ITU Telecommunication Standardization Sector **LBC-96-302**
Study Group 15, Working Party 15/1 **November 1996**
Expert's Group on Very Low Bitrate Video Telephony

Title: Proposal on H.263+ support for flexible frame rates, frame sizes and pixel aspect ratios

Source: CLI
Purpose: Proposal
Authors: George Campbell, Sen-ching Cheung, Smita Gupta, Dan Klenke
Contact: Sen-ching Cheung (sccheung@clix.com)
325 E. Plumeria Drive
San Jose, CA 95134
Voice: +1 408 922-5489
Fax: +1 408 922-4770

---

1. Introduction

During the London LBC meeting, a key technical area of supporting more pixel aspect ratios (PARs) and frame rates was defined for consideration in the H.263+ video group **[LBC-96-264]**. We proposed a scheme with minimal modifications to the H.263+ header to accommodate all the suggestions collected during the meeting **[Cheung]**. Sullivan commented on our proposal and suggested that a more flexible scheme, which also allows maximal interoperability, should be used **[Sullivan]**. Those comments are well-received. Indeed, when a new capability is proposed, it should be as general and flexible as possible but still capable to quickly identify the best common operating mode in the presence of multiple incoherent entries. This paper provides an alternative scheme to what proposed in **[Sullivan]**. We believe this scheme to be flexible enough to achieve the best quality possible without too much burden on the implementations, but yet capable to attain maximal interoperability. To facilitate the presentations, some pseudo-H.245 code will be used to illustrate the exchange and hierarchy of capabilities.

2. Overview of Syntax Change in H.263+ and H.245

For H.245, all the additions will be in **H263VideoCapability :**

```
H263VideoCapability                    ::=SEQUENCE
{
        -- existing capabilities unchanged

        customClock                    SEQUENCE
        {
                baseClock              INTEGER (1..31), -- unit 1/599.40 Hz
                sqcifMPI               INTEGER (1..1023) OPTIONAL, -- multiples of baseClock
                qcifMPI                INTEGER (1..1023) OPTIONAL, -- multiples of baseClock
                cifMPI                 INTEGER (1..1023) OPTIONAL, -- multiples of baseClock
                cif4MPI          INTEGER (1..1023) OPTIONAL, -- multiples of baseClock
                cif16MPI               INTEGER (1..1023) OPTIONAL  -- multiples of baseClock
        } OPTIONAL,
```

```
        customFrameSize                 SEQUENCE
        {
                frameWidthBound         INTEGER (0..511) ,      -- units 4 pixels
                frameHeightBound        INTEGER (0..287),       -- units 4 pixels
        } OPTIONAL,

        pixelAspectInformation          INTEGER(1..15) OPTIONAL
        ...
}
```

For H.263+, the changes include the addition of new fields to Custom Source Format (**CSFMT**) **[LBC-96-263]** and extensions to various temporal reference fields:

Custom Source Format (CSFMT) (29 bits)
Bit 1-4 **Pixel Aspect Information: Range [1:31],**
Bit 5-13 **Frame Width Indication: Range [0,..,511]; Number of pixels per line = (FWI+1)*4,**
Bit 14 **Equal to "1" to prevent start code emulation,**
Bit 15-19 **Base Clock (BC) : "0" = 29.97 Hz, Range [1..31] in unit of 1/599.40Hz,**
Bit 20-28 **Frame Height Indication: Range [0,..,287]; Number of lines = (FHI+1)*4,**
Bit 29 **Equal to "1" to prevent start code emulation.**

Temporal Reference Extension (TRE) (5 bits)
Bit 1-4 **Higher order extension to Temporal Reference to form a 13-bit number. The unit is 1/599.40 Hz but it can only be incremented in multiples of BC/599.40 It only presents when Base Clock in Custom Source Format is non-zero.**

(Immediately after Temporal Reference for B-pictures field)
Temporal Reference Extension for B-pictures (TRE_B) (5 bits)
Bit 1-4 **Higher order extension to Temporal Reference for B-pictures to form a 8-bit number. The unit is 1/599.40Hz but it can only be incremented in multiples of BC/599.40. It only presents when Base Clock in Custom Source Format is non-zero and TR_B is present.**

The above only highlights the major changes in each standard. Detailed modifications should be worked out when there is a consensus within the committee on the proposed concepts.

3. Flexible Frame Rates

Support of flexible frame rates consists of two parts : a custom clock, **baseClock** in which the source coder is operating at and the Minimum Picture Interval (MPI) which is specified in terms of this custom clock. The custom clock is provided through a fraction (1/1 through 1/31) of a fast clock (600 000/1001 = 599.40Hz = 20×29.97Hz). The advantage of using such a fast clock is that many commonly used clock rates can be derived from or closely approximated by it. Also it allows clock rates higher than the existing 29.97 Hz. In H.263+ picture header, both the Temporal Reference and Temporal Reference for B-pictures fields will be augmented by a 5-bit field in order to use any custom clock rates. These temporal reference fields will be in the unit of 1/599.40 seconds regardless of what custom clock rate is used. However they can only be incremented in multiples of the custom clock period. Such formulation provides a roughly constant period of time before the temporal references are recycled, independent of the choice of custom clock.

To allow maximal interoperability and simple capability exchange procedure, we suggest the following rules in real-time conferencing standards :

1. If the **customClock** MPI is present, the corresponding standard MPI must also be present.
2. The **customClock** MPI must represent a real time interval no smaller than that of the corresponding standard MPI unless the value of the corresponding standard MPI is one. (i.e. **customClock** MPI ≥ standard MPI × 20.) This will ensure that the **customClock** MPI will always provide a slower maximum frame rate than the corresponding standard frame rate unless the standard frame rate already reaches its maximum.
3. When the **customClock** MPI represents the same real time interval as the standard MPI, the **customClock** MPI is preferred.
4. In a multipoint conference where all particpants support custom clocks, custom frame rates will be used only when the common clock, represented by the least common multiple of all custom clocks can be represented within the possible range. Otherwise the standard 29.97 Hz will be used instead.

4. Flexible Frame Sizes

customFrameSize represents the capability of using any frame size smaller than the frame dimension specified by ,

frameWidth × **frameHeight**,

where **frameWidth** = ((**frameWidthBound**+1)×4), and

frameHeight = ((**frameHeightBound** +1)×4).

One of the concerns in **[Sullivan]** is that casual usage of custom frame size may lead to implementations of systems whose sources are slightly simpler than the standard format with comparable quality. Even though the market will ultimately decide whether such “sub-standard” implementations will be successful, it may not be of the best interest for the standard committee to encourage such activities. On the other hand, the huge gaps in implementation complexities between higher formats like CIF and 4CIF, 4CIF and 16CIF make some forms of intermediate formats very attractive. As a compromise, we propose the following restrictions on the use of **customFrameSize** :

Let **Width** and **Height** be the dimensions of the largest standard format declared in the capset.

1. When the dimension indicated by **frameWidth** and **frameHeight** is less than **16CIF**, then the followings must be satisfied:
   - 1.1. **frameWidth** < 2 × **Width**,
   - 1.2. **frameHeight** < 2 × **Height,**
   - 1.3. **frameWidth** × **frameHeight** < 2 × **Width** × **Height**.

   As a result, custom frame size must be significantly smaller than the next larger standard format for which not to be declared. The custom format will use the MPI of the smallest standard format which satisfies the above relations. We will call this standard format the *reference format*.
2. When the dimension indicated by **frameWidth** and **frameHeight** is equal to or bigger than **16CIF**, then the MPI of **16CIF** must be declared and used for this custom format. In this case, the *reference format* is **16CIF.**

In a multipoint conference scenario, custom format can be used if all the participants support this capability and share the same *reference format*. The common custom frame size will be the intersection of the custom frame sizes from all the capsets.

5. Custom Pixel Aspect Ratios

Pixel aspect ratio (PAR) describes the ratio of the width to the height of a single pixel. For example, to display a CIF picture on a display with aspect ratio 4:3, the pixel aspect ratio is (4/3)× (288/352) = 12:11, which is the only PAR supported in H.263. This PAR is used in the PAL and SECAM television system. For NTSC television system, the use of this PAR requires resampling which always leads to degradation of video quality and increase in implementation complexity.

Historically, the CIF-type format was developed by the (then) CCITT as an international compromise between countries which use NTSC television systems (for its choice of 29.97 Hz frame rate) and those which use PAL or SECAM systems (for the PAR). However, the fact is NO known PAL/SECAM-based H.261 system converts 25 Hz input to 29.97 Hz. On the other hand, every NTSC-based H.261 does convert its 240 line/field input to 288 lines. During the London meeting, it has been suggested that H.263+ would benefit if it could support pictures with square pixels so that resampling will not be necessary when both ends of a conference use square pixels. We believe the same philosophy should be applied to NTSC systems.

Sullivan argued that the standard committee is not in favor of recommendations which recognize "regional" standards such as NTSC, PAL or SECAM **[Sullivan]**. But in fact, the most well-known ITU video standard, ITU-R BT.601 gives equal recognitions to 525-line (NTSC) and 625-line (PAL/SECAM) video systems. ITU-T H.262, developed jointly with ISO/IEC recognizes these standards as well. The inclusion of H.262 in the determined draft revision of the H.320-series specifically outlines the use of "SIF25" and "SIF30" type picture formats. Thus the failure to include recommendations for proper use of these formats will likely limit the utility of H.263+ as a standard rather than eliminate the use of equipments that uses these PARs. Hence, we are proposing the inclusion of both square PAR (1:1) and NTSC PAR (10:11) to the H.263+ header by using a single 4-bit index, **pixelAspectInformation**, to a PAR table. The table is designed in a way to allow maximal compatibility with that from the H.262 **[H.262]**. To ensure maximum interoperability, the standard recommendations should mandate that any standard format which allows non-CIF PAR must also support CIF PAR.

pixel aspect information	pixel aspect ratio
0000	forbidden
0001	1:1 (Square PAR)
0010	reserved
0011	reserved
0100	reserved
0101	12:11 (CIF PAR)
0110	10:11 (NTSC PAR)
0111	reserved
...	...

1111	reserved

Table : H.263+ Pixel Aspect Ratio Table

6. References

[Cheung] S. Cheung (CLI). *H.263+ Extended Functionality : More PAR and frame rates.* LBC Advanced Video Email Reflector. August, 96.

[LBC-96-263] H.263+ Video Group. *Altered Sections of H.263 for Draft Text of H.263+*. July, 96.

[LBC-96-264] H.263+ Video Group Chair. *H.263+ Video Group Meeting Report.* July, 96.

[H.262] ITU-T Recommendation H.262 - *Generic Coding of Moving Pictures and Associated Audio: Vido - ISO/IEC 13818-2*. 1996.

[Sullivan] G. Sullivan (PictureTel) *Re: Additional frame rates and pixel aspect ratios in H.263+*. LBC Advanced Video Email Reflector. Sept, 96