

SOURCE: CHAIRMAN OF THE SPECIALISTS GROUP ON CODING FOR VISUAL TELEPHONY
TITLE : REPORT OF THE SIXTH MEETING IN MONTREAL (JUNE 17-20, 1986)

CONTENTS

1. General
2. Document for the meeting
3. Source format
4. Transformer
5. Source coding algorithm
6. Video multiplex coding
7. Transmission coding
8. Multipoint considerations
9. Flexible prototype hardware specifications
10. Intellectual property
11. Future work plan

1. General

The Specialists Group met in Montreal from June 17 to 20, 1986, at the kind invitation of BNR, Canada. Welcoming address was delivered on behalf of the hosting organization by Mr. Prasada.

The list of participants appears at the end of this report.

At the final session, Chairman thanked the inviting organization for the meeting facilities provided and the excellent organization.

2. Documents for the Meeting

2.1 Normal Documents

#103R: REPORT OF THE FIFTH MEETING IN TOKYO (CHAIRMAN)

Points agreed upon and/or for further study are recorded on all the aspects in order to provide backgrounds for the discussion in this meeting.

#104 : REVISED SPECIFICATIONS FOR THE REFERENCE MODEL (CHAIRMAN)

Revision was proposed to the specifications agreed in Tokyo (Annex 2 to Document #103R) based on the preliminary works in Europe and Japan.

#105 : NUMBER OF CODED PELS PER LINE (NTT, KDD, NEC, FUJITSU)

Reduction of coded pels per line to 352 for luminance and 176 for chrominance is supported, since it discards only 1% or less of active picture area. It is proposed that the parameter should be defined as a part of source coding algorithm, not as modification of CIF.

#106 : PROGRAMMABLE TRANSFORMER (NTT, KDD, NEC, FUJITSU)

Descriptions of programmable transformer and specific parameters for DCT are presented. Operation order, accuracy of operation (12 bits) and truncation methods at each stage are proposed with respect to compatibility.

#107 : REFERENCE MODEL SIMULATION (NTT, KDD, NEC, FUJITSU)

Though the reference model was first implemented according to Document #104, reproduced pictures were not good. Hence, improvements were made so that the reference model gives reasonable quality for algorithm comparison work. Such elements as block attribute coding, quantizer control, pre-filter and filter in the coding loop were improved in four steps to clarify the effectiveness of each element with numerical data and processed pictures.

#108 : PROPOSAL ON CODING ALGORITHM (NTT, KDD, NEC, FUJITSU)

From the simulation results and hardware considerations, the quantization/coding strategy consisting of the following elements are proposed for the flexible hardware.

- a. Classification operated on unquantized transform coefficients.
- b. Transmitted coefficients are determined as a function of class and assigned distortion.
- c. Different quantizers for DC and AC components.
- d. Multiple VLC code sets for quantizer index.

Coding control in such large unit as picture is also proposed. Comments on other aspects are given.

#109 : DISCARDING OF SIGNIFICANT COEFFICIENTS (NTT, KDD, NEC, FUJITSU)

Use of human-visual perception for adaptive quantization/coding was suggested. Three such methods were experimented on the reference model; isolated coefficients discarding, motion vector control discarding, and adaptive quantization. Processed pictures are presented with SNR and information rate data.

#110 : FILTER INSIDE THE CODING LOOP (NTT, KDD, NEC, FUJITSU)

Shaping filter in the local decoder loop is experimented on the reference model. Since adaptively controlled shaping filter gives improvements in terms of SNR and picture quality, its inclusion in the prototype hardware is proposed. Filter type and adaptive control method are for further study.

#111 : SOME OBSERVATIONS ON CODING CONTROL (NTT, KDD, NEC, FUJITSU)

Taking note of the variable frame dropping capability required for the $n \times 384$ kbit/s codec, it is proposed that distortion control should be carried out in unit of 1 frame and supplementary control in each block should be based on classification.

#112 : MOTION COMPENSATION FOR CHROMINANCE SIGNALS (NTT, KDD, NEC, FUJITSU)

The three methods described in Section 9/Document #103R,

- derived from luminance
- none
- independent

are being compared. Since the performance depends on image contents, the second method (no MC) is preferred because of its simplicity in hardware.

#113 : VIDEO MULTIPLEX (NTT, KDD, NEC, FUJITSU)

Data structure in video multiplex coding and measures against transmission error are discussed to conclude that;

- a. Block attribute data should be encoded in collective forms as a part of GOB data from ease of VLC design and hardware simplicity.
- b. From the analysis of transmission error effects, the error correction approach is favored because error resilience considerations become complicated in the scheme using interframe difference and motion compensation.
- c. Plural VLC code sets can be employed without increasing vulnerability to channel error.
- d. To cope with remaining errors after error correction, video data should be refreshed periodically and/or on demand.

#114 : CODING OF MC VECTORS (NTT, KDD, NEC, FUJITSU)

Two motion vector encoding methods, absolute and block-difference, are compared with respect to error resilience, coding efficiency and hardware complexity. Integral and fractional vectors are also compared. It is proposed that motion vectors with integer accuracy be encoded in the form of difference between two successive vectors.

#115 : COMMENTS ON DIGITAL INTERFACE FOR $n \times 384$ kbit/s CODEC (NTT, KDD, NEC, FUJITSU)

Document #77 by FRG has been reviewed for the 1544 kbit/s interface. The interface and frame structure should conform to the existing recommendations. The following two points need considerations.

- a. Allocation of H0 channels in the 1544 kbit/s interface:
Example in Annex A/I.431 is supported.
- b. Synchronous/asynchronous indication bit: Provision in the H0 channel will be favored because of the capacity lack in 1544 kbit/s frame structure.

#116 : REVISION OF THE SPECIFICATION FOR THE FLEXIBLE PROTOTYPE HARDWARE (NTT, KDD, NEC, FUJITSU)

As a summary of the contributions #105 - #115, modification and addition for Annex 4 to Document #103R are proposed.

#117 : ONES DENSITY RESTRICTION AS RELATED TO $n \times 384$ kb/s VIDEO CODECS (USA)

It is suggested that the Specialists Group consider alternatives to allow operation of the $n \times 384$ kbit/s codec on a restricted channel existing in North America. The restriction is a '1' in each 8 bit channel byte.

#118 : MULTIPOINT CONFERENCING WITH $n \times 384$ kbit/s CODECS (USA)

The following provisions should be considered in the $n \times 384$ kbit/s codec specification.

- a. The codec must operate in a simplex mode.
- b. The codec should operate in a fast update mode when the communication channel is switched.
- c. The encryption must be self synchronizing.
- d. An option to remove audio from the encryption and FEC fields may be provided.

#119 : REDUCED CHROMINANCE RESOLUTION (FRANCE, FRG, ITALY, NETHERLANDS, SWEDEN, UK)

Subsampling for chrominance signals in the pel domain is not supported from hardware and coding efficiency considerations.

#120 : A HARDWARE ASPECT OF MULTI-PEL PREDICTORS (UK)

Considering the hardware impact and simulation results, it is proposed that the hardware specification should initially be for a single pel access picture store, but with the proviso that this will be changed later if required.

#121 : ACCURACY OF TRANSFORM ARITHMETIC (UK, SWEDEN, NETHERLANDS)

From the simulation results, the followings are proposed to ensure both absolute and relative accuracies.

- a. The matrix multiplier implementation of the inverse transform should use TDC1010 or functionally equivalent devices.
- b. The hardware should allow matrix coefficients to be stored to 16 bits.
- c. The most significant 16 bits by truncation should be carried between the two 1-D transforms.
- d. The output of the inverse transform should be rounded to 9 bits giving pel domain values in the range -256 to 255.

#122 : VLC SET PROPOSAL FOR $n \times 384$ kbit/s (UK, SWEDEN, FRANCE)

Considering the statistical data obtained through the reference model simulation on Split-Trevor and Miss America sequences, the following proposals are made.

- a. Initially the code set in Table 1 should be adopted for all variable length coded data in the codec.
- b. A GBSC should consist of 16 zeros followed by Line of Block Number and Type information.
- c. A PSC should be as the GBSC above except with the Line of Block Number set to zero.
- d. The detail of the VLC set should be programmable for all code words within the constraint of a maximum code word length is 16.

#123 : SPECIFICATION FOR THE FLEXIBLE PROTOTYPE n x 384 kbit/s
VIDEO CODEC (FRANCE, FRG, ITALY, NETHERLANDS, SWEDEN, UK)

As a result of further studies in Europe, amendments and extensions for Annex 4 to Document #103R are proposed.

#124 : INTELLECTUAL, PROPERTY CONSIDERATIONS (BT, GEC)

Examination of the reference model and applicable patents are reported. Since a number of patents appear to cover some fundamental parts of the reference model, it is proposed that the holders of such patents be formally asked to clarify their position on any patents held, and to make a clear policy statement as to whether non-discriminatory licenses will be made available. It is further proposed that they should be asked to state precisely the terms of such licenses (royalty fees etc.) by 1st October 1986.

#125 : A BLOCK STRUCTURE FOR THE COMMON INTERMEDIATE FORMAT (FRG, UK)

After discussing various approaches to get an integer number of blocks for R-Y and B-Y signals, omission of 8 luminance pels and 4 chrominance pels is proposed with a treatment of those discarded pels.

#126 : A HYBRID CODER WITH EQUIVALENT FILTERS (FRG, NETHERLANDS)

An improvement of picture quality by using filters in the local decoder is presented.

#127 : FRACTIONAL MOTION VECTORS FOR MOVEMENT COMPENSATING PREDICTION IN
HYBRID CODERS (FRG, NETHERLANDS)

A study on fractional motion vectors is reported.

- a. A vector finer than 0.5 provides only a slight gain.
- b. The encountered savings in bit rate are approximately 6%.
- c. The picture quality is improved slightly compared to the reference modes.

Since the hardware of a codec with fractional motion vectors is more complex than a scheme with integer vectors, it has to be carefully considered whether or not this is payed back by a significant better picture quality.

#128 : RESULTS ON REFERENCE SIMULATION (SWEDEN)

The reference model has been implemented. The picture quality was poor, but better than expected. Some statistical data are presented.

#129 : MOTION COMPENSATED PREDICTION (SWEDEN)

The following three motion compensation strategies are compared in simulation.

- a. Integer pel displacement without filtering
- b. Integer pel displacement with filtering
- c. Fractional pel displacement

Simulation results show a significant picture quality improvement by using strategy b.

#130 : IMPROVEMENTS ON THE REFERENCES MODEL (SWEDEN)

A simulation with the sequence Split-Trevor was carried out to find that;

- a. A significant improvement was achieved by introducing multi-predictor for non-zero motion vector block.
- b. A small improvement was achieved by using the motion vectors estimated for the luminance for prediction of chrominance.
- c. A small improvement was achieved by introducing adaptive scanning with three classes; zig-zag, horizontal and vertical.
- d. Some improvement was gained by changing the quantizer step size only after each GOB.

#131 : MOTION VECTORS FOR THE CHROMINANCE (NETHERLANDS, SWEDEN)

Application of chrominance vectors derived from luminance vectors show a little improvement (1-3% bits saving) but hardly visible picture quality improvement. However, its use in the flexible hardware is proposed for future considerations.

#132 : COMMENT REFERENCE MODEL SPECIFICATIONS -SCENE CUT DETECTION-(NETHERLANDS, BTRL)

The three scene cut detection methods are discussed; the previous frame difference, the frame difference and the criterion used in the reference model. Since simulations show that the operation of the scene cuts can significantly influence the subjective quality of the results, and the scene cut is an unusual case, a priori knowledge use for scene cuts is proposed for simulation.

#133 : STATISTICS ON THE REFERENCE MODEL (NETHERLANDS, FRANCE)

The following results are presented.

- a. Occurrence frequency of each Block Type
- b. Bit consumption
- c. Occurrence frequency of each quantizer index
- d. Distribution of number of non-zero coefficients

#134 : EXAMPLE STATISTICS ON THE IMPROVED REFERENCE MODEL (FRANCE, NETHERLANDS, FRG)

Occurrence frequency of each Block Type and bit consumption are given for the following three improvements.

- a. Quantizer step transmitted as a GOB attribute and use of triangular VLC.
- b. Improvement of a. + intra mode only for the start and the

scene cut

c. Non-linear quantizer and filter in the loop

#135 : SOME CONSIDERATIONS ON MULTIPOINT VIDEOCONFERENCING (FRANCE)

The influence of the MCU definition and functions on the coded design are considered concerning the following aspects.

- a. MCU functions
- b. Clock synchronism
- c. Video switching
- d. Data continuity
- e. Encryption

#136 : ACTION ON $m \times 64$ kbit/s IN SGXV AND SPECIALISTS GROUP (BT)

It is proposed that the Specialists Group should actively begin studies on $m \times 64$ kbit/s codec with the aim of encouraging SGXV to adopt at least a structure for that codec and produce a recommendation in 1988 alongside that for $n \times 384$ kbit/s.

#137 : CLASSIFICATION TECHNIQUE (FRANCE, SWEDEN, ITALY)

A provision of 8 scanning classes is suggested for flexible hardware to determine the shortest way reaching the non-zero coefficients in the transformed block. It implies 3072 bits PROM and some comparison circuits and may include zonal techniques.

#138 : STATISTICS RELATED TO THE USE OF THE 'DCT' TRANSFORM IN THE SIMULATION OF THE 'OKUBO' REFERENCE MODEL (BT)

#139 : STATISTICS RELATED TO THE USE OF THE 'HAAR' TRANSFORM IN THE SIMULATION OF THE 'OKUBO' REFERENCE MODEL

Statistics of Split-Trevor sequence such as bits used per frame, buffer status, block type, number of transmitted coefficients are presented for information.

2.2 Temporary Document

- No.1 Agenda (Chairman)
- No.2 Available Documents and Discussion Points (Chairman)
- No.3 Major Differences in the Two Proposals for Flexible Prototype Hardware Specification (Chairman)
- No.4 Proposal on Motion Vectors (France)
- No.5 Block Type (Sweden)
- No.6 Proposal for the Hardware Specification (France)
- No.7 Draft Report of the Sixth Meeting in Montreal (Chairman)
- No.8 Action Points for Simulation Work (Small Group on Simulation)
- No.9 Specification for the Flexible Prototype $n \times 384$ kbit/s Video Codec (Small Group on Hardware Specification)
- No.10 A Family of VLC's for 384 kbit/s CODEC (France)
- No.11 Intellectual Property (Chairman)

3. Source Format

3.1 Number of Coded Pels per Line (#105, #125)

Considering that the number of chrominance samples (180 in CIF) can not be divided by 8, which is the number describing the block size, and the impact of removing 8 luminance pels and 4 chrominance pels from CIF on the picture, the meeting agreed on the following:

- 1) The number of coded pels per line:
Y : 352
C_R: 176
C_B^R: 176
- 2) This reduction is defined as a part of the source coding algorithm.
- 3) Position of these coded pels for luminance in the video signal is depicted in Figure 1. The pels denoted by X are associated to block level at the decoder.

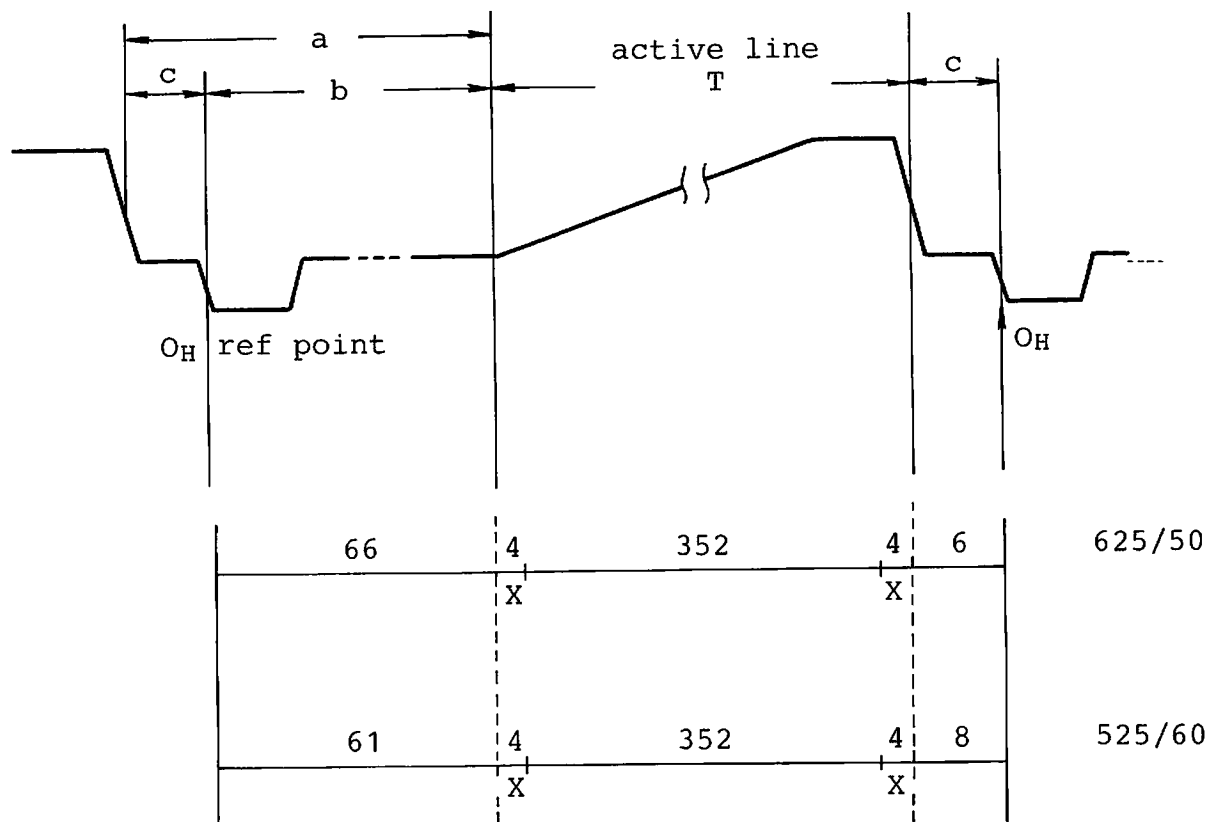


Figure 1

3.2 Subsampling of Chrominance Signals (#119)

According to the proposal made in Document #119, it was agreed that subsampling in the pel domain is not used. When further reduction of chrominance information, spatial pre-filtering or transform domain filtering may be applied.

4. Transformer (#106, #121)

Description of programmable transformer specification was discussed concerning the operation order, accuracy of operation and truncation methods. The following conclusion was obtained.

- 1) To guarantee hardware compatibility, it is necessary to specify the hardware structure and the number of bits at each stage with the rules for discarding bits.
- 2) The hardware should allow matrix coefficients to be stored to 16 bits.
- 3) The most significant 16 bits by truncation should be carried between the two 1-D transforms.
- 4) The output of the inverse transform should be rounded to 9 bits giving pel domain values in the range -256 to 255.

The details are given in 1.2.4/Annex 4 to this report.

5. Source Coding Algorithm

Reference was made to #104, #107-112, #120, #122, #126-134, #137-139. Along with the presentation of the documents, various demonstrations of simulation results were also given. An overview of the tape demonstrations is listed here.

- a. Improved reference model and zonal technique (Japan; #107, #108)
- b. Conditional discarding of significant coefficients and filter in the loop (Japan; #109, #110)
- c. Coding control (Japan; #111)
- d. Motion compensation for chrominance (Japan; #114)
- e. Filter in the loop and motion compensation for fractional displacement (FRG; #126, #127)
- f. Improvements on the reference model (Sweden; #130)
- g. Improvements on the reference model (Netherlands; #131, #133)
- h. Improvements on the reference model (France; #134)
- i. DCT and Haar transform in reference model (UK; #138, #139)

5.1 Reference Model Simulation Results (#104, #107, #128, #130, #133-134, #138-139)

The reference model was implemented in several laboratories and numerical data were presented. Possible causes for slight differences among documents need be analyzed by reviewing these results. A necessity of common format to present the numerical results was pointed out.

It was felt that the reference model specified in Document #104 is not giving desirable performance. The action is described in 2/Annex 1 to this report.

5.2 Scene Cut (#132)

Though methods of scene cut are not subject to compatibility, practical ones await further study. Since the scene cut processing affects the picture quality to a considerable extent in such short time of simulation test sequences, it was agreed to use a priori knowledge method for future simulation work.

Use of optional camera switching signal from the videoconferencing terminal equipment was recognized as convenient for the encoder operation. This point, however, is out of this Specialists Group responsibility and should be forwarded to WPXV/1.

5.3 Quantizer (4.3/#108)

There was a comment that the quantizer for AC component of intra mode should better be with dead zone = 0.5g in Fig.3(b)/Document #108.

5.4 Coding Strategy (#108, #109, #122, #130, #137)

The means to efficiently encode the transform coefficients were discussed in the following aspects; determination of coefficients to be transmitted, classification, VLC code sets and coding table.

The meeting got a common understanding that this coding strategy is the field where flexibility be left so that the results of the future study can be incorporated.

To get a compromise solution between difference proposals in Documents #108 and #122+#137, Mr. Temime undertook a coordination. The results are reflected in 1.2.5, 1.2.7 and 2.2.5 of Annex 4 to this report.

- 1) To determine the area of transmitted coefficients, the method proposed in Document #137 covers the method proposed in Document #108. Thus, it is agreed to use the former method for flexible hardware specification.
- 2) In order to efficiently encode the non-zero coefficients, the use of multiple VLC code sets were proposed in Document #108. After extensive discussion on coding efficiency, error resilience and hardware implication, the meeting agreed on the use of multiple VLCs in hardware specification on 'future inclusion' basis. Namely, the initial hardware is constructed using one VLC, but it must be easily extended to 4 VLCs.
- 3) As for the number of VLC code sets for different information, it was recognized that the following five information should be variable length coded.
 - Quantizer index
 - Motion Vector
 - Block address
 - Block type
 - EOB/class

- 4) EOB code is one of the VLC code words used for quantizer index.
- 5) Related to the use of multiple VLCs, the error protection philosophy was discussed. Some members stressed that error resilience should be maintained to confine the effect of transmission errors in the block without relying on FEC. The other members were of the opinion that the error resilience design becomes complicated in motion compensated interframe coding environments, and the hardware for FEC is not a burden, thus FEC use is practical. The meeting, however, recognized the desirability of error resilience if it is easily designed without hurting coding efficiency.

5.5 Quantizer Control (#111)

Document #111 suggested that large unit basis quantizer control such as frame basis one is necessary to improve the picture quality. It was confirmed that this is possible using Picture Attribute or GOB Attribute, and even in that case, the block basis control is still necessary for forced update.

The coding control methods to cope with the following situations were raised as problem;

- Scene cut where two times or more bits per frame are spent compared to steady state frame
- Variable frame dropping rate

As a possible method, the use of input buffer before coding loop and display buffer after decoding was suggested in 1.3/Document #116. Relation to the transmission buffer and other possible methods need be clarified.

5.6 Motion Compensation

5.6.1 Integer vs fractional vectors (#114, #120, #127, #129)

After discussion of coding efficiency and hardware implication, it was agreed that fractional vectors be used on 'future inclusion' basis.

5.6.2 Motion vector for Chrominance (#112, #130, #131)

After comparison of no use and use of motion vectors derived from luminance vectors in respect to coding efficiency and hardware implication, the meeting agreed to apply 'future inclusion' principle to this problem.

5.6.3 Restriction of Motion Vector Range at the Picture Boundary

To avoid the motion vector point out the pels outside the picture, it was agreed that predictor must not use pels outside the 352 x 288 picture.

5.7 Filter in the Coding Loop (#110, #126, #130)

Two possibilities of inserting filter in the coding loop were proposed; in front of the frame store (#110) and after the frame store (#126). Taking into account the fact that this filter is likely to improve the picture quality and also taking into account hardware implication, the

meeting decided that the hardware should allow future inclusion of either method.

5.8 Action Plan

Small meeting was coordinated by Mr. Guichard to set an action plan of the simulation work toward the next meeting. The results are attached as Annex 1 to this report.

6. Video Multiplex Coding

6.1 Block Attribute and Data Arrangement (#113)

The following two schemes were compared in terms of schemes of hardware simplicity, correlation usage, error resilience.

Scheme 1: Block basis arrangement

Scheme 2: GOB basis arrangement

After confirming scheme 1 is easily implementable, the meeting agreed on adopting scheme 1 for the block attribute and data arrangement.

6.2 Motion Vector Coding (#114)

After a comparison of absolute vector transmission (5.8/#103R) and difference vector transmission (#114), the following agreements were obtained.

- 1) Motion vectors are transmitted differentially, with respect to the MV of the previous block. The first block MV is related to AMV: 0, 0. The hardware should allow the Global MV in the GOB attribute to be also the basis of differential coding of MV.
- 2) Differential Motion Vectors are coded using a VLC (two entries).
- 3) Maximum length of VLC for differential motion vectors is 16.

There are two alternatives for zero motion vector coding. One is to code it as Block Attribute, the other is as Vector Data. An example of Block Attribute code set is shown in Annex 2 to this report. The choice is for further study.

7. Transmission Coding

7.1 Time Slot Assignment for H0 Channels and Synchronous/asynchronous Indication (#115)

The proposals in Document #115 were supported by the meeting. Incidentally, the suggestion described in 3/#115 was taken note of as a necessary measure for the compatibility checks at laboratories.

7.2 One's Density Restriction (#117)

The meeting recognized the impact of one's density restriction in North America, particularly the stricter restriction for H0 channel, and the necessity of consideration. USA was requested to propose some cures at

the next meeting taking into account the frame structure already agreed for the n x 384 kbit/s codec.

8. Multipoint Considerations (#118, #135)

Document #118, together with #135, was recognized as useful input for the Specialists Group's work. The extract is attached as Annex 3 to this report.

France was requested to provide a proposal for the Application Channel code assignment including 'Freeze Picture Request' and 'Fast Update Request'.

9. Flexible Prototype Hardware Specification (#116, #121)

As a conclusion of this meeting, Annex 4 to Document #103R was revised as Annex 4 to this report. This drafting task was coordinated by Mr. Morrison.

Several items raised in Document #116, such as optional demand refresh function, spare attribute bits, byte structure for Picture and GOB headers, could not be discussed due to lack of meeting time and also lack of time required for consideration by other members.

10. Intellectual Property (#124)

It was agreed to attempt to generate a list of known patents which are covered by the Hardware Specification. Each core member would search in his own country and endeavour to obtain a statement from holders of those patents considered relevant, before the next meeting. These statements would be in response to a 3 stage question (Note) based on the fifth paragraph of Document #124. This exercise would be ongoing as and when changes are made to the Hardware Specification.

- Note:
1. Do any of your patents cover the Flexible Hardware Specification?
 2. If yes, will licenses be made available on a non-discriminatory basis?
 3. If yes to 2, what will be the precise terms of such license (royalty fees, initial payments etc.)?

11. Future Work Plan

11.1 Correspondence Procedures

Though the discussion results of this meeting are reflected in the revision of the flexible prototype hardware specification (see Annex 4 to this report), further amendments or clarifications to this specification may be required in the course of study toward the next meeting. These should be agreed through correspondence to enable the earliest possible hardware provision. The meeting agreed on the following correspondence procedures.

- 1) Any formal proposal to amend or clarify the hardware specification is

first sent to Chairman.

- 2) Chairman distributes the received proposal to all the core members with a notice of dead line for comments.
- 3) If the proposal is agreed, Chairman issues an appropriate document if necessary.
- 4) Decisions made through correspondence are confirmed at the next meeting.

11.2 Study on m x 64 kbit/s CODEC (#136)

After considering the proposals contained in Document #136, the meeting agreed on the following points concerning the m x 64 kbit/s codec study.

- 1) The Specialists Group will start discussion on the m x 64 kbit/s codec at the next meeting.
- 2) Priority for the Specialists Group activity is placed on the n x 384 kbit/s codec.
- 3) As a guideline, half a day will be spent for the m x 64 kbit/s codec at the next meeting.
- 4) Such general aspects as;
 - application
 - picture format
 - compatibility with n x 384 kbit/s codec
 - bit rate
 - frame structure

will be discussed at the next meeting. Contributions are requested.

11.3 Future Meetings

1) Seventh meeting

- November 11(Tu) - 14(Fr), 1986
- Nürnberg
- Topics
 - Source coding algorithm
 - Video multiplex coding
 - Transmission coding
 - Frame structure
 - Compatibility check at laboratories and international field trial
 - General aspects of m x 64 kbit/s codec

2) Eighth meeting

- U.S.A.

Annex 1: Action plan for simulation
Annex 2: VLC for Block Type
Annex 3: Multipoint considerations
Annex 4: Flexible hardware specification

LIST OF PARTICIPANTS
(Montreal; June 17 - 20, 1986)

<u>Chairman</u>	S. Okubo	- NTT
<u>Core Members</u>		
F. R. of Germany	J. Speidel	- PKI/TEKADE
	G. Zedler	- FTZ
Canada	B. Mitani	- Teleglobe Canada (acting for S. K. Sharma)
	S. Sabri	- BNR
U.S.A.	B. G. Haskell	- AT&T Bell Lab.
	R. A. Schaphorst	- DIS
	A. Tabatabai	- Bell CORE (acting for J. A. Bellisio)
France	J. Guichard	- CNET
	J. P. Temime	- CNET
Italy	M. Guglielmo	- CSELT (acting for L. Chariglione)
Japan	M. Kaneko	- KDD (acting for Y. Hatori)
	N. Mukawa	- NTT
Netherlands	F. Booman	- DNL
United Kingdom	R. Nicol	- BT
	N. Shilston	- GEC (acting for D. Bonnie)
Sweden	H. Brusewitz	- Swedish Telecom Admin. (acting for P. Weiss)
<u>Assisting Experts</u>		
F. R. of Germany	W. Geuen	- FTZ-FI
Canada	A. Golembiowski	- BNR (Secretary)
U.S.A.	C. Clapp	- Pictel Corp.
	S. Ericsson	- Pictel Corp.
	G. J. Pearson	- CLI
	L. Renick	- CLI
Japan	T. Koga	- NEC
	K. Matsuda	- Fujitsu
United Kingdom	D. G. Morrison	- BT

Annex 1 ACTION POINTS FOR SIMULATION WORK

1. Presentation of the Statistics

It is agreed to present:

- R.M.S. and SNR ($10\log_{10} \frac{255^2}{MS}$) for the luminance.
- Mean value of the step size.
- Mean value of the number of non-zero coefficients.
- Mean value of the number of zeroes before the last non-zero coefficient.
- For the rest of the presentation, see p.3/Document #133 (the first frame and the switch are excluded when calculating the mean value).
- The same figures have to be provided for the 15th encoded picture:

Original picture	1	2	3	4	5	6	7	8	9	..
Coded picture	1		*		2		3		4	..

* omitted after the start

2. Reference Model Version 2 (RM2)

- a. The quantizer step size is transmitted as a GOB attribute (GOB not counted).
- b. Decision inter/intra: replace 128 with the mean value of the block.
- c. A priori knowledge for scene cut: intra mode only, global (picture) attribute.
- d. VLC for block attribute: see p.1/paragraph 3 of Document #107.
- e. VLC for coefficients: see TABLE 4/Document #122. EOB = 010. Maximum values +/- 70. Clipping must be introduced for coefficients whose absolute value is greater than 70.
- f. Detection of motion vector: see Figure 2/p.5 of Document #107.
- g. For the rest, see Document #104.

3. Presentation of Split-Trevor Sequence

When presenting this sequence, it is suggested to have:

- a) the complete sequence
- b) only the Split part
- c) only the Trevor part

4. Quantization Strategy

- non-linear quantizer
- classification
- multiple VLC, max 4 VLCs, max 16 bits

5. Filtering in the Loop

6. Motion Vector

- differential method
- post processing for motion vector
- fractional displacement

7. Coding Control

8. Pre- and post-processing

9. Error Resilience

10. Block Addressing

Points 4-8 and 10 have to be compared respectively to the Reference Model 2.

End

Annex 2 VLC EXAMPLES FOR BLOCK ATTRIBUTE

Luminance

Alternative 1: Zero difference motion vector is coded as Block Attribute

Code	Block Type	Transmitted Data

1	Fixed	-
001	Interframe	QI
010	MC, Coded, DMV≠0	DMV, QI
011	MC, Coded, DMV=0	QI
0001	MC, Not Coded, DMV=0	-
00001	MC, Not Coded, DMV≠0	DMV
00000	Intraframe	QI

Alternative 2: Zero difference motion vector is coded as vector data.

Code	Block Type	Transmitted Data

1	Fixed	-
01	Interframe	QI
001	MC, Coded	DMV, QI
0001	MC, Not Coded	DMV
0000	Intraframe	QI

Chrominance

Code	Block Type	Transmitted Data

1	Fixed	-
00	Interframe	QI
01	Intraframe	QI

End

Annex 3 Multipoint Considerations for n x 384 kbit/s CODEC

1. CLOCK SYNCHRONISM

This is easily obtained if the whole network is synchronous. If this is not the case (eg. one subnetwork is not synchronous), the MCU must act as a synchronization source and impose the same clock (its own one) to all codecs. If the MCU is connected to a synchronous network, then the MCU can have its own clock or take its clock from any incoming port.

To impose its clock to all codecs, the MCU may use the "Synchronous operation bit" defined in Document #77. As said during the last Tokyo meeting, the best place for such a bit is in the Service Channel of the frame structure for the following reasons:

- TSO is not end-to-end transparently transmitted.
- For SG XVIII, ISDN is a synchronous network and there is no need to modify G.704 or I.431.
- Even if the network would be capable of handling this bit in TSO, it will be difficult and expensive to have this function implemented in all digital exchanges at the same time in all countries.

2. VIDEO SWITCHING

Two solutions are possible. The first one, similar to the one used in the first generation codecs, limits the exchange of information by using one-way only protocols, ie an action is done while a signal describing this action is sent without any command/response protocol. The second one, uses a message-type protocol where every action is preceded by a request and an acknowledge, which introduces unacceptable delay in video switching. The first solution is preferred.

The absence of degradation during the switching may be obtained by the following procedure:

- 1) The MCU when deciding to switch (automatically or on chairman/conferee request) asks all decoders to freeze their received video ("Freeze Picture Request").
- 2) It then executes the switching from video signal A to video signal B. Degradations occurring in all decoders (buffer resync, video multiplex resync) are not visible since the picture is frozen. Decoders then regain all synchronizations.
- 3) It then asks the switched-in encoder A to empty its transmit buffer and encode its next picture in intraframe mode with such coding parameters as to avoid buffer overflow ("Fast Update Request").
- 4) When receiving a picture attribute indicating intraframe mode, the decoders resume normal video decoding.

Points 2) or 3) may occur simultaneously or be inverted.

It is proposed that those "Freeze Picture Request" (FPR) and "Fast Update Request" (FUR) may be coded with one bit each in the Application Channel of the Audio Time Slot.

3. DATA CONTINUITY

For the moment, two kinds of data transmission are envisaged: one at 64 kbit/s using the data time slot, and the other at 320 or 256 kbit/s replacing the moving video.

A solution which does not involve any message channel or command/response protocol, eg by using BAS monitoring, is not viable since it is delicately implementable and seriously restricts the service. The following procedure is proposed:

- i) Codec A sends a request to transmit data to the MCU and clears its data channel.
- ii) The MCU asks all codecs to clear their data channel in turn.
- iii) All codecs clear their data channel.
- iv) The MCU, having received all acknowledgements, inserts the A data channel in all outgoing ports (except A) and sends A an authorisation to transmit.
- v) A transmits data. When it is over, A sends a request to end data transmission to the MCU.
- iv) The MCU withdraws the A data channel from outgoing ports and sends an acknowledge to A. It then sends to all other codecs a request to resume video transmission in the data channel.

Note 1: This protocol does not impose A to be seen by everybody before the data transmission begins.

Note 2: This protocol only considers one-way data transmission. For two-way transmission, the protocol must be completed to identify which codec is answering and to insert its data time slot in the stream going to A.

4. ENCRYPTION

This is by far the most complex problem since audio signals cannot anymore be decoded, "(n-1)*summed" and recoded. Continuous presence for audio must then be lost and audio must be switched separately. It does not seem possible to switch audio and video together since audio requires fast switching (some ms) while video switching is inherently slow (some tenths of a second). Also the video switching criterion may be different from the audio one. Anyway some basic assumptions may be made:

- Audio, data and video must be ciphered separately to allow independent switching.
- Unciphered audio power must be available to the MCU.
- Only a memoryless and/or self synchronizing cipherment algorithm can be used to allow switching.
- Those considerations do not prevent the use of regular initialisation vectors.

Annex 4

SPECIFICATION FOR THE FLEXIBLE PROTOTYPE n x 384 kbit/s VIDEO CODEC

The aim has been to produce a specification which essentially is the minimum implementation with which all laboratories will comply. The specification will provide considerable scope for optimization and experimentation.

We would expect that the final Recommendation to be produced in early 1988 would be significantly different as it would reflect advances made during 1986/87. It would also have the areas of flexibility removed.

An outline block diagram of the codec is given in Fig. 1.

CONTENTS

1. Source Coder
2. Video Multiplex Coder
3. Transmission Coder

1. Source Coder

1.1 Source Format

1.1.1 The format to be coded is 288 lines, 30000/1001 (approximately 29.97) non-interlaced pictures per second - the Common Intermediate Format. The tolerance on the picture frequency is ± 50 ppm.

1.1.2 Pictures are coded in component form, these components being luminance (Y) and two colour difference signals (C_R , C_B). These three components and the codes representing their sampled values are as defined in CCIR Rec. 601.

Black = 16

White = 235

Zero colour difference = 128

Peak colour difference = 16 and 240

Codes outside the above ranges will be accepted but may be modified by the coder to avoid emulation of reserved codewords.

1.1.3 The luminance sampling structure is 288 lines per picture, 360 pels per line in an orthogonal sampling arrangement. The colour difference sampling parameters are 144 lines, 180 samples per line, orthogonal. Both C_R and C_B samples are sited such that their block boundaries coincide with luminance block boundaries. (See Table 1 and Fig. 2)

1.1.4 Before coding the number of samples per line is reduced to 352 for luminance and 176 for each colour difference signal by discarding 8 pels and 4 pels respectively from the source as shown in Figure 3.

1.2 Video Coding Algorithm

The video coder algorithm is shown in generalized form in Figure 4. The main elements are prediction, block transformation, quantization and classification.

1.2.1 The prediction is inter-picture. The predictor may incorporate movement compensation. The predictor requires a picture store with one pel access for each pel being predicted. (Note: multi-pel prediction* may be added later, introducing a requirement for a multiple access picture store configuration.)

1.2.2 Motion compensation is optional at the coder. The decoder will accept one motion vector for each luminance block of size 8 pels by 8 lines. The maximum motion vector is ± 15 pels and ± 15 lines. Only integer values of the horizontal and vertical components of the vector are currently considered. Fractional values* may be added later. (Note: The encoding method for transmission of motion vectors may restrict the vectors to a subset of all these possible values that the decoder hardware can accommodate. This will be specified later.)

A positive value of the horizontal component of the displacement vector signifies that the prediction is formed from pels in the previous picture which are spatially to the right of the pels being predicted.

A positive value of the vertical component of the displacement vector signifies that the prediction is formed from pels in the previous picture which are spatially below the pels being predicted.

Motion compensation vectors are restricted such that all prediction pels are taken from the coded picture area.

1.2.3 Coder

The prediction error (inter-picture mode) or the original picture is subdivided into 8 pel by 8 line blocks which are segmented as transmitted or non-transmitted. The method of choosing between inter or intra mode is not defined in this specification, nor is the segmentation criterion is not part of the specification, both being left open to equipment designers and may be varied dynamically as part of the data rate control strategy.

Transmitted blocks are coded by a transform based scheme. Coding block size is 8 lines by 8 pels (64 sample values) for Y , C_R and C_B .

1.2.4 Transformer

The transformer shall be implemented in a flexible manner so that a number of different transforms, or configurations of a particular transform in terms of bits per stage, can be investigated. All hardware should be equivalent in performance to classical matrix multiplication.

The 2-D transform is implemented as the equivalent of 2 independent 1-D transforms.

All coefficients in the forward and inverse transforms shall be programmable to 16 bit resolution.

For the purposes of compatibility it is only necessary to specify the inverse transform. As a preliminary specification the hardware should allow for 12 bit resolution in (ie coefficients) and 9 bits out.

The matrix multiplier implementation of the inverse transform uses 16 by 16 multipliers. No loss of accuracy occurs in the accumulator. The matrix coefficients are stored to 16 bits. The most significant 16 bits by truncation are carried between the 1-D transforms. The output of the inverse transform is rounded to 9 bits giving pel domain values in the range -256 to 255. The order of the 1-D transformers is yet to be specified.

1.2.5 Classification

The classification method itself, ie the way the choice of the parameters to encode a block is made, is not a matter of recommendation. However, the range of possibilities must be the same at the encoder and the decoder. Also the way the parameters are transmitted must be specified (see Section 2).

Each class is described by a table (PROM) of dimension 64 x 8 bits. In this table 6 bits (for coefficients 0 up to 63) are used to describe how many of and the order in which the coefficients are transmitted. The other 2 bits are reserved for future use to change the coding characteristics of the various coefficients of the block (eg related to prediction, quantization, variable length code set). The number of such PROM tables to be implemented is 8.

1.2.6 Quantizer

The number of inverse quantizers provided in the decoder shall be 32. Each quantizer has 12 bits input and up to 12 bits out.

(The same or different for luminance/chrominance?
Adaptive/non-adaptive)

1.2.7 VLC

The VLC to be used to encode the quantization index of AC components is given in Table 4/Document #113. This VLC incorporates one codeword for the EOB and presents some properties to facilitate auto-synchronization. The maximum wordlength is 16 bits.

The hardware must be designed to be easily extended to 4 VLCs controlled as a function of class, quantizer and sequency.

1.3 Data Rate Control and Subsampling Modes

The exact method of assessing the encoder data generation rate need not be specified but the specified maximum size of encoder buffer must incorporate an allowance for latency in the assessment and control loop. Hence, any requirement to constrain overall system delay may effectively preclude some schemes.

Control information is carried by side information - not derived recursively from received data.

Horizontal subsampling	- picture, group of blocks or block basis?
Vertical subsampling	- picture, group of blocks or block basis?

- Temporal subsampling - picture basis only. Interpolated pictures are not placed in the picture memory.
- Quantizer selection - see also section 1.4.
- Block significance criterion - not part of the specification

The coding algorithm will automatically permit the full quality of the source format specified in section 1.1 to be realized on still pictures.

1.4 Forced Updating

There is no separate coding scheme for forced update. This function is achieved merely by forcing the use of the intra-picture mode of the coding algorithm. Since the decoder cannot distinguish between normal and forced update blocks, then there are no parameters to specify which are unique to forced update. The bit rate allotted to forced updating, the sequence in which blocks are updated etc, are not specified.

In order that the quality of forced update blocks can be sufficiently high at all times, quantizer selection must not be dependent solely on buffer fill state.

2. Video Multiplex Coder

2.1 Tasks

The video multiplex coder has the following tasks:

- 1) Block address coding
- 2) Video data formatting and serializing
- 3) Synchronization - picture/line
- 4) Motion vector coding
- 5) Side channels for indicating dynamic coding parameters eg subsampling modes, quantizers, buffer state etc. (Permanent or transient channels? Transient channels require care in switched multipoint.)

2.2 Video Multiplex Arrangement

2.2.1 Picture Start Code (PSC)

Unique sequence of f bits [Buffer state] [Temporal ref.] [Type]

The buffer state is a 6 bit number representing the encoder buffer fullness in 1Kbit units at the beginning of this picture.

The temporal reference is a 3 bit number representing the time sequence, in Common Intermediate Format picture periods, of a particular picture.

Type is a Variable Length Code which allows block attributes to be applied to all blocks within a picture.

All PSCs are transmitted. Two successive PSCs indicate the intervening picture has been dropped.

2.2.2 Group of Blocks Start Code (GBSC)

Unique sequence of g bits [Group Number] [Type]

A group of blocks consists of two lines of luminance blocks, one line of C_R blocks and one line of C_B blocks. See Fig. 5.

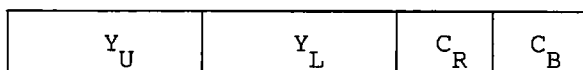


Fig. 5

Subsampling aspects are open for further consideration.

The group number is a 5 bit number, ranging from 00001_2 to 10010_2 , representing the vertical spatial position, in units of groups, of the current group of blocks.

Type is a Variable Length Code which allows block attributes to be applied to all blocks within a group of blocks.

All GBSCs are transmitted.

2.2.3 [Block Address]

Block Address indicates the position of a transmitted block within a group of blocks. (A decision is urgently needed on how this is achieved. Possible methods include relative addressing of only transmitted blocks and one bit per block for all blocks.)

The range of absolute block addresses for the upper line of luminance blocks (Y_U) is 0 to 43.

The range of absolute block addresses for the lower line of luminance blocks (Y_L) is 44 to 87.

The range of absolute block addresses for C_R blocks is 88 to 109.

The range of absolute block addresses for C_B blocks is 110 to 131.

2.2.4 Block Type

A Variable Length Code representing the type of the block.

- 1) Intra-picture coded block
- 2) Inter-picture coded block
- 3) Motion compensated block
 - Differential Motion Vector (DMV) $\neq 0$
 - DMV = 0
- 4) Motion compensated with coded residue
 - DMV $\neq 0$
 - DMV = 0
- 5) Future expansion on
- 6) Future expansion off

The decoder shall be designed to discard all data between types 5) and 6) and also between type 5) and the next GBSC. The exact method the decoder uses to discard this data is still to be defined.

2.2.5 Motion Vector Coding

- 1) Motion vectors are transmitted differentially, with respect to the MV of the previous block. The first block MV is related AMV: 0,0. The hardware should allow the Global MV in the GOB attribute to be also the basis of differential coding of MV.
- 2) Differential Motion Vectors are coded using a VLC (2 entries).
- 3) Maximum length of VLC for differential motion vectors is 16.

2.2.6 Block Data

Data specifying motion vectors, scanning class, quantizer type, transform coefficients and end of block marker (EOB). Although some of these may not be present, they are always transmitted in the following order.

.....
.....
.....
Transform coefficients
EOB

EOB is one of the codewords in the VLC set used for quantized transform coefficients. It is always present for every block for which coefficients are transmitted.

Each class number is transmitted with 3 bits (VLC?). This number might be reduced if less than 8 classes are needed. The quantization index is then transmitted using codewords as indicated in 1.2.7. This is followed by an EOB codeword as specified in 1.2.7.

2.3 Multipoint Considerations

2.3.1 Freeze Picture Request (FPR)

On receipt of a FPR, conveyed in bit of timeslot the decoder shall freeze its received video until the next picture attribute indicating intra-picture mode is encountered.

2.3.2 Fast Update Request (FUR)

On receipt of a FUR, conveyed in bit of timeslot of the encoder shall empty its transmit buffer and encode its next picture in intra-picture mode with such coding parameters as to avoid buffer overflow.

2.3.3 Data Continuity

The protocol adopted for ensuring data continuity in a multipoint connection will be handled by the message channel.

3. Transmission Coder

3.1 The transmission coder assembles all data and interfaces to the digital line transmission system.

3.1.1 The data rate is $n \times 384$ kbit/s where n is an integer between 1 and 5, both inclusive.

3.1.2 The coder channel output clock rate source shall be switchable between either a free running internal source or a source synchronized to the received data from the network. The mechanism for this switching is bit of the application channel.

3.1.3 When in free running mode the tolerance on output clock rate will be ± 50 ppm of nominal.

3.1.4 When in synchronized mode the synchronism should be maintained when the frequency of the received data clock is within ± 50 ppm of nominal.

3.2 Framing Structure (See Fig. 6)

As per CCITT Study Group XV WP XV/1 Doc. #58 plus the following coding for the applications channel:

List of codec attributes/facilities/parameters needing transmission from transmitter to receiver. No return path is assumed. Currently this list includes only encryption.

Operability with audioconferencing. Timeslot positioning to CCITT Rec I.431.

3.3 Video Data Buffering

The size of the transmission buffer at the transmitter is switchable from 8 Kbits to 64 Kbits, both inclusive, in steps of 8Kbits. ($K=1024$) Buffer size should be related to the transmission rate (overall - not video) to ensure acceptable system delay.

3.4 Video clock justification

Not provided.

3.5 Optional full spatial resolution mode data for quasi-stationary pictures.

To be specified later if required.

3.6 Audio

As per CCITT Draft Rec. G.72X type 2 terminal.

The audio channel is carried by the first time slot.

Flexible testbed hardware need only incorporate?

(64 kbit/s A-law
64 kbit/s u-law
56 kbit/s sub-band ADPCM according to CCITT Draft Rec.
G.72X)

3.7 Error Handling

Video coding strategy to be error resilient without internal or external error corrector. Note that demand refresh can be implemented using the Fast Update Request of 2.3.2.

3.8 Encryption

3.9 Data transmission

Framing structure to allow 2 data ports of 64 kbit/s each, though picture quality constraints may require only one to be available at 384 kbit/s. In this case time slot 4 is used.

3.10 Network Interface

Access will be at the primary rate interface with vacated time slots. CCITT Recommendation I.431.

For 1544 kbit/s interfaces the default H0 channel is timeslots 1 to 6. Interface code is B8ZS or AMI.

For 2048 kbit/s interfaces the default H0 channel is timeslots 1-2-3-17-18-19. Interface code is HDB3.

END

* Hardware to be capable of these additions later.

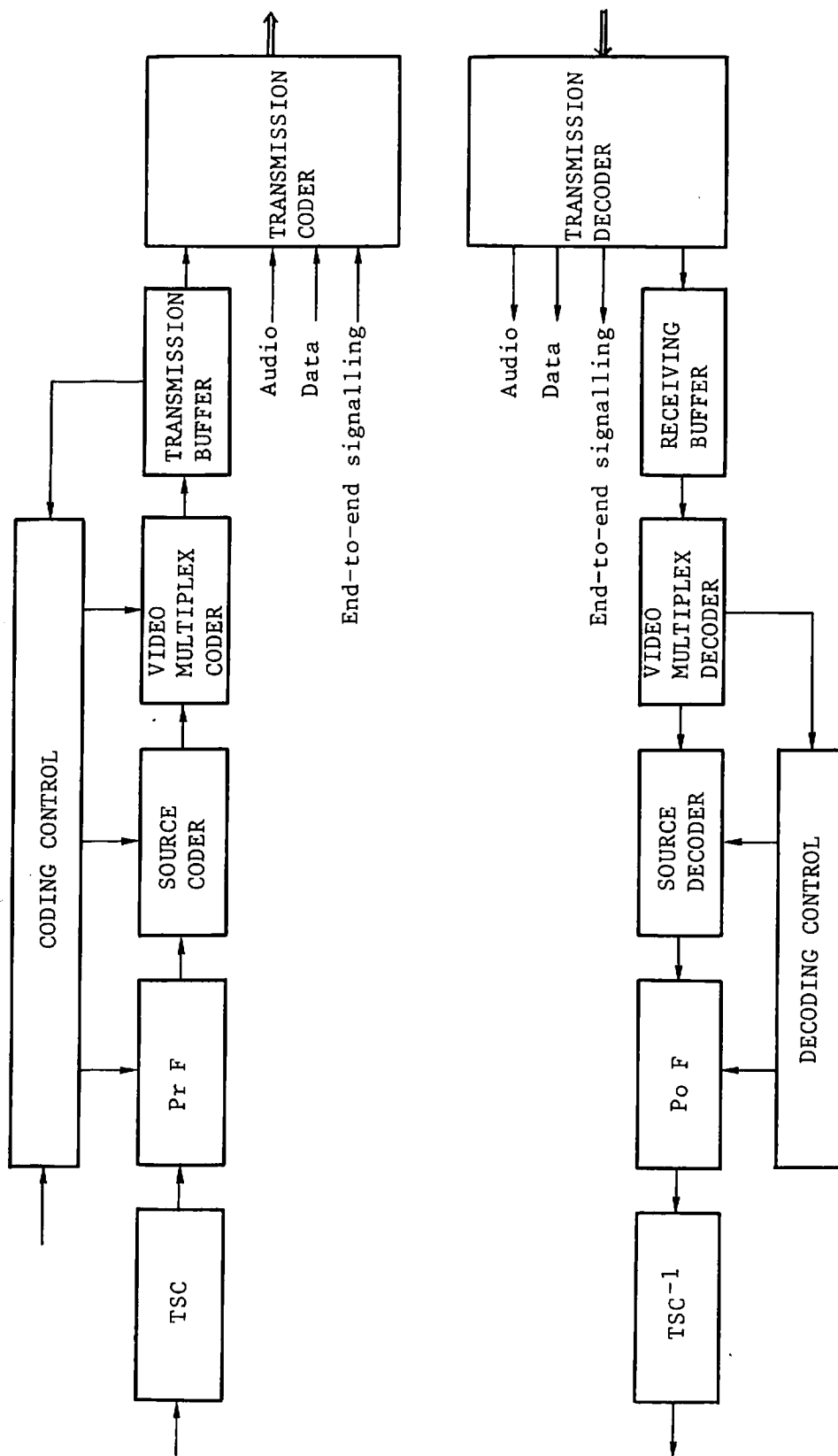


Fig. 1

Table 1. Basic parameters for the new generation n x 384 kbit/s CODEC

Items	Parameters
1. Reference point	Point B in Fig. 1/Annex 1 to COM XV-R 4
2. Baseband signals and their levels	Y, R-Y, B-Y, as defined in CCIR Rec. 601
3. Number of pels per line	Y: 360 (Note 1) R-Y: 180 B-Y: 180
4. Number of lines per field	Y: 288 (Note 2) R-Y: 144 B-Y: 144
5. Field frequency	Y, R-Y, B-Y: 29.97 Hz
6. Interlace	Y, R-Y, B-Y: 1:1
7. Sampling structure	Y, R-Y, B-Y: orthogonal, positioning of R-Y and B-Y samples share the same block boundaries with Y samples as shown in Fig. 2

Note 1: Active line duration is approximately 53 us.

Note 2: Active field duration is approximately 18.4 ms (for 625/50 systems) and approximately 15.2 ms (for 525/60 systems).

Note 3: The common intermediate format defines the maximum attainable spatial and temporal resolution in the codec. Effective resolution may eventually be reduced by some coding operating modes.

Note 4: The common intermediate format is a logical specification to ensure compatibility among codecs. Hence, it might not appear at the physical interface points in the codec.

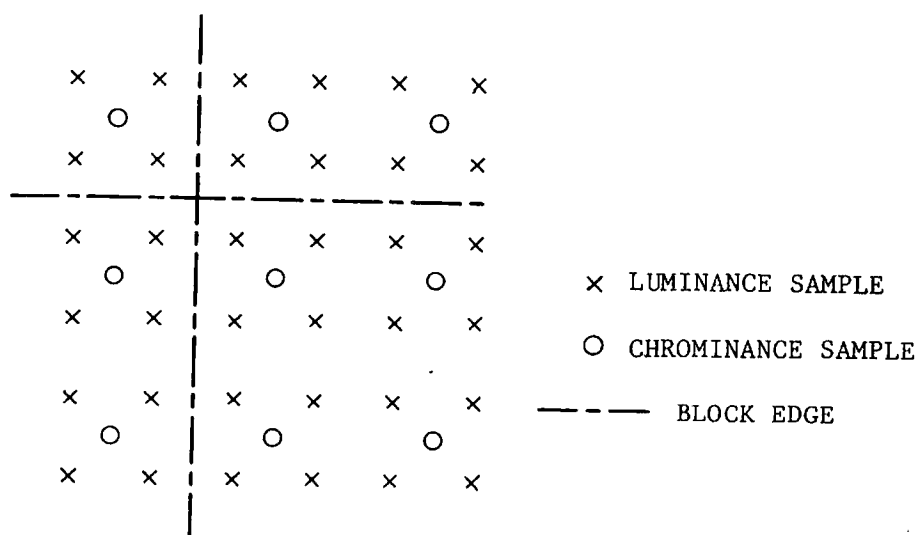


Fig. 2 Positioning of Luminance and Chrominance Samples

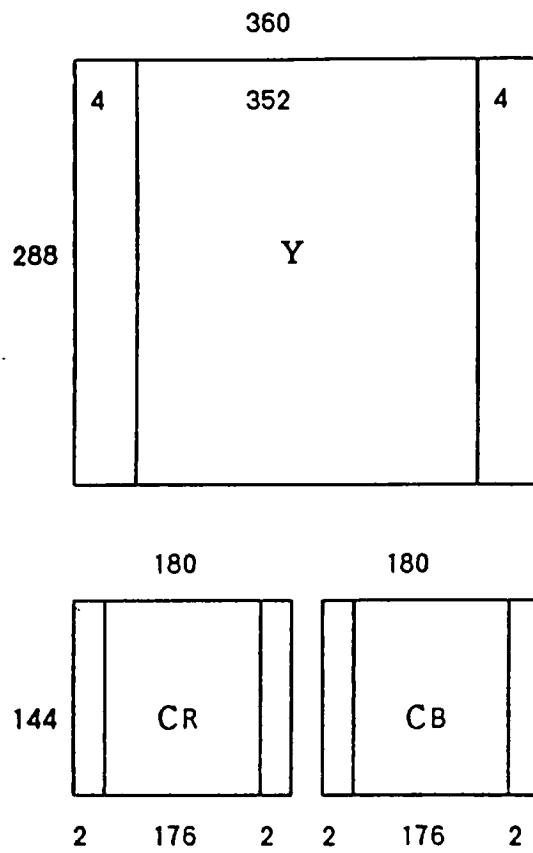
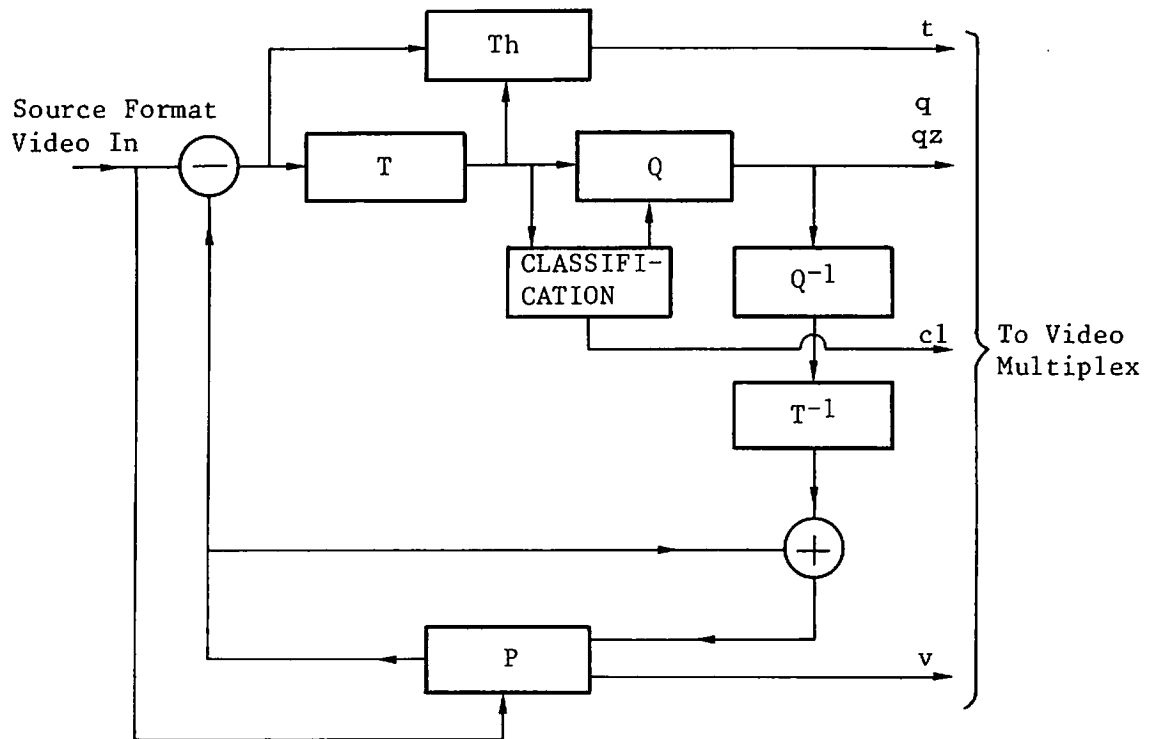
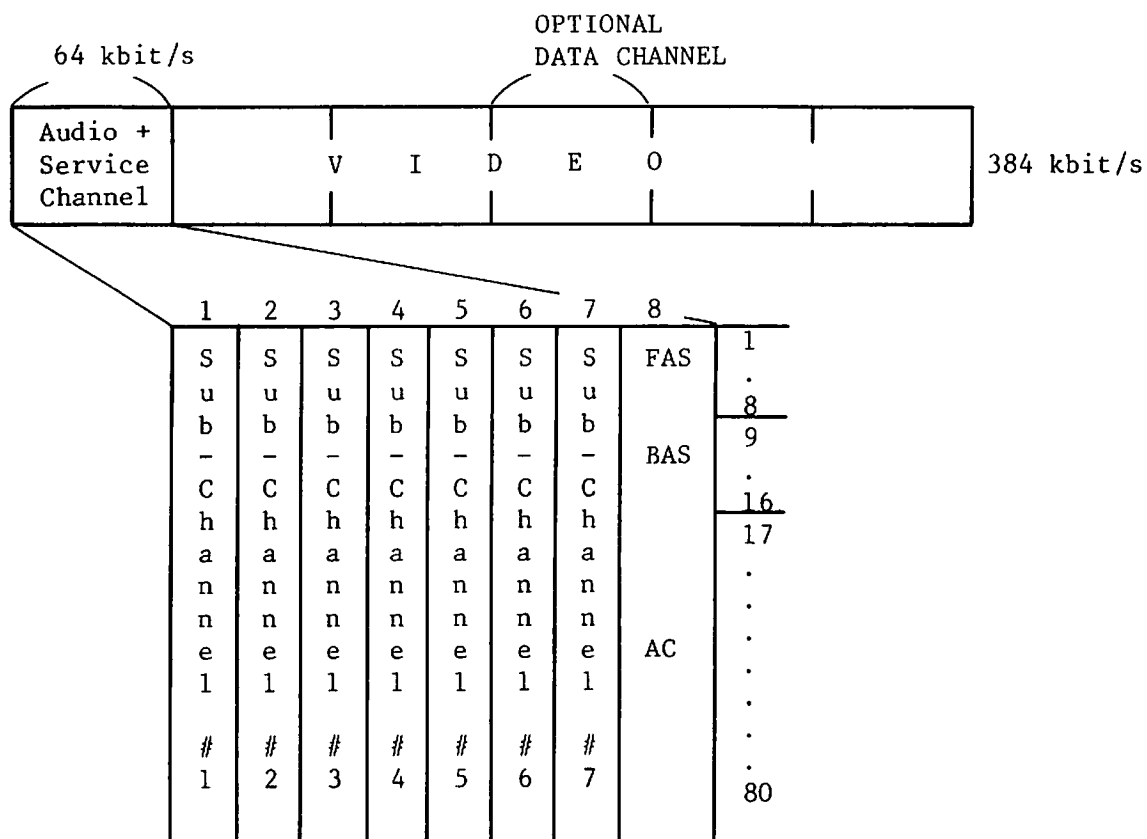


Figure 3 Definition of significant pel area



Th	Threshold
T	Transform
Q	Quantizer
Q ⁻¹	Inverse Quantizer
T ⁻¹	Inverse Transform
P	Predictor
t	Flag for transmitted/not transmitted block
q	Quantizing index for transform coefficients
qz	Quantizer indication
v	Motion vector
cl	Classification index

Figure 4



FAS: Frame Alignment Signal (note 1)
BAS: Bitrate Allocation Signal
AC: Applicaion Channel

Note 1: The block termed as FAS contains also other information than for frame alignment purposes.

Fig. 6 Frame Structure for $n \times 384$ kbit/s codec (in case of $n = 1$)