SDP syntax for H.263 options

STATUS OF THIS MEMO

This document is an Internet-Draft. Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

To view the entire list of current Internet-Drafts, please check the "iid-abstracts.txt" listing contained in the Internet-Drafts Shadow Directories on ftp.is.co.za (Africa), ftp.nordu.net (Northern Europe), ftp.nis.garr.it (Southern Europe), munnari.oz.au (Pacific Rim), ftp.ietf.org (US East Coast), or ftp.isi.edu (US West Coast).

ABSTRACT

This document defines the SDP syntax for 1998 version of H.263 codec options and parameters for IETF multimedia conferencing architecture. It is often useful to know beforehand (in call setup phase) what features of H.263 the other end supports. This document specifies the format specific parameters for H.263, which exists in a=fmtp:<format> <format specific parameters> as defined in the SDP document. Moreover, the usage of this feature with SIP and SAP is specified.

1. INTRODUCTION

Internet multimedia conferencing is becoming a reality and new protocols have been defined to provide co-operation between different applications. Current IETF protocols SIP [3] and SAP [4] are quite widely used to simple MBone conferencing and Internet unicast conferencing is also taking its first steps. Today, the most widely used video codec in this environment is the ITU-T H.261 video standard, but a better and more efficient low-bit rate video standard, ITU-T H.263 [1] (aka H.263+ and H.263v2) is becoming more and more popular.

However, the usage of H.263 with current IETF multimedia conferencing architecture is difficult, since improved efficiency of H.263 depends on use of coding options and parameters which must be told somehow to the other end. SIP and SAP protocols can deliver this information, but current SDP [2] format does not specify standardised way of representing these options and parameters.
This document specifies the SDP syntax which describes H.263 options and parameters to be delivered by some signaling protocol (e.g. SIP or SAP). The syntax presented in this document uses SDP "a=fmtp" attribute which is meant for carrying codec specific parameters.

These options are especially useful in SIP. In SAP it is not clear whether they have any use. In SIP, these format specific parameters are decoder properties or wishes, and in SAP they informative announcement about forthcoming video options and parameters. The SAP rules are applied also if a multicast session is advertised in a web page in SDP format.

The syntax in this document is text-based and uses the ISO 10646 character set in UTF-8 encoding (RFC 2279 [5]). This syntax includes one line and it is terminated by CRLF.

Syntax is described in an augmented Backus-Naur form (BNF) described in detail in RFC 2234 [6].

2. H.263 CODEC PARAMETERS

For further description about these H.263 parameters, please refer to ITU documents [1].

The context in which SDP syntax for H.263 is represented is SDP "a" attribute:

```
a=fmtp 34 SDP_263_syntax
```

"34" is RTP payload number for H.263.

The syntax itself is defined as follows:

```
SDP_263_syntax = *capability CRLF | request CRLF | CRLF
capability = 1*Size SP | 1*Params SP | 1*Options SP
```

Here as usual, SP means space and *term means that zero or more instances of term may be present. 1*term means that one or more instances of term may appear.

SDP_263_syntax may be either capability announcement, intra request during video session or empty line when having only CRLF (end of line). Request is explained in chapter 4.

2.1 Picture information and MPI

H.263 bitstream supports many picture sizes and different frame rates. Supported picture sizes and their corresponding minimum picture interval (MPI) values are combined with "=" symbol.

MPI is an integer value (1..32) and it means that maximum picture (frame) rate is (29.97/MPI) frames/sec. Bigger MPI value means slower picture rate.

Augmented BNF syntax for size related H.263 parameters is given below.

```
Size= "SQCIF" "%" mpi | "QCIF" "%" mpi | "CIF" "%" mpi | "CIF4" "%" mpi | "CIF16" "%" mpi | "XMAX" "%" xmax SP "YMAX" "%" ymax SP "MPI" "%" mpi

mpi=1*2DIGIT
```
Explanations:

SQCIF:
Sub-QCIF picture size and its MPI value.
MPI is integer value between 1 and 32.

QCIF:
QCIF picture size and its MPI value (1..32).

CIF:
CIF picture size and its MPI value (1..32).

CIF4:
CIF4 picture size and its MPI value (1..32).

CIF16:
CIF16 picture size and its MPI value (1..32).

Arbitrary picture sizes (XMAX,YMAX,MPI):
These parameters mean that terminal is capable and/or willing
to support arbitrary picture sizes. Both picture sizes must be divisible
by 4. These X and Y values are the maximum of allowed picture sizes with
corresponding MPI value. All three words must exist together in this
order, or none is allowed to exist.

More than one mode can exist in the same line (see example later).
At least one picture mode should be present.

2.2 Other parameters

Params = "PAR" = par_a :" par_b | "CPCF" = cpcf | "MaxBR" = maxbr | "BPP" = bpp | "HRD" = hrd

par_a = 1*DIGIT
par_b = 1*DIGIT
cpcf = 1*DIGIT "." 1*DIGIT
maxbr = 1*DIGIT
bpp = 1*DIGIT

Explanations:

Arbitrary Pixel Aspect Ratio (PAR):
Par_a and par_b are integers between 0 and 255.
Default ratio is 12:11 if not otherwise specified.

Arbitrary (Custom) Picture Clock Frequency (CPCF):
Cpcf is floating point value. Default value is 29.97.

MaxBitRate (MaxBR):
Maximum video stream bitrate, presented with units of
100 bits/s. MaxBR value is an integer between 1..19200.
BitsPerPictureMaxKb (BPP):
Maximum amount of kilobits allowed to represent a single picture frame, value is specified by largest supported picture resolution, see [1]. If this parameter is not present, then default value, that is based on the maximum supported resolution, is used. BPP is integer value between 0 and 65536.

Hypothetical Reference Decoder (HRD):
See annex B of H.263 specification.

These parameters are separated by space.

3. H.263 CODEC OPTIONS

These options are presented by the character of the annex which describes the corresponding feature.

Syntax for options is given below:

Options = "D" = "E" | "F" | "G" | "I" | "J" | "K" = "L" = "M" | "N" = "O" = "P" | "Q" | "R" | "S" | "T"

annex_d = "1" | "2"
annex_k = "1" | "2" | "3" | "4"
annex_l = "1" | "2" | "3" | "4" | "5" | "6" | "7"
annex_n = "1" | "2" | "3" | "4"
annex_o = "1" | "2" | "3"

Here #term means comma separated list of term.

3.1 H.263v1 Options

Annex D: UnRestricted motion Vector option.
D is 1 and/or 2 see later in chapter 3.2.

Annex E: Syntax based Arithmetic Coding.
It is expected that this annex is not widely used because of the complexity and IPR-issues.

Annex F: Advanced Prediction.

Annex G: PB Frames.

3.2 H.263v2 Options

Below are the new codec options which are defined in 1998 version of H.263.

Annex D: UnRestricted motion Vector.
This mode is applied differently depending whether H.263v1 or H.263v2 is used. 1 means that H.263v1 is used and annex D is applied correspondingly. 2 means that H.263v2 is used. Terminal can support both modes.
Example: D=1,2
Annex I: Advanced Intra Coding mode.

Annex J: Deblocking Filter mode

Annex K: Slice Structured mode
This mode can have up to four possible modes. It can support none, some or all of these modes. The modes are as follows:
1: slicesInOrder-NonRect
2: slicesInOrder-Rect
3: slicesNoOrder-NonRect
4: slicesNoOrder-Rect
Example: K=1,2,4

Annex L: Supplementary Enhancement Information Specification
This mode has several features.
The possibly supported features are:
1: full picture freeze
2: partial picture freeze and release
3: resizing partial picture freeze
4: full picture snapshot
5: partial picture snapshot
6: video segment tagging
7: progressive refinement
Example: L=1,6

Annex M: Improved PB-frames mode

Annex N: Reference Picture Selection mode
Four choices (modes) is available:
1: NEITHER: In which no back-channel data is returned from the decoder to the encoder,
2: ACK: In which the decoder returns only acknowledgment messages,
3: NACK: In which the decoder returns only non-acknowledgment messages, and
4: ACK+NACK: In which the decoder returns both acknowledgment and non-acknowledgment messages.
Example: N=2
For further details, see [1].

Annex O: Temporal, SNR, and Spatial Scalability mode
This annex has three possible choices (temporal, SNR, spatial). The mode is represented by integer number which refers to temporal, SNR or spatial scalability, respectively. These three submodes can be available at the same time.
Example: O=2,3

Annex P: Reference Picture Resampling
Explanation: Following submodes are available:
1: dynamicPictureResizingByFour
2: dynamicPictureResizingBySixteenthPel
3: dynamicWarpingHalfPel
4: dynamicWarpingSixteenthPel
Example: P=1,3
See [1] for further details.

Annex Q: Reduced-Resolution Update mode

Annex R: Independent Segment Decoding mode

Annex S: Alternative INTER VLC mode
Annex T: Modified Quantization mode

The presence of annex character indicates the availability of that feature.

The actual interpretation of these words is defined differently for one-way and two-way negotiations (see chapters 5 and 6).

Note that some annexes can not be used together. H.263 specification gives detailed rules for this issues and it should be followed.

4. H.263 REQUEST FEATURES IN SDP

It is sometimes useful to request intra update from other end during video session. This update might be intra picture update or GOP (Group of Blocks) update.

request= intra | gob

intra= "I-UPDATE"
Explanation: Intra picture should be sent as soon as possible. This must NOT be present in the capability exchange phase. It must be alone in the a=fmtp line when it is used (during video session).

gob= "GOB-UPDATE" ":" First "," Amount
First=1*DIGIT
Amount=1*DIGIT

Explanation: Group of Blocks are requested and they should be sent as soon as possible. "First" is the first GOB to be requested and "Amount" is the number of requested GOBs. This must NOT be present in the capability exchange phase. It must be alone in the a=fmtp line when it is used (during video session).

Example:

a=fmtp 34 GOP-UPDATE=1,3
Here first 3 GOPs are updated.

When receiving this, video encoder should send intra picture update as soon as possible. This feature is signaled only during video session, not before it.

5. USAGE OF SDP H.263 OPTIONS AND PARAMETERS WITH SIP

This document does not specify actual SIP signaling. There is no need to negotiation because it is quite needless since it is enough to tell the preferred decoder options and parameters and let the other end to decide actual options and parameters as long as the other party sends only such options and parameters which are advertised. This scheme keeps the IETF multimedia terminal simple. This syntax should be sent with SIP INVITE and corresponding status response (200 ok).

These characters exist only if the sender of this SDP message is able or willing to decode those. E.g. If a terminal is capable of decoding
SAC and AP options, it can put E F in the end of <format specific parameters>. Then the other party knows it, and can use those options in its encoder.

Picture sizes and MPI:
Supported picture sizes and their corresponding minimum picture interval (MPI) information can be combined. All picture sizes can be advertised to other party, or only some subset of it. Terminal announces only those picture sizes (with their MPIs) which it is willing to receive. For example, MPI=2 means that maximum (decodeable) picture rate per sec is about 15.

Parameter occurring first is the most preferred picture mode to be received, and last is the least preferred (but still supported) one.

These words are present in SDP line only if terminal is willing to decode the picture size with corresponding MPI. If terminal is not willing to receive some mode, it is not present in the list.

Example of the usage of these words:
CIF=4 QCIF=3 SQCIF=2 XMAX=360 YMAX=240 MPI=2

This means that sender hopes to receive CIF picture size, which it can decode at MPI=4. If that is not possible, then QCIF with MPI value 3, if that is neither possible, then SQCIF with MPI value =2. It is also allowed (but least preferred) to send arbitrary picture sizes (max 360x240) with MPI=2.

Note that most encoders support at least QCIF and CIF fixed resolutions and they are expected to be available almost in every H.263-based video application.

MaxBR and BPP parameters:
Both these parameters are useful in SIP. MaxBitRate is video decoder property, hence it differs from SDP b:<modifier> <bandwidth-value> attribute which refers more to application’s total bandwidth (an application consists often of both audio and video).

BitsPerPictureMaxKb is needed especially for decoder buffer size estimation to reduce the propability of video buffer overflow.

Also PAR, CPCF and HRD parameters can be applied in SIP.

Below is an example of H.263 SDP syntax in SIP message.

a=fmtp 34 CIF=4 QCIF=2/MaxBR=1000/E F

This means that the sender of this message can decode H.263 (RTP payload type 34) bitstream with following options and parameters: Preferred resolution is CIF (its MPI is 4), but if that is not possible then QCIF size is ok. Maximum receivable bitrate is 100 kbit/s (1000*100 bit/s) and SAC and AP options can be used.

Intra picture update can be provided during session with new SIP INVITE message with same Call-id like this:
...

m=video 9999 RTP/AVP 34
6. USE OF SDP H.263 SYNTAX AND OPTIONS WITH SAP

SAP announcements are one-way only. All H.263 options can be used to mean that sending terminal is going to use these options in its transmitted H.263 stream. It is just an informal message.

Usually only one picture size (with its MPI) exists. However, since it is possible for a video source (terminal) to change its picture size during session, several picture sizes can exist in the parameter list. First one is the original picture size to be used in the beginning of the session.

Intra requests MUST NOT be used with SAP.

Example with SAP:
a=fmtp 34 CIF=2/D=1 F

The video source is sending an H.263 bitstream and picture size is CIF, MPI=2 and URV and SAC options are used. This kind of announcement can be used e.g. in the MBone.

7. SECURITY CONSIDERATIONS

Security is truly believed to be irrelevant to this document.

8. CHANGES FROM DRAFT-01:

- Syntax is represented with augmented BNF
- H.263 option acronyms are changed to annex characters (e.g. "SAC"->"E")
- some parameter names are now shorter (e.g. BitsPerPictureMaxKb -> BPP) in order to shorten SDP description and to simplify it
- new parameters as suggested by Gary Sullivan: HRD, PAR, CPCF
- intra updates are now possible during video session (with new SIP INVITE-message)
- MaxBR is no longer mandatory plus many similar minor changes
- H.263v2 options are added. Some of these have parameters (e.g. annex O has three possible values (temporal/SNR/spatial scalability) so the syntax is e.g: O=2,3

9. OPEN QUESTIONS

The use of profiles (like in MPEG2 etc) to simplify the most common group of options.
- this seems unnecessary..
AUTHORS’ ADDRESSES

Harri Honko, Petri Koskelainen
Nokia Research Center
P.O.Box 100
FIN-33721 Tampere
Finland
e-mail: harri.honko@research.nokia.com,
petri.koskelainen@research.nokia.com,
jouni.salonen@ntc.nokia.com

References:

[1] International Telecommunication Union.
Video Coding for Low Bitrate Communication, ITU-T Recommendation H.263,
ITU 1998


specifications: ABNF," RFC 2234, Internet Engineering Task Force,