner. The most significant bit of first octet is $\mathbf{r}_1$, the least significant bit in the same octet is $\mathbf{r}_8$, the most significant bit in second octet is $\mathbf{r}_9$ and so on.

The bits of the codec parameters are numbered in the same way as they are numbered in the respective codec specifications [2], [5], [8].

5.1 GSM 06.10 full rate codec

The GSM full rate codec has frame length of 260 bits. In the canonical encoding used for RTP, the bits are packed beginning from the most significant bit. Every frame is coded into one 33 octet (264 bit) buffer. Every such buffer begins with a 4 bit magic value or signature (0xD, binary 1101), followed by the MSB encoding of the fields of the frame.

In table 1, the RTP buffer bits are numbered in the big-endian manner from $\mathbf{r}_1$ to $\mathbf{r}_{264}$.

The bits in GSM 06.10 codec parameters are numbered in little-endian order starting from 0, i.e. the least significant bit in each parameter is number 0. The first octet thus contains 1101 in the 4 most significant bits ($\mathbf{r}_1$ - $\mathbf{r}_4$) and the 4 most significant bits of $\mathbf{LAR_c(0)}$ (bits 5, 4, 3 and 2) in the 4 least significant bits ($\mathbf{r}_5$-$\mathbf{r}_8$).

5.1.1 Encoding of speech frames

The normal GSM 06.10 speech frames are encoded for RTP according to table 1. This is the canonical RTP encoding as specified in the IETF AVT profile (see annex A).

Table 1: The order of GSM 06.10 full rate speech codec parameters in the canonical RTP encoding

Parameter	Bits	Bit No. (MSB-LSB)
Signature	4	r1 - r4
$LAR_c(0)$	6	r5 - r10
$LAR_c(1)$	6	r11 - r16
$LAR_c(2)$	5	r17 - r21
$LAR_c(3)$	5	r22 - r26
$LAR_c(4)$	4	r27 - r30
$LAR_c(5)$	4	r31 - r34
$LAR_c(6)$	3	r35 - r37
$LAR_c(7)$	3	r38 - r40

Parameter	Bits	Bit No. (MSB-LSB)
N_c	7	r41 - r47
b_c	2	r48 - r49
M_c	2	r50 - r51
x_{maxc}	6	r52 - r57
$x_{Mc}(0)$	3	r58 - r60
$x_{Mc}(1)$	3	r61 - r63
...		
$x_{Mc}(12)$	3	r94 - r96

Parameter	Bits	Bit No. (MSB-LSB)
N_c	7	r97 - r103
b_c	2	r104 - r105
M_c	2	r106 - r107
x_{maxc}	6	r108 - r113
$x_{Mc}(0)$	3	r114 - r116
$x_{Mc}(1)$	3	r117 - r119
	...	
$x_{Mc}(12)$	3	r150 - r152

Parameter	Bits	Bit No. (MSB-LSB)
N_c	7	r153 - r159
b_c	2	r160 - r161
M_c	2	r162 - r163
x_{maxc}	6	r164 - r169
$x_{Mc}(0)$	3	r170 - r172
$x_{Mc}(1)$	3	r173 - r175
	...	
$x_{Mc}(12)$	3	r206 - r208

Parameter	Bits	Bit No. (MSB-LSB)
N_c	7	r209 - r215
b_c	2	r216 - r217
M_c	2	r218 - r219
x_{maxc}	6	r220 - r225
$x_{Mc}(0)$	3	r226 - r228
$x_{Mc}(1)$	3	r229 - r231
	...	
$x_{Mc}(12)$	3	r262 - r264

5.1.2 Encoding of silence indication frames

The SID frame is encoded according to ETS 300 963 [3].

The signature of the full rate SID frame shall be `0xD`.

A SID frame is identified by a SID code word consisting of 95 bits of $\mathbf{x_{Mc}(i)}$ parameters in the error protection class I, which all shall be 0 (see table 2 in ETS 300 909 [11]). These bits are:

- the most significant bit (bit number 2) of all $\mathbf{x_{Mc}(0..12)}$ parameters in every sub-frame;
- the middle bit (bit number 1) of the $\mathbf{x_{Mc}(0..12)}$ parameters in sub-frames 1, 2 and 3;

and

- the middle bit (bit number 1) of the $\mathbf{x_{Mc}(0..3)}$ parameters in sub-frame 4.

In other words, the following bits in the canonical RTP encoding form the SID code word:

> $\mathbf{r}_{58}$, $\mathbf{r}_{59}$, $\mathbf{r}_{61}$, $\mathbf{r}_{62}$, $\mathbf{r}_{64}$, $\mathbf{r}_{65}$, $\mathbf{r}_{67}$, $\mathbf{r}_{68}$, $\mathbf{r}_{70}$, $\mathbf{r}_{71}$, $\mathbf{r}_{73}$, $\mathbf{r}_{74}$, $\mathbf{r}_{76}$, $\mathbf{r}_{77}$, $\mathbf{r}_{79}$, $\mathbf{r}_{80}$, $\mathbf{r}_{82}$, $\mathbf{r}_{83}$, $\mathbf{r}_{85}$, $\mathbf{r}_{86}$, $\mathbf{r}_{88}$, $\mathbf{r}_{89}$, $\mathbf{r}_{91}$, $\mathbf{r}_{92}$, $\mathbf{r}_{94}$, $\mathbf{r}_{95}$, $\mathbf{r}_{114}$, $\mathbf{r}_{115}$, $\mathbf{r}_{117}$, $\mathbf{r}_{118}$, $\mathbf{r}_{120}$, $\mathbf{r}_{121}$, $\mathbf{r}_{123}$, $\mathbf{r}_{124}$, $\mathbf{r}_{126}$, $\mathbf{r}_{127}$, $\mathbf{r}_{129}$, $\mathbf{r}_{130}$, $\mathbf{r}_{132}$, $\mathbf{r}_{133}$, $\mathbf{r}_{135}$, $\mathbf{r}_{136}$, $\mathbf{r}_{138}$, $\mathbf{r}_{139}$, $\mathbf{r}_{141}$, $\mathbf{r}_{142}$, $\mathbf{r}_{144}$, $\mathbf{r}_{145}$, $\mathbf{r}_{147}$, $\mathbf{r}_{148}$, $\mathbf{r}_{150}$, $\mathbf{r}_{151}$, $\mathbf{r}_{170}$, $\mathbf{r}_{171}$, $\mathbf{r}_{173}$, $\mathbf{r}_{174}$, $\mathbf{r}_{176}$, $\mathbf{r}_{177}$, $\mathbf{r}_{179}$, $\mathbf{r}_{180}$, $\mathbf{r}_{182}$, $\mathbf{r}_{183}$, $\mathbf{r}_{185}$, $\mathbf{r}_{186}$, $\mathbf{r}_{188}$, $\mathbf{r}_{189}$, $\mathbf{r}_{191}$, $\mathbf{r}_{192}$, $\mathbf{r}_{194}$, $\mathbf{r}_{195}$, $\mathbf{r}_{197}$, $\mathbf{r}_{198}$, $\mathbf{r}_{200}$, $\mathbf{r}_{201}$, $\mathbf{r}_{203}$, $\mathbf{r}_{204}$, $\mathbf{r}_{206}$, $\mathbf{r}_{207}$, $\mathbf{r}_{226}$, $\mathbf{r}_{227}$, $\mathbf{r}_{229}$, $\mathbf{r}_{230}$, $\mathbf{r}_{232}$, $\mathbf{r}_{233}$, $\mathbf{r}_{235}$, $\mathbf{r}_{236}$, $\mathbf{r}_{238}$, $\mathbf{r}_{241}$, $\mathbf{r}_{244}$, $\mathbf{r}_{247}$, $\mathbf{r}_{250}$, $\mathbf{r}_{253}$, $\mathbf{r}_{256}$, $\mathbf{r}_{259}$ and $\mathbf{r}_{262}$.

In the SID frames, the $\mathbf{LAR_c(i)}$ parameters are replaced by the **mean** $\mathbf{(LAR(i))}$ values. The block amplitude values $\mathbf{x_{MAXc}(0..3)}$ are replaced by the **mean** $\mathbf{(x_{MAX})}$ value, repeated four times.

All remaining bits are set to zero when sending and ignored when receiving a SID frame.

5.1.3 Encoding of scrambled speech frames

The use of the **scrambled** field is for further study.

5.2 GSM 06.20 half rate codec

The GSM half rate codec has frame length of 112 bits. Every frame is encoded into one 14 octet (112 bit) buffer. There shall be no signature.

The bits in RTP buffer are numbered from $\mathbf{r}_1$ (the most significant bit of first octet) to $\mathbf{r}_{112}$ (the least significant bit of last octet). Within the GSM 06.20 codec parameter bits are numbered in big-endian manner.

The order of occurrence of the codec parameters in the buffer is the same as order of occurrence over the Abis as defined in annex B of ETS 300 969 [5].

The discontinuous transmission (ETS 300 972 [7]) and the generation of comfort noise (ETS 300 971 [6]) is integral part of the half rate codec. RTP implementations shall be able to support DTX according to [6], [7].

5.2.1 Encoding of speech frames

There are two alternative formats of a ETS 300 969 [5] speech frames. The first form is for codec mode 0 (unvoiced speech), the second form is for modes 1, 2 and 3 (voiced speech).

Table 2: The order of GSM 06.20 half rate speech codec parameters in RTP buffer (MODE=0)

Parameter	No. of bits	Bit No. (MSB-LSB)
R0	5	r1 - r5
LPC1	11	r6 - r16
LPC2	9	r17 - r25
LPC3	8	r26 - r33
INT_LPC	1	r34
MODE	2	r35 - r36
CODE1_1	7	r37 - r43
CODE2_1	7	r44 - r50
GSP0_1	5	r51 - r55
CODE1_2	7	r56 - r62
CODE2_2	7	r63 - r69
GSP0_2	5	r70 - r74
CODE1_3	7	r75 - r81
CODE2_3	7	r82 - r88
GSP0_3	5	r89 - r93
CODE1_4	7	r94 - r100
CODE2_4	7	r101 - r107
GSP0_4	5	r108 - r112

Table 3: The order of GSM 06.20 half rate speech codec parameters in RTP buffer (MODE=1, 2 or 3)

Parameter	No. of bits	Bit No. (MSB-LSB)
R0	5	r1 - r5
LPC1	11	r6 - r16
LPC2	9	r17 - r25
LPC3	8	r26 - r33
INT_LPC	1	r34
MODE	2	r35 - r36
LAG_1	8	r37 - r44
CODE1	9	r45 - r53
GSP0_1	5	r54 - r58
LAG_2	4	r59 - r62
CODE2	9	r63 - r71
GSP0_2	5	r72 - r76
LAG_3	4	r77 - r80
CODE3	9	r81 - r89
GSP0_3	5	r90 - r94
LAG_4	4	r95 - r98
CODE4	9	r99 - r107
GSP0_4	5	r108 - r112

5.2.2 Encoding of silence indication frames

The half-rate codec SID frame is encoded according to the ETS 300 971 [6].

A SID frame is identified by a SID codeword consisting of 79 bits which are all 1. The parameters in table 4 have to be set as shown in order to mark a frame as a SID frame.

Table 4: SID codeword for half rate speech codec

Parameter	No. of bits	Value (Hex)
INT_LPC	1	1_{16}
MODE	2	3_{16}
LAG_1	8	FF_{16}
CODE1	9	$1FF_{16}$
GSP0_1	5	$1F_{16}$
LAG_2	4	F_{16}
CODE2	9	$1FF_{16}$
GSP0_2	5	$1F_{16}$
LAG_3	4	F_{16}
CODE3	9	$1FF_{16}$
GSP0_3	5	$1F_{16}$
LAG_4	4	F_{16}
CODE4	9	$1FF_{16}$
GSP0_4	5	$1F_{16}$

5.2.3 Encoding of scrambled speech frames

The use of the scrambled field is for further study.

5.3 GSM 06.60 enhanced full rate codec

The ETS 300 726 [8] enhanced full rate codec has frame length of 244 bits. In the encoding used for RTP, the bits are packed beginning from the most significant bit. Every frame is coded into one 31 octet (248 bit) buffer. Every such buffer shall begin with a 4 bit signature (**`0xC`**, binary **`1100`**), followed by the MSB encoding of the fields of the frame.

The bits in buffer are numbered in the big-endian manner, starting from $\mathbf{r}_1$ (the most significant bit of first octet) and finishing to $\mathbf{r}_{248}$ (the least significant bit of last octet). The ETS 300 726 [8] speech codec parameter bits are numbered in little-endian manner starting from 0. The first octet in buffer thus contains 1101 in its 4 most significant bits ($\mathbf{r}_1 - \mathbf{r}_4$) and the bits 6, 5, 4 and 3 from the index of 1st LSF submatrix in its 4 least significant bits ($\mathbf{r}_5 - \mathbf{r}_8$).

The order of occurrence of the codec parameters in the buffer is the same as order of occurrence within the speech frame, as defined in the table 5 of ETS 300 726 [8]. The redundancy bits specified by ETS 300 909 [11] are not included in the RTP encoding.

The discontinuous transmission (ETS 300 729 [10]) and the generation of comfort noise (ETS 300 728 [9]) is integral part of the half rate codec. RTP implementations shall be able to support DTX according to [9], [10].

5.3.1 Encoding of speech frames

The normal ETS 300 726 [8] speech frames are encoded for RTP according to table 5.

Table 5: The order of GSM 06.20 enhanced full rate speech codec parameters in RTP buffer

Parameter	Bits	Bit No. (MSB-LSB)
Encoding signature	4	r1 - r4
index of 1st LSF submatrix	7	r5 - r11
index of 2nd LSF submatrix	8	r12 - r19
index of 3rd LSF submatrix	8	r20 - r27
sign of 3rd LSF submatrix	1	r28
index of 4th LSF submatrix	8	r29 - r36
index of 5th LSF submatrix	6	r37 - r42
Sub-frame 1		
adaptive codebook index	9	r43 - r51
adaptive codebook gain	4	r52 - r55
sign for 1st and 6th pulses	1	r56
position of 1st pulse	3	r57 - r59
sign for 2nd and 7th pulses	1	r60
position of 2nd pulse	3	r61 - r63
sign for 3rd and 8th pulses	1	r64
position of 3rd pulse	3	r65 - r67
sign for 4th and 9th pulses	1	r68
position of 4th pulse	3	r69 - r71
sign for 5th and 10th pulses	1	r72
position of 5th pulse	3	r73 - r75
position of 6th pulse	3	r76 - r78
position of 7th pulse	3	r79 - r81
position of 8th pulse	3	r82 - r84
position of 9th pulse	3	r85 - r87
position of 10th pulse	3	r88 - r90
fixed codebook gain	5	r91 - r95
Sub-frame 2		
adaptive codebook index (relative)	6	r96 - r101
same parameters as b52 - b95	44	r102 - r145
Sub-frame 3		
same parameters as b43 - b95	53	r146 - r198
Sub-frame 4		
same parameters as b96 - b145	50	r199 - r248

5.3.2 Encoding of silence indication frames

The SID frame is encoded according to ETS 300 728 [9].

A SID frame is identified by a SID codeword consisting of 95 bits which are all 1. The parameters in table 4 have to be set as shown in order to mark a frame as a SID frame.

Table 6: SID codeword for enhanced full rate speech codec

Parameter	Bits	Value
LTP LAG 1	2	3_{16}
LTP LAG 2	3	7_{16}
LTP LAG 3	2	3_{16}
LTP LAG 4	4	F_{16}
LTP GAIN 1	3	7_{16}
LTP GAIN 2	3	7_{16}
LTP GAIN 3	4	F_{16}
LTP GAIN 4	4	F_{16}
PULSE 1 of 1st subfr.	4	F_{16}
PULSE 2 of 1st subfr.	4	F_{16}
PULSE 3 of 1st subfr.	4	F_{16}
PULSE 4 of 1st subfr.	4	F_{16}
PULSE 5 of 1st subfr.	2	3_{16}
PULSE 1 of 2nd subfr.	4	F_{16}
PULSE 2 of 2nd subfr.	4	F_{16}
PULSE 3 of 2nd subfr.	4	F_{16}
PULSE 4 of 2nd subfr.	4	F_{16}
PULSE 5 of 2nd subfr.	2	3_{16}
PULSE 1 of 3rd subfr.	4	F_{16}
PULSE 2 of 3rd subfr.	4	F_{16}
PULSE 3 of 3rd subfr.	4	F_{16}
PULSE 4 of 3rd subfr.	4	F_{16}
PULSE 5 of 3rd subfr.	2	3_{16}
PULSE 1 of 4th subfr.	4	F_{16}
PULSE 2 of 4th subfr.	2	3_{16}
PULSE 3 of 4th subfr.	4	F_{16}
PULSE 4 of 4th subfr.	4	F_{16}
PULSE 5 of 4th subfr.	2	3_{16}

The quantization indices of the LP parameters are replaced by the quantization indices derived from the averaged LSF parameter vector $\mathbf{f}^{mean}$. The encoding of the quantization indices is defined in ETS 300 728 [9], subclause 5.1.

The fixed codebook gain quantization indices are replaced by the quantization index derived from the averaged fixed codebook gain value g_c^{mean}, encoded as defined in ETS 300 728 [9] subclause 5.1, repeated four times inside the frame.

All remaining bits are set to zero when sending and ignored when receiving a SID frame.

5.3.3 Encoding of scrambled speech frames

The use of the **scrambled** field is for further study.

Annex A (informative): GSM 06.10 RTP Profile for Audio and Video Conferences with Minimal Control

The following excerpt is from [18].

GSM

GSM (group speciale mobile) denotes the European GSM 06.10 provisional standard for full-rate speech transcoding, […], which is based on RPE/LTP (residual pulse excitation/long term prediction) coding at a rate of 13 kb/s […].

Blocks of 160 audio samples are compressed into 33 octets, for an effective data rate of 13,200 b/s.

General Packaging Issues

The GSM audio encoding used in the RTP protocol uses the same calculation and order of fields in the frame as the ACM codec supplied by Microsoft, however, the packing of those fields in every GSM frame differs.

In the Microsoft codec, every two frames (320 samples) are packed into a buffer of 65 octets (520 bits). The packing begins from the least significant bit of every octet, and the least significant bits of the field are packed first. For instance, if the first field is F1 and contains 6 bits, and the second field is F2 which also contains 6 bits, they are packed in the following way: the first octet contains F1 in bits 0-5, and the lower two bits of F2 in bits 6-7. The second octet contains the high 4 bits of F2 in bits 0-3, and the 4 least significant bits of F3 in bits 4-7, and so on. The 33rd octet contains the last 4 bits of the first GSM frame in its lower 4 bits (0-3) and the first 4 bits of the next frame (the lower 4 bits of F1) in bits 4-7.

In the GSM encoding used by RTP, the bits are packed beginning from the most significant bit. Every 160 sample GSM frame is coded into one 33 octet (264 bit) buffer. Every such buffer begins with a 4 bit signature (0xD), followed by the MSB encoding of the fields of the frame. The first octet thus contains 1101 in the 4 most significant bits (4-7) and the 4 most significant bits of F1 (2-5) in the 4 least significant bits (0-3). The second octet contains the 2 least bits of F1 in bits 6-7, and F2 in bits 0-5, and so on. The order of the fields in the frame is as follows:

GSM variable names and numbers

So if F.i signifies the i^{th} bit of the field F, and bit 0 is the most significant bit, and the bits of every octet are numbered from 0 to 7 from left to right, then in the RTP encoding we have:

```
              Octet
                     Bit 0     Bit 1      Bit 2      Bit 3
                     Bit 4     Bit 5      Bit 6      Bit 7
----------------------------------------------------------
                0      1          1          0          1
                     LARc0.0    LARc0.1    LARc0.2    LARc0.3
                1    LARc0.4    LARc0.5    LARc1.0    LARc1.1
                     LARc1.2    LARc1.3    LARc1.4    LARc1.5
                2    LARc2.0    LARc2.1    LARc2.2    LARc2.3
                     LARc2.4    LARc3.0    LARc3.1    LARc3.2
```

Table A.1: Ordering of GSM variables

field	name	bits	field	name	bits
1	LARc[0]	6	39	xmc[22]	3
2	LARc[1]	6	40	xmc[23]	3
3	LARc[2]	5	41	xmc[24]	3
4	LARc[3]	5	42	xmc[25]	3
5	LARc[4]	4	43	Nc[2]	7
6	LARc[5]	4	44	bc[2]	2
7	LARc[6]	3	45	Mc[2]	2
8	LARc[7]	3	46	xmaxc[2]	6
9	Nc[0]	7	47	xmc[26]	3
10	bc[0]	2	48	xmc[27]	3
11	Mc[0]	2	49	xmc[28]	3
12	xmaxc[0]	6	50	xmc[29]	3
13	xmc[0]	3	51	xmc[30]	3
14	xmc[1]	3	52	xmc[31]	3
15	xmc[2]	3	53	xmc[32]	3
16	xmc[3]	3	54	xmc[33]	3
17	xmc[4]	3	55	xmc[34]	3
18	xmc[5]	3	56	xmc[35]	3
19	xmc[6]	3	57	xmc[36]	3
20	xmc[7]	3	58	xmc[37]	3
21	xmc[8]	3	59	xmc[38]	3
22	xmc[9]	3	60	Nc[3]	7
23	xmc[10]	3	61	bc[3]	2
24	xmc[11]	3	62	Mc[3]	2
25	xmc[12]	3	63	xmaxc[3]	6
26	Nc[1]	7	64	xmc[39]	3
27	bc[1]	2	65	xmc[40]	3
28	Mc[1]	2	66	xmc[41]	3
29	xmaxc[1]	6	67	xmc[42]	3
30	xmc[13]	3	68	xmc[43]	3
31	xmc[14]	3	69	xmc[44]	3
32	xmc[15]	3	70	xmc[45]	3
33	xmc[16]	3	71	xmc[46]	3
34	xmc[17]	3	72	xmc[47]	3
35	xmc[18]	3	73	xmc[48]	3
36	xmc[19]	3	74	xmc[49]	3
37	xmc[20]	3	75	xmc[50]	3
38	xmc[21]	3	76	xmc[51]	3

Bibliography

The following material, though not specifically referenced in the body of the present document, gives supporting information.

ITU-T Recommendation X.680: "Information Technology – Abstract Syntax Notation One (ASN.1) – Specification of basic notation".

ITU-T Recommendation X.691: "Information Technology –ASN.1 Encoding Rules –Specification of Packed Encoding Rules (PER)".

ETS 300 960: "Digital cellular telecommunications system; Full rate speech; Processing functions (GSM 06.01)".

ETS 300 962: "Digital cellular telecommunications system; Full rate speech; Substitution and muting of lost frames for full rate speech channels (GSM 06.11)".

ETS 300 965: "Digital cellular telecommunications system (Phase 2+); Voice Activity Detector (VAD) for full rate speech traffic channels (GSM 06.32)".

ETS 300 967: "Digital cellular telecommunications system (Phase 2+); Half rate speech; ANSI-C code for the GSM half rate speech codec (GSM 06.06)".

ETS 300 968: "Digital cellular telecommunications system (Phase 2+); Half rate speech; Test sequences for the GSM half rate speech codec (GSM 06.07)".

ETS 300 970: "Digital cellular telecommunications system; Half rate speech; Substitution and muting of lost frames for half rate speech traffic channels (GSM 06.21)".

ETS 300 973: "Digital cellular telecommunications system; Half rate speech; Voice Activity Detector (VAD) for half rate speech traffic channels (GSM 06.42)".

ETS 300 723: "Digital cellular telecommunications system; Enhanced Full Rate (EFR) speech processing functions; General description (GSM 06.51)",

ETS 300 724: "Digital cellular telecommunications system; ANSI-C code for the GSM Enhanced Full Rate (EFR) speech codec (GSM 06.53)".

ETS 300 725: "Digital cellular telecommunications system (Phase 2+); Test sequences for the GSM Enhanced Full Rate (EFR) speech codec (GSM 06.54)".

ETS 300 727: "Digital cellular telecommunications system; Substitution and muting of lost frame for Enhanced Full Rate (EFR) speech traffic channels (GSM 06.61)".

ETS 300 730: "Digital cellular telecommunications system; Voice Activity Detector (VAD) for Enhanced Full Rate (EFR) speech traffic channels (GSM 06.82)".

History

Document history		
V1.1.1	August 1998	Publication

ISBN 2-7437-2415-3
Dépôt légal : Aout 1998