Document No. 35
Date September 1985

CCITT STUDY GROUP XV
Working Party on Visual Telephony
Specialists Group on Coding for Visual Telephony

Source : The Netherlands

Title : Simulations on source coding with a COST-sequence

1. Introduction

At the Holmdel meeting of the Specialists Group it has been decided to use a COST 211-bis sequence for simulation purposes. The sequence has been distributed in July 1985 by Mr. Nicol (BT). The covering letter gave already the most important specifications in order to achieve comparable simulation results. This document describes the tape format (as an annex 1) and the " pre-processing " of the digital data to obtain an uniform picture format. Some general information is given about the COST 211-bis project ("Redundancy Reduction Techniques for Coding of Broadband Video Signals "). A number of source coding schemes evaluated by the COST 211-bis Simulation Subgroup, will be presented in seperate documents.

2. COST 211-bis Tape-exchange Format.

The tape has been produced by BTRL in 1983 and was found to be critical. It contains a typical splitscreen scene followed, after a hard cut by a full screen sequence of a single person. There is quite a lot of movement (particulary in the so called Trevor White sequence). The signal to noise ratio is rather low. Unfortunately there are some small errors visible caused by equipment defects during the production of the original tapes.

Experience learned that the format is suitable for all digital computers.
The tape can be used provided one will mention that it has been produced on behalf of the COST 211-bis project.
Annex 1 gives a full description of the format.

3. Test sequence (see also letter Mr. Nicol)

The original images follow CCIR rec. 601

frame rate : 25
sampling frequency : 13.5 MHz (Y)
sampling frequency : 6.75 MHz (U,V)

The images on tape are of 624 pels by 576 lines.

Pre-processing in order to achieve a common reduced picture format:
Horizontally subsample : 2:1

Start with field drop : non-interlaced 25 frame rate

Frame subsample : 2:1

Resulting in a frame rate of 12.5 (for the reduced format) and 288 lines per frame.

A horizontal window of 304 pels is coded selected from the 312 original pels.

64 Kbit/s is reserved for audio and codec to codec information.

320 Kbit/s for video information (error protection and overhead included).

Demonstration bitrate= 304/360 * 320 = 270 Kbit/s

Start with splitscreen, the first frame being mid-grey.
An empty buffer is assumed for this frame.

4. Comparison of results on source coding.

To compare simulation results on source coding it is important to use:
- the same test sequence.
- the same initial conditions (state of the buffer etc.).
- The same device to display results (high quality VTR or digital tape are recommended).

It is also important to take into account:

- the number of calculations (hardware implementation)
- translatebility between bitrates.

As an annex 2: two forms used so far to compare results on coding schemes.

5. Some information about the COST 211-bis project

The Memorandum of Understanding has been signed in 1983. Ten European countries take part: Belgium (1985) , Denmark, Finland, France, Federal Republic of Germany, Italy, the Netherlands, Norway, Sweden and United Kingdom.
During the last few years a great amount of work was dedicated to the evaluation of different source coding techniques for two typical ranges of bitrates:
- 34 Mbit/s up to 140 Mbit/s for applications with broadcast picture quality
- 64 Kbit/s up to 2 Mbit/s (8) for typical videophone,videoconference and ENG applications.

The COST-group has contacts with other international organizations ie. other COST-projects, different subgroups of CEPT, CCIR an EBU.
Most of the work so far is based on the results of the Simulation Subgroup. The Subgroup has prepared proposals for candidate sampling schemes for compatible digital transmission of TV-signals at 140, 70 and 34 Mbit/s. A Hardware Subgroup started its work on this subject in 1985.

For low bitrates simulations go on as well. Most of the laboratories involved have equipment for the processing of digital video sequences (Ampex PTD, Data-Q) connected to DEC, GOULD-SEL or HP digital computers.

COST 211bis EXCHANGE FORMAT FOR PICTURE MATERIAL BY COMPUTER TAPE

1 - This format is intented for the exchange of picture material between COST 211bis countries using magnetic tapes (1600 bpi according to ISO Standard 3788). This format defines some labels for a universal use of exchanging picture material. Specific values are given for the case of exchanging picture material digitized according to CCIR Rec. 601 in Annex 1.

2 -The multi-volume file structure following the ISO 1001 Standard* will be as follows :

Single-volume file

```
VOL1 HDR1 HDR2*---File A---*EOF1 EOF2**
```

Multi-volume file

```
VOL1 HDR1 HDR2*---first section of File A---*EOV1 EOV2**
VOL1 HDR1 HDR2*---last section of File A---*EOF1 EOF2**
```

Multi-file volume

```
VOL1 HDR1 HDR2*---File A---*EOF1 EOF2 * HDR1 HDR2 *---
                                  ---File B---* EOF1 EOF2**
```

Multi-volume multi-file

```
VOL1 HDR1 HDR2*---File A---* EOF1 EOF2* HDR1 HDR2*---
                             ---first section of File B---*EOV1 EOV2**
VOL1 HDR1 HDR2*---intermediate section of File B---
                                  *EOV1 EOV2**
VOL1 HDR1 HDR2*---last section of File B---*EOF1 EOF2*HDR1 HDR2*---
                                  ---File C---*EOF1 EOF2**
```

with :

VOL1 - Volume Header label (system label VOL1)

HDR1 - First file Header label (system label HDR1)

HDR2 - Second file Header label (system label HDR2)

---

* - Tape mark

EOV1 - First End of Volume

EOV2 - Second End of Volume

EOF1 - First End of File

EOF2 - Second End of File

3 - The labels should correspond to ISO 1001 with some restrictions :
- two system header labels and no user label (see Annex A of ISO 1001).
- only ASCII-characters should be used for the labels.
- in some cases, a "restricted set of characters" should be used, being defined as only (ASCII) upper-case letters (A-Z) and numbers (0-9) with a letter as first character.

4 - A description of the labels can be found in the Annex 2 of this document (ISO 1001 Standard). Only restrictions to this Standard will be listed below.

a) Volume header label (system label VOL1)

CP	Field name	L	Content
5-10	Volume identifier	6	restricted character set (content should reflect the origine of the tape : e.g. "CCETT").
11	Accessibility	1	must be a space, to indicate that there are no restrictions
51	Digital standard version	1	numeric character 1

with CP = character position
L(ength) = number of ASCII-characters

b) first file header label (system label HDR1)

CP	Field name	L	Content
5-21	File identifier	17	characters at CP5-10 should be taken from the restricted set.
22-27	File set identifier	6	must be identical to the volume identifier of VOL1
28-31	File section number	4	numeric characters starting at 0001 and incrementing by 1 for each additional volume, with respect to the first volume on which the file begins.
32-35	File sequence number	4	file number within a file set (single or multi-volume);consists of numeric characters, starting at 0001 that indicates the position of this file with respect to the first file of the set.
48-53	Expiration date	6	must be a space followed by "00000" (five zeros) to indicate no significant expiration date
54	Accessibility	1	must be a space (no restriction).

c) **second file header** (system label HDR2)

CP	Field name	L	Content
5	Record format	1	"F" (=fixed length) should be used
6-10	Block length	5	at max, 4096 bytes at min, 320 bytes must be a multiple of 4 Five numeric characters specifying the max. number of characters/block.
11-15	Record length	5	one record per block should be used (CP11-15 content = CP 6-10 content). Numeric characters indicating the record length
51-52	Buffer offset	2	must be "00" (no offset)

5. The EOF1-EOF2 labels and the EOV1-EOV2 labels are described in the Annex 2 (same restrictions as for HDR1 and HDR2).

6. The first data record should contain relevant information describing the picture material (only ASCII characters being used). Its structure is as follows :

FIRST DATA RECORD

CP	Field name	Length	Content
1 - 16	Picture name	16	Characters at CP1-6 should be taken from the restricted set
17 - 20	Number of recorded lines per field NL1	4	All these number should be integers and they relate to the first component of the picture material (see CP 33-34)
21 - 24	Number of recorded pixels per line NP1	4	
25 - 28	Number of lines per record	4	
29 - 32	Number of frames	4	(See note 7)
33 - 34	Nature of the picture	2	"BW" : one component (B/W) "CC" : three component (RGB) "CD" : three component (YC_BC_R) "Dx" : four components (document prints) ; x free "Px" : composite PAL (x="M" for PAL/M) "Sx" : composite SECAM (x="L" for SECAM/L) "Nx" : composite NTSC (Same)
35	Movie/video	1	"M" : movie ; "V" video
36 - 39	Field rate	4	Numerical value identical for all components.
40	Interlace ratio	1	

First data record (continued)

CP	Field name	Length	Content
41 - 48	Sampling frequency	8	sampling frequency of the first component (kHz)
49	Horizontal sampling ratio 1 (S1)	1	first component (see note 1)
50	Horizontal sampling ratio 2 (S2)	1	second " "
51	Horizontal sampling ratio 3 (S3)	1	third " "
52	Horizontal sampling ratio 4 (S4)		fourth " "
53	Line subsampling ratio 1 (L1)	1	first component (see note 2)
54	Line subsampling ratio 2 (L2)	1	second " "
55	Line subsampling ratio 3 (L3)		third " "
56	Line subsampling ratio 4 (L4)	1	fourth " "
57	Sampling pattern 1(P1)	1	first component (see note 3)
58	" " 2(P2)	1	second " "
59	" " 3(P3)	1	third " "
60	" " 4(P4)	1	fourth " "
61	Horizontal sampling phase 1 (H1)	1	first component (see note 4)
62	" sampling phase 2 (H2)	1	second " "
63	" sampling phase 3 (H3)	1	third " "
64	" sampling phase 1 (H4)	1	fourth " "
65	Vertical sampling phase 1 (V1)	1	first component (see note 5)
66	" sampling phase 2 (V2)	1	second " "
67	" sampling phase 3 (V3)	1	third " "
68	" sampling phase 4 (V4)	1	fourth " "
69 - 160	(reserved)	92	spaces (until other specifications are issued)
161 - RL	(comments) . . .	-	See Note 6. After comments, null characters should be used till the end of the record. RL is the number of bytes of the record (see CP11-15 of HDR2).

Note 1 : Integers S1, S2, S3, S4 should be used (values 0-9), the effective sampling frequency Fi of component i being as follows : Fi = (Si/S1)* F1 (F1 is given at CP 41-48).

Note 2 : Subsampling ratio relative to line frequency. Integers L1, L2, L3, L4 should be used (values 0-9), the effective number of line per field NLi of component i being as follows : NLi = (1/Li)*NL1 (NL1 is given at CP 17-20).

Note 3 : "Ø" = orthogonal ; "L" = Line quincunx ; "F" = field quincunx.

Note 4 : Integers H1, H2, H3, H4 should be used (values 0-9) indicating the distance between the first pixel of the recorded line (component 1) and the first pixel of component i. The unit is defined in number of pixels (the sampling frequency is given at CP41-48). This is a necessary information when horizontal subsampling is used for component i.

Note 5 : Integers V1, V2, V3, V4 should be used (values 0-9) indicating the distance in number of lines between the first line of the recorded field (component 1) and the first line of component i. This is a necessary information when line subsampling is used for component i.

Note 6 : Comments should give all complementary informations on the picture material e. g. number of line per frame of the original signal, aspect ratio, ... The use of a not allowed character in CP 49-68 indicates that the corresponding paramater is described in the comments.

Note 7 : The number of recorded pixels per line NP1 must be choosed in such a way that, for multicomponents signal, all the NPi=NP1*Si/S1 are integer numbers. In the same way, NL1 must be choosed in order to ensure that all the NLi=NL1*L1/Li are integer numbers.

7. The actual data starts from the second data record.

8. Each frame is described by the following successive "fields" :

Frame k	field 1 component 1	
Field 1 of the frame	field 1 component 2 field 1 component 3 field 1 component 4	if required
	. . .	
	field j component 1	
Field j of the frame	field j component 2 field j component 3 field j component 4	if required
	. . .	
	field n component 1	
Field n of the frame	field n component 2 field n component 3 field n component 4	if required

The value n is defined at CP 40 of the first data record. Within one field, each component is scanned separately and line by line.

Within one frame, field j is above field j+1, which means that line 1 of field j is above line 1 of field j+1.

Each record contains an integer number of lines, all related to the same field and the same component. For the first component, the value R1 is given at CP 25-28 of the first data record. For the other components, each record contains $Ri = \left[\frac{S1}{Si}\right] R1$ lines where $[x]$ means "integer part "of x" (the Si are given at CP 49-52 of the first data record). If necessary, the end of the record is filled with zeros. If the last record of field j component i contains less than Ri lines, this record is filled by lines of zeros.

ANNEX 1 – Examples of parameters values describing the picture material.

This Annex is intended to give the values of the parameters describing the picture material when digitized using the CCIR 4:2:2 standard from Rec. 601 (luminance signal sampled at 13.5 MHZ, both color difference signals at 6.75 MHz ; 8 bits per sample) in the case of 625 l. 50 Hz standards.

Two sets of parameters will be given, one corresponding to the full active format (288 lines of 720 samples per field), the other to the slightly reduced format (268 lines of 675 samples per field) that is used by the ARTISTE system from CCETT, due to the capacity of the AMPEX DMPTD disks which is slightly below than full format one. Only the relevant parameters are described below.

CP	Field name	Content for CCIR Rec. 601 digitalization	
		full format	ARTISTE format
Second file header label (HDR2)			
6-10	Block length	"␣3600"	"␣4044"
11-15	Record length	"␣3600"	"␣4044"
First and second end of file labels (EOF1, EOF2)			
	same as for HDR1 and HDR2		
First and second end of volume labels (EOV1, EOV2)			
	same as for HDR1 and HDR2		
First data record			
17-20	number of recorded lines per field	"␣288"	"␣268"
21-24	number of recorded pixels per line	"␣720"	"␣674"
25-28	number of lines per record	"␣␣␣5"	"␣␣␣6"
33-34	nature of the picture	"CD"	"CD"
36-39	field rate	"␣␣50"	"␣␣50"
40	interlace ratio	"2"	"2"
41-48	Sampling frequency	"␣␣␣13500"	"␣␣␣13500"

49	S1	"4"	"4"
50	S2	"2"	"2"
51	S3	"2"	"2"
52	S4	" "	" "
53	L1	"1"	"1"
54	L2	"1"	"1"
55	L3	"1"	"1"
56	L4	" "	" "
57	P1	"Ø"	"Ø"
58	P2	"Ø"	"Ø"
59	P3	"Ø"	"Ø"
60	P4	" "	" "
61	H1	"0"	"0"
62	H2	"0"	"0"
63	H3	"0"	"0"
64	H4	" "	" "
65	V1	"0"	"0"
66	V2	"0"	"0"
67	V3	"0"	"0"
68	V4	" "	" "

P.S. : . For the Rec. 601 full format, we get :
R1 = 5 (5 lines of component Y per record)
R2 = 10 (10 " " " C_B " ")
R3 = 10 (10 " " " C_R " ")
. For the Rec. 601 ARTISTE format, we get :
R1 = 6 (6 lines of component Y per record)
R2 = 12 (12 " " " C_B " ")
R3 = 12 (12 " " " C_R " ")

COMPARISONS OF CODING METHODS BASED ON:
==========

Picture quality:

Ranking, but with comments about major impairments.
(Pre/post processing used).

Picture quality if for example 5-10% of the bitrate must be used for other purposes.

Can the picture quality be significantly better by using new "tricks" ?

Implementation:

Ranking with comments on major drawbacks.

Number of calculations in critical parts.

General structure (parameter list).

Translatability:

Between bitrates, concepts.

Network implications:

New or old:

Is the method well understood ?
Possible to standardise in a few years time?

CODING SCHEME:

A: NEVER
B: SELDOM
C: SOMETIMES
D: OFTEN
E: ALWAYS

IMPAIRMENTS	RATING	comments
1) blocking		
2) granular noise		
3) slope overload		
4) false contours		
5) edge business		
6) spatial blur		
7) temporal blur		
8) dirty window		
9) recovery time		
10) motion business (jerkiness)		
11) amplitude reproduction .		