CCITT SGXV
Working Party XV/1
Specialists Group on Coding for Visual Telephony

Document #346R
March 25, 1988

SOURCE: CHAIRMAN OF THE SPECIALISTS GROUP ON CODING FOR VISUAL TELEPHONY
TITLE : REPORT OF THE TWELFTH MEETING IN THE HAGUE (March 22-25, 1988)

CONTENTS

1. General

The Specialist Group met in The Hague, The Netherlands, from March 22 to 25, 1988 at the kind invitation of The Netherlands PTT. Welcoming address was delivered by Mr. Groen on behalf of the hosting organization.

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

At the closing session, the following changes of core members were announced;

Bellcore	from Mr. H. Gharavi	to Mr. A. Tabatabai,
KDD	from Mr. Y. Hatori	to Mr. M. Wada,
NTT	from Mr. N. Mukawa	to Mr. Y. Kato,
STA	from Mr. P. Weiss	to Mr. H. Brusewitz.

Furthermore, Chairman thanked the hosting organization for the meeting facilities provided and for the excellent operation of the meeting.

2. Documents for the Meeting (TD2)

For this meeting, 29 normal documents and 11 temporary documents have been available. Annex 1 shows the outline of each document.

3. Tape Demonstrations (TD3)

A number of tape demonstrations as listed in Annex 2 were presented during the meeting.

4. Discussion on nx384 kbit/s Codec

4.1 Progress of Flexible Hardware experiments (#319, #327, #331, #337; TD11)

Current status of each hardware development in Japan, France, UK and FRG/Netherlands was reported with documents and tape demonstrations. DIS test signal was correctly decoded by four Flexible Hardwares. Four connections between independently designed hardwares were made between the following projects;

- UK and Japan (back to back in UK)
- Two projects in Japan (back to back)

- France and UK (via satellite)
- France and Japan (via satellite, Japanese hardware placed in UK)

and basic compatibility was confirmed.

During the questions and answers, some members expressed concerns with the gain of motion compensation. This point needs further investigation.

4.2 Transform part specification (Annex 3/#317R; #321, #326, #340, #345; TD5)

Three documents (#321, #326, #340) were presented toward the approach of specifying mismatch error and refresh rate. Mr. Haskell reported the correspondence work with transform chip design proposers aiming at a practical specification, including his draft specification together with indication that at least four design meet this specification. During the discussion, the following observations were made;

- At lower bit rates, the affect of IDCT mismatch errors is reduced because of coding error domination and more fixed blocks.
- Scaling of input to IDCT may affect the mismatch error.
- For input test data for IDCT, worst case vectors are not useful, random data input looks sufficient. As far as a few random data tried are concerned, specification on how to generate them is not necessary.
- A combination of slightly different IDCTs which result in anti-correlated mismatch errors may represent a worst case.
- Biased error is a problem since it accumulates in a monotonic way.

Taking these information and observations into account, the meeting agreed to seek the 'preferred approach' in Annex 3 to Doc #317R, and asked Mr. Haskell to chair a small group meeting to formulate a specification for mismatch error and refresh rate as well as action points.

The outcome is in Annex 3 to this report. As to the DCT chip standardization, Mr. Carr undertook to organize a meeting of DCT chip manufacturers to be held in Martlesham Heath. Furthermore, Mr. Liou undertook to act as a liaison between this group and the IEEE Circuit and Systems Society.

4.3 Source Coding

4.3.1 Quantization (#329)

As a measure to cope with the clipping which occurs in small step size intra mode blocks in Flexible Hardware, necessary dynamic ranges of the quantizer were analyzed for the worst case of computer graphic type sharp images. Based on these values, VLC code sets meeting the requirements were suggested.

Since there was a proposal in the previous meeting addressed to the same problem (Doc. #285), the members were requested to compare these two solutions and reach a conclusion for Rec. H.12x.

4.3.2 Loop filter (#320, #327)

Experimental results regarding the on/off control methods and range of blocks to be filtered were provided, suggesting the use of motion vector controlled filtering and 2:1 finer quantization for chrominance. For finer quantization for chrominance, European members expressed that they had no evidence. The meeting recognized more experience is required.

As to the filtering area comparison, 8x8 vs 10x10, French results indicated that there were no significant difference between them. The meeting decided to include 8x8 filtering in the draft Rec. H.12x.

4.4 Video multiplex coding

4.4.1 Transform coefficient coding (#322)

Comparison of 2-D VLC and 1-D VLC was presented with tape demonstration, concluding that it is difficult to observe difference, but further study is required to obtain more firm conclusion. It was pointed out that the reason might be the code set being not optimized due to the hardware restriction.

Further study is required.

4.4.2 Motion vector coding (#323)

Coding method and VLC code sets were proposed for inclusion in the draft Rec. H.12x. For the tracking range, a different proposal is contained in Doc. #334. The meeting decided this matter would be discussed in the drafting group meeting for Rec. H.12x.

There was a discussion on the comparison between absolute and differential coding, concluding that experimental results were awaited, regarding vulnerability to transmission errors in particular.

4.4.3 Block attribute coding (#328)

Signalling methods for loop filter, classification index, and quantizer were addressed to allow Rec. H.12x for a bit rate of nx384 kbit/s as well as mx64 kbit/s. After some questions and answers, it was decided these matters were dealt with in the drafting group meeting.

4.4.4 Picture Headers for dropped pictures (#324)

The meeting agreed not to transmit Picture Headers for dropped pictures, but extend the value of modulo for Temporal Reference.

4.5 Transmission buffer (#324)

Doc. #324 gave further discussion for a buffer control scheme to cope with the mixture of pre-coding and post-coding buffers, which was proposed in Doc. #287. Extending the interpretation of 'buffer size' to the maximum coded data per picture was also proposed.

Further study through hardware experiments are expected.

4.6 Transmission coding

4.6.1 Error correcting codes (#325, #335)

Information on two error correcting codes, BCH and Reed-Solomon, was given. A question was made on what burst errors can be corrected by BCH codes. Another question was if it is possible to collect more accurate information concerning transmission errors in the channel.

The meeting recognized that experimental results on whether error resilience

approach be practical with the current motion compensated hybrid coding at 384 kbit/s were awaited. A suggestion was also made that we need careful consideration from system aspects point of view.

Mr. Carr announced that a European expert group was established for error correction, and its work would be reported at the next meeting.

4.6.2 Transmission clock frequency (#332)

Clarification was given to the previous Doc. #293 with the intention that the codec need not distinguish whether the network is synchronous or asynchronous. Experimental results are awaited to establish if the idea is practical.

4.6.3 Available networks for 384 kbit/s (#338)

Information of French networks were presented. During questions and answers, a problem was raised on what would happen in interworking if a network provides 6xB channels for 384 kbit/s transmission and the other network provides a H0 channel.

4.6.4 Value of n

Chairman raised a question for clarification whether the codec conforming to Rec. H.12x should operate for all values of n. After some discussion, the meeting has the following common understanding;

- Recommendation covers the whole range of n.
- The values of n which each codec implements is not the matter of Recommendation.
- A specific operational value of n for a session is determined by the communication procedure, e.g., AV.242.
- It is expected that the codec having the maximum value of n1 can operate for all the values from 1 to n1.

4.7 Updating Draft Rec. H.12x (Annex 5/#317R; #323, #324, #333, #334; TD6)

European position on outstanding items (Annex 6/#317R) was provided in Doc #333. Mr. Hatori presented Japanese position on corresponding items. In addition to those items which were confirmed through hardware experiments or other studies, the meeting considered inclusion of some items which allow operation at bit rates lower than 384 kbit/s (see §5.7.2/Doc. #317R).

Concerns were expressed that the inclusion relevant to lower bit rate operability should be based on hardware verification and that too much complexity should not be introduced for this purpose.

Before asking Mr. Morrison to coordinate a drafting group for Rec. H.12x, the meeting further discussed on the underlying issue concerning the scope of this Recommendation, which had been discussed in Tokyo.

Extensively discussed point was whether the concept of 'px64 kbit/s (p=1-30)' could be judged as practical at this stage. European members thought yes, while American and Japanese members thought no. After discussion, the meeting decided to accelerate the study for mx64 kbit/s, recognizing that we lack supporting study results for coding performance. One round of 'divergence' phase activities (cf Fig. 1/#317R) are requested before the next meeting in

September so that we will be able to share a common view on this issue at that time.

As to some items under study but with only a limited number of alternatives, the meeting decided to include them in the report of this meeting as well as in the progress report of the Specialists Group to Working Party XV/1 so that the information be published and distributed through the official CCITT channel.

The outcome of the drafting group chaired by Mr. Morrison is in Annex 4 (Draft H.12x) and Annex 5 (Sections of H.12x where a common view has not been reached) to this report.

4.8 Testing tool for equipment conforming to Recommendation (TD4)

Chairman raised the necessity of some tools to aid developing codecs conforming to Rec. H.12x. Though only test signal generators were described in TD4, analyzers for the coded signal were also suggested during the discussion. The meeting recognized that the idea was worth considering and volunteers to develop these testing tools were requested. Mr. Schaphorst undertook to submit a contribution at the next meeting.

4.9 Intellectual property (#342)

Statements from participating organizations were collected. In addition to these statements, Mr. Speidel expressed that Philips can follow the policy 'B' in Doc #342, and Mr. Devemieux expressed that his organization and Alcatel would submit a statement of policy 'B'.

Based on these statements, the meeting reached an agreement that participating organizations advise the Group of information concerning their patents relevant to nx384 kbit/s and mx64 kbit/s codec standardization as soon as those patents are applied.

In case that new members join the group, they are requested to follow the same policy and provide statements.

5. Discussion on mx64 kbit/s CODEC

5.1 Picture format (§6.2/#317R; #318, #341)

Japan and USA proposed that a medium resolution (4/9CIF) picture format should be added to the two formats agreed in Tokyo, full CIF and 1/4 CIF. The consideration was that it provides the minimum acceptable spatial resolution for three persons seated side by side and a good spatio-temporal trade-off. Europe expressed the opinion that this addition of the third format is not necessary. Argued points were;

- Evolution from low resolution to high resolution,
- Simple standard to avoid confusion, in multipoint communication in particular,
- Proliferation of local standards,
- Customer's choice,
- Stimulation to industry by single standard,
- etc.

It was recognized that the crucial question is which of full CIF or medium

resolution format provides more acceptable coding performance for mx64 kbit/s in the coming few years. This point should be clarified by carrying out the study.

Consequently the meeting decided to defer concluding until more simulation works would be obtained. Members are requested to include this study item in speeding up the mx64 kbit/s study described in §4.7 of this report.

In the outline draft Rec. H.12y, the picture format part contains the two formats agreed in Tokyo, replacing the two notes (Note 1, Note 2 in §6.2/#317R) with "The possibility of a third format of 4/9 (2/3 x 2/3) CIF is under study".

Other aspects of picture format raised in Doc #318 were also left for further study because of discussion time shortage.

5.2 Activities in T1Y1 (#341)

Progress of T1Y1 activities toward mx56/64 kbit/s video telephony ANSI standards were introduced. Study results on frame structure, picture format, network aspects and transmission formats are contained.

Chairman informed the meeting that Rec. H.221 for frame structure and AV.242 for audiovisual communication procedures would be elaborated in Working Party XV/1 meeting in April, 1988.

5.3 Coding algorithm study

5.3.1 Test sequence (#336)

It was pointed out " Salesman" sequence has some problems, and "Miss America" suffers also 15Hz flicker. Mr. Haskell undertook to take a look and report the results to those whom he sent the sequence.

5.3.2 Latest achievements (#330, #344)

Updated information on adaptive zonal scanning at 64 kbit/s was presented with tape demonstration.

Coding performance of RM5 was also presented with tape demonstration.

5.3.3 Study plan (#339, #343, #344; TD8)

As a study result in Europe, Reference Model 5 was presented in a complete document. The meeting recognized this model forms a good yardstick for the coming six months study in the Specialist Group.

Chairman proposed that coding schemes to be presented at the next meeting be accompanied by detailed description of the algorithm so that other parties can verify them. Doc. #339 is an example of the detailed description. The meeting accepted this proposal.

In order to make the achievements directly comparable at the next meeting and to make "convergence" easier, the meeting asked Mr. Guichard to coordinate a small group to set up rules and action points for the simulation work. The outcome is in Annex 6 of this report.

5.4 Outline draft Recommendation H.12y (TD9)

It was agreed to put forward the text in Annex 7 of this report to Study Group XV so that it will be annexed to a question of the next study period.

6. Others

6.1 Publicity in GLOBECOM'88 (TD10)

Four papers reflecting the activities of this group will be submitted for presentation in the GLOBECOM'88 session "Standard for Video Coding".

6.2 Liaison with ISO group

The meeting recognized the necessity of close liaison with ISO/TC97/SC2/WG8 in order to avoid the situation that these two video coding related groups will deal with the same field and produce different Recommendations.

7. Future Meeting

- 13th meeting hosted by France: September 19-22, 1988 in Paris.

Annexes

Annex 1: Documents for the Hague meeting
Annex 2: List of tape demonstrations
Annex 3: Action points for the transform part specification
Annex 4: Draft Recommendation H.12x
Annex 5: Sections of H.12x where a common view has not been reached
Annex 6: Report on algorithms study action group
Annex 7: Outline draft Recommendation H.12y

LIST OF PARTICIPANTS
(The Hague; March 22-25, 1988)

Chairman	S. Okubo	- NTT, Japan
Core Members		
F. R. of Germany	J. Speidel	- PKI
	G. Zedler	- FTZ
Canada	D. Lemay	- BNR
USA	A. Tabatabai	- BELLCORE (acting for H. Gharavi)
	B.G. Haskell	- AT&T Bell Laboratories
	R.A. Schaphorst	- DIS
France	G. Eude	- CNET
	J. Guichard	- CNET
Italy	M. Guglielmo	- CSELT (acting for L. Chariglione)
Japan	Y. Hatori	- KDD
	Y. Kato	- NTT (acting for N. Mukawa)
Korea	J-H. Jeong	- ETRI (acting for J-S. Lee)
Norway	G. Bjoentegaard	- Norwegian Telecom
The Netherlands	R.H.J.M. Plompen	- DNL
United Kingdom	R.C. Nicol	- BTRL
	N.L. Shilston	- GEC Video Systems
Sweden	H. Brusewitz	- Swedish Telecom Admin. (acting for P. Weiss)
Assisting Experts		
F. R. of Germany	W. Geuen	- FTZ-FI
USA	S. Ericsson	- PictureTel
	D.N. Hein	- CLI
	M.L. Liou	- Bellcore
	L.V. Rennick	- CLI
France	D. Devimeux	- SAT
Japan	K. Matsuda	- FUJITSU
	T. Murakami	- Mitsubishi Electric
	M. Ohta	- NEC
The Netherlands	P.J.M. Nooij	- DNL
	B.F. Shuurink	- DNL
United Kingdom	M.D. Carr	- BTRL
	T.S. Duffy	- GEC Video Systems
	R. Forsyth	- INMOS
	D. G. Morrison	- BTRL
	H. Yassaie	- INMOS

DOCUMENTS FOR THE HAGUE MEETING

Normal Documents

#318 mx64 kb/s CODEC PICTURE FORMAT (JAPAN)

In addition to the two formats agreed at the previous Tokyo meeting, full CIF and 1/4 CIF, a medium resolution format, 4/9 CIF is proposed. The following aspects of the picture format are also proposed;

- Temporal resolution: j/29.97 sec, j=1,2,3,4.
- Larger formats cover all the smaller ones.
- Negotiation determines spatial parameter first, then temporal parameter.
- Two to one horizontal and vertical subsampling for chrominance.

#319 PROGRESS REPORT ON JAPANESE HARDWARE PROJECTS (NTT, KDD, NEC, FUJITSU)

Progress of two Flexible Hardware projects in Japan are reported. Both are working in full mode, except some errors in the second project due to lack of entropy codec processing power. It was confirmed that the two type of codecs are compatible in the back-to-back connection test.

#320 FILTER IN THE CODING LOOP (NTT, KDD, NEC, FUJITSU)

Control methods and applied blocks for the loop filter were experimented with halving the quantizer step size for chrominance. It is concluded that finer quantization reduces dirty window noise as was described in Doc. #285 and control methods (MC vector or side information based) or applied blocks (luminance only or both of luminance and chrominance) do not make significant difference in picture quality.

#321 PRECISION FOR IDCT CALCULATION AND REFRESH CYCLE (NTT, KDD, NEC, FUJITSU)

IDCT mismatch is estimated by analyzing the worst case error. Taking the ideal IDCT as a baseline, the mismatch error is calculated as the sum of coding and decoding IDCT errors. Each IDCT error E is suggested to be less than 1/12 + 0.0027, while the refresh rate is specified as "interframe coding must not be continued more than n(=120) times at any blocks except fixed ones".

#322 TWO-DIMENSIONAL VLC FOR TRANSFORM COEFFICIENTS (NTT, KDD, NEC, FUJITSU)

A type of two-dimensional VLC has been implemented in the Flexible Hardware. The performance difference is reported to be difficult to distinguish between 1-D and 2-D VLCs.

#323 CODING OF MOTION VECTORS (NTT, KDD, NEC, FUJITSU)

The following specifications for motion vector coding are proposed;

- Differential coding,
- One dimensional coding for each component,
- Tracking range; +/-15 pels horizontal, +/-7 pels vertical,
- Two VLC code sets as shown in this document.

#324 TRANSMISSION BUFFER CONTROL (NTT, KDD, NEC, FUJITSU)

The variable delay scheme for receiving buffer control presented in Doc. #287 is further discussed. It is concluded that the scheme ensures secure operation in the mixture of pre-coding and post-coding buffers and has possibility of maintaining constant delay when picture dropping rate is up to a certain value. It is proposed that no picture headers for dropped pictures are sent, but Time Reference for coded picture is counted in mod 16. Furthermore, interpretation of "buffer size" is extended to include the maximum coded data per picture.

#325 A DOUBLE ERROR CORRECTION BCH ENCODER/DECODER LSI (MITSUBISHI ELECTRIC CORP)

An 8k-gate Gate Array implementation of BCH error correcting codecs is reported for information. It can provide double error correction for block length of up to 511 bits with the help of two external components. Only test samples are now deliverable.

#326 IDCT MISMATCH (CLI)

The results of an IDCT mismatch experiment run at 320 and 768 kbit/s are presented. The coding performance is shown as SNR versus the refresh period and MME defined in Doc. #281. As a conclusion, it is suggested that the required number of coded frames for refresh could be relaxed to 120 frames, or 8 seconds at 15 frames per second.

#327 HARDWARE RESULTS ON THE FLEXIBLE CODEC (FRANCE)

Operation at 384 kbit/s and 2048 kbit/s are demonstrated together with comparison of 8x8 and 10x10 loop filters. Working with "Miss America" which was simulated and implanted in PROMs is also demonstrated.

#328 SIGNALLING OF LOOP FILTER, CLASSIFICATION INDEX AND QUANTISER IN A px64 kbit/s CODEC - p=1,2,...,30 (FRG)

Reduction of overhead bits are proposed to allow a 384 kbit/s codec to operate at 64 kbit/s;

- TYPE 1 indicates whether loop filtering is indicated explicitly by block type or implicitly by motion vectors.
- TYPE 1 indicates whether multiple scanning classes are used or not.
- Defining the quantiser for Q equal subgroups of a GOB separately, where Q is given in QUANT 1.

#329 METHODS TO REDUCE CLIPPING EFFECTS (FRG)

In order to cope with clipping which occurs for the Flexible Hardware VLC code set if blocks are intra coded with step size of 4, two code sets providing +/-512 levels and +/-387 levels respectively are discussed. As a conclusion, use of two code sets is suggested; Flexible Hardware set for inter, and the set giving +/-387 levels for intra, because it is simple, maintains error resistancy against PSC or GBSC emulation, and does not decrease coding efficiency significantly.

#330 ADAPTIVE ZONAL SCANNING AT 64 kbit/s (UK)

Updated experimental results are given for the scheme where the zone used for any particular block is a function of current quantizer, indicating a small change in quality from a model based on RM4 with no zonal coding. It is concluded that care must be exercised in determining an optimum trade off between quantizer resolution and spatial resolution.

#331 VIDEO TAPE DEMONSTRATION OF FLEXIBLE HARDWARE (BT)

A list of tape demonstration for coded and decoded pictures.

#332 NETWORK SYNCHRONISATION - CLARIFICATION OF DOCUMENT #293 PRESENTED IN TOKYO (UK, FRANCE)

Clarification is given for the problem and a possible solution proposed in Doc. #293, particularly for the required range of the encoder clock frequency.

#333 EUROPEAN POSITION ON OUTSTANDING ITEMS FOR H.12x (FRANCE, FRG, ITALY, NETHERLANDS, NORWAY, SWEDEN, UK)

Summary of European position on the 39 items listed as needing further study in Annex 6 to Doc. #317R.

#334 UPDATES TO DRAFT RECOMMENDATION H.12x (FRANCE, FRG, ITALY, NETHERLANDS, NORWAY, SWEDEN, UK)

Changes to the current text (Annex 5 to Doc.#317R) are proposed according to the position described in Doc. #333.

#335 SOME NOTES ON REED-SOLOMON CODES (PICTURETEL)

Performance and implementation of Reed-Solomon error correcting codes are provided. For double error correction, block length is 2040 bits with redundancy of 1.6 %. The required number of operations for error finding from four syndromes in the receiver is 150 per block.

#336 COMMENTS ON SEQUENCE "SALESMAN" (SWEDEN, FRANCE)

It is pointed out that the sequence SALESMAN has the following problems;

- Line No.1 has the value of 0 for Y, Cr and Cb.
- Each of 288 chrominance lines have unique values.
- 15 Hz flicker is visible due to one noise pattern added to odd pictures and the other pattern to even pictures. Actions are proposed for these problems, as well as for dealing with the window for source coding and temporal subsampling.

#337 PROGRESS OF THE COMPATIBILITY CHECK BETWEEN BT AND JAPANESE CODECS (BT, NTT, NEC)

Progress of a back-to-back compatibility check between UK and Japan Flexible hardwares are presented, reporting that basic compatibility has been confirmed except motion compensation and adaptive scanning for interframe processing. Further progress may be given during the meeting.

#338 AVAILABLE NETWORKS IN FRANCE FOR px64 kbit/s BEARER SERVICES - VIDEOCONFERENCE APPLICATIONS (FRANCE)

The following three networks are available in France for px64 kbit/s bearer services;

- 2 Mbit/s leased lines,
- "Transdyn" providing unrestricted px64 kbit/s,
- Switched circuits on ISDN providing 64 kbit/s.

In the third case, synchronization equipment is required to obtain px64 kbit/s which uses a negotiation procedure at the first 64 kbit/s channel and H.221 framing at the remaining channels, resulting in provided data rate of (p - delta)x64 kbit/s. As a conclusion, it is stressed that a compatible coding of the signals at different data rate is very attractive.

#339 DESCRIPTION OF REF. MODEL 5 - RM5 (THE NETHERLANDS, FRANCE, FRG, ITALY, SWEDEN, BTRL)

In order to investigate px64 kbit/s (p=1-30) coding algorithm, Reference Model 5 is completely described. RM5 is based on RM4, with modifications in block attribute coding (macro block), motion compensation and loop filtering also for chrominance, buffers size dependent on the value of p, etc.

#340 MISMATCH ERROR ACCUMULATION IN A FAST DCT (TOSHIBA)

IDCT mismatch errors for a fast algorithm and a reduced bit precision matrix algorithm are investigated against the baseline IDCT defined in Doc. #281 at 384 and 768 kbit/s using RM4. From the SNR versus time data, the following conclusions are drawn out;

- The way in which mismatch error is accumulated may vary significantly from algorithm to algorithm, thus MME alone may not completely specify how error accumulates in the decoder.
- Quantizer step size can have a great effect on the appearance and accumulation of MME.

#341 mx56/64 kbps VIDEO TELEPHONY; FRAME FORMAT, PICTURE FORMAT, CODING AND NETWORK ASPECTS LEADING TO COMPATIBLE STANDARD (USA)

A material is provided for possible standards, covering all aspects of mx64 kbit/s videophone; video source format, video source coder, audio source coder,network aspects, transmission coder and communication procedure. In the video source format, an optional medium resolution format of 240(192) lines x 256 pels for luminance is defined in addition to the mandatory 1/4 CIF. 4:1 horizontal subsampling of chrominance is proposed. As to hypothetical reference connections, five different cases are considered; 56, 2x56, 64, 2x64 and (128 to 384) kbit/s transmission, leading to five different transmission formats. Videophone communication procedure is mentioned in addition to audiovisual procedure AV.221.

#342 PATENT INFORMATION DISCLOSURE POLICIES - SECOND ISSUE (CHAIRMAN)

A collection of statements received by March 22, 1988.

#343 SIMULATION RESULTS AT 384 kbit/s BASED ON RM5 MACRO BLOCK SCHEME (FRANCE)

Numerical data at 384 kbit/s are provided to demonstrate the capability of RM5 for px64 kbit/s. Data for RM4 are also given for comparison.

#344 SIMULATION RESULTS AT 64 kbit/s BASED ON RM5 MACRO BLOCK SCHEME (FRANCE)

Numerical data at 64 kbit/s are provided to demonstrate the capability of RM5 for px64 kbit/s.

#345 PROGRESS ON INVERSE IDCT ACCURACY SPECIFICATION (AT&T BELL LABS)

Contains copies of correspondence between parties interested in chip manufacture, algorithm simulation and flexible hardware experimentation.

Temporary Documents

No. 1 Agenda (Chairman)
No. 2 Available documents (Chairman)
No. 3 List of tape demonstrations (Chairman)
No. 4 Conformance checker (Chairman)
No. 5 Action points (Small group on transform part specification)
No. 6 Draft Recommendation H.12x (Drafting group)
No. 7 Draft report of the twelfth meeting in The Hague (Chairman)
No. 8 Report on algorithm study action group (Small group on algorithm study)
No. 9 Outline draft Recommendation H.12y (Chairman)
No.10 Publicity in GLOBECOM88 (Chairman)
No.11 Progress of the compatibility check between BT and Japanese codecs - 2nd report (BT, NTT, NEC)

LIST OF TAPE DEMONSTRATIONS (March 22, 1988)

Topics	Source	Document
a. Flexible Hardware performance - Full coder at 384 and 2048 kbit/s - Full decoder at 384 kbit/s (DIS pattern, Miss America) - Comparison of 8x8 and 10x10 loop filters	France	#327
b. Hardware progress at 384 kbit/s	NL/FRG	
c. Clipping effects with present hardware specification	NL/FRG	#329
d. Zonal coding	UK	#330
e. Flexible Hardware performance - BT coder to NTT decoder - DIS pattern generator to BT decoder - BT coder to BT decoder - NTT coder to BT decoder - BT coder to BT decoder at 64 kbit/s	UK	#331
f. Flexible Hardware performance - DIS pattern generator to the 1st project decoder - DIS pattern generator to the 2nd project decoder - The 2nd project coder to the 1st project decoder - The 1st project coder to the 2nd project decoder	Japan	#319
g. Filter in the loop - Blocks to be applied - Finer quantization for chrominance	Japan	#320
h. IDCT mismatch (MME=0.32)	USA	#326
i. Mismatch error of IDCT (Hardware processed) - Closed loop - Open loop with step size of 4	Japan	#321
j. Two-dimensional VLC	Japan	#322
k. Buffer control	Japan	#324
l. RM5 at 384 kbit/s	France	#343
m. RM5 at 64 kbit/s	France	#344
n. Impairments on "Salesman" found through 2x2 transform	France	#336

LETTER TO IDCT CHIP DESIGN PROPOSERS

To: Respondents to CCITT Request for IDCT

Dear Sirs and Mesdames,

Thank you for your responses so far to the CCITT request for IDCT chip information. Based on further computer simulations and hardware studies, we think we are nearer to a performance specification for ISDN video codecs (see attached). Could you examine the proposed method and let us know your thoughts? The peak error constraint is fairly firm, whereas the others are subject to relaxation as we gain more experience.

Please send responses to:
Barry Haskell
AT&T Bell Labs - 4c538
Holmdel, NJ 07733 USA
telephone: +1 201 949 5459
telefax : +1 201 949 3697

with a copy to
Mr. Sakae Okubo
NTT Human Interface Labs
telefax: +81 468 59 2812.

Also on this subject - we believe the error constraints could be relaxed somewhat more if an industry - standard IDCT could be defined. Therefore, we invite you to come to a meeting (details attached) to discuss the issues and perhaps define a standard. It would be most useful if such a standard were completed in 1988.

CCITT Specialists Group on Visual Telephony

Proposed Specification for Inverse DCT Chips

1. Generate random integer pixel data values in the range -L to +H. Arrange into 8x8 blocks. Data sets should be generated for (L=256, H=255), (L=H=5) and (L=H=500??).

2. For each 8x8 block, perform a separable, orthonormal, matrix multiply, Forward Discrete Cosine Transform (FDCT) using at least 32-bit floating point accuracy.

3. For each block, round the 64 resulting transformed coefficients to the nearest integer values. Then clip them to the range -2048 to +2047. This is the 12-bit input data to the inverse transform.

4. For each 8x8 block of 12-bit data produced by step 3, perform a separable, orthonormal, matrix multiply, Inverse Discrete Cosine Transform (IDCT) using at least 32-bit floating point accuracy. Round the resulting pixels to the nearest integer, and clip to the range -256 to +255. These blocks of 8x8 pixels are the "reference" IDCT output data.

5. For each 8x8 block of 12-bit data produced by step 3, use the proposed IDCT chip or an exact-bit simulation thereof to perform an Inverse Discrete Cosine Transform. Clip the output to the range -256 to +255. These blocks of 8x8 pixels are the "test" IDCT output data.

6. For each of the 64 IDCT output pixels, measure the peak, mean and mean square error between the "reference" and "test" data.

7. For any pixel, the peak error should not exceed 1 in magnitude.
   For any pixel, the mean square error should not exceed 0.06
   Overall, the mean square error should not exceed 0.02
   For any pixel, the mean error should not exceed 0.015 in magnitude.
   Overall, the mean error should not exceed 0.010 in magnitude.

8. All-zeros in must produce all-zeros out.

9. Rerun the measurements using exactly the same data values of step 1, but change the sign on each pixel.

Meeting of DCT chip manufactures.

Venue: British Telecom Research Laboratories
Martlesham Heath
Ipswich
Suffolk
UK

Date/Time: Start 14:00 24 May 1988
Finish 15:00 25 May 1988

Agenda:

1. Description of DCT specification problem and its impact on standards.

2. Brief presentation from manufactures describing their work concerning DCT chips to date.

3. Discussion of manufactures proposals.

4. Conclusions and preparation of work plan.

Hotel Arrangements/Attendance:

Please let Mike Carr know by Friday 13th May if you wish to participate in this meeting, giving details of Hotel requirements (arrival dates etc).

M D Carr BTRL Martlesham

Telephone: +44 473 644504
Fax: +44 473 643791
Telex: 988811 (VSTEL G)

DRAFT RECOMMENDATION H.12x

CODEC FOR AUDIOVISUAL SERVICES AT nx384 kbit/s

Contents

The CCITT,

considering

that there is significant customer demand for videoconference service;

that circuits to meet this demand can be provided by digital transmission using the H0 rate or its multiples up to the primary rate;

that ISDNs are likely to be available in some countries that provide a switched transmission service at the H0 rate;

that the existence of different digital hierarchies and different television standards in different parts of the world complicates the problems of specifying coding and transmission standards for international connections;

that videophone services are likely to appear using basic ISDN access and that some means of interconnection of videophone and videoconference terminals should be possible;

that Recommendation H.120 for videoconferencing using primary digital group transmission was the first in an evolving series of recommendations;

appreciating

that advances are being made in research and development of video coding and bit rate reduction techniques which will lead to further recommendations for videophone and videoconferencing at multiples of 64kbit/s during subsequent study periods, so that this may be considered as the second in the evolving series of recommendations;

and noting

that it is the basic objective of CCITT to recommend unique solutions for international connections;

recommends

that in addition to those codecs complying to Recommendation H.120, codecs having signal processing and interface characteristics described below should

be used for international videoconference connections.

Note 1: Codecs of this type are also suitable for some television services where full broadcast quality is not required.
Note 2: Equipment for transcoding from and to codecs according to Recommendation H.120 is under study.
Note 3: It is recognised that it is essential to provide interworking between nx384 kbit/s codecs and mx64 kbit/s codecs as defined in the H series Recommendations. Interworking will be on the basis of mx64 kbit/s, where the values of m are under study.

1. Scope

This Recommendation describes the coding and decoding methods for audiovisual services at the rates of nx384 kbit/s, where n is 1 to 5. Possible extension of this scope to meet the objective in Note 3 above is under study.

2. Brief Specification

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

2.1 Video input and output

To permit a single recommendation to cover use in and between 625 and 525 line regions, pictures are coded in one common intermediate format. The standards of the input and output television signals, which may, for example, be composite or component, analogue or digital and the methods of performing any necessary conversion to and from the intermediate coding format are not subject to recommendation.

2.2 Digital output and input

Digital access at the primary rate of 1544 or 2048 kbit/s is with vacated timeslots in accordance with Recommendation I.431.

Interfaces using ISDN basic accesses are under study. (Recommendation I.420)

2.3 Sampling frequency

Pictures are sampled at an integer multiple of the video line rate. This sampling clock and the digital network clock are asynchronous.

2.4 Source coding algorithm

A hybrid of inter-picture prediction to utilize temporal redundancy and transform coding of the remaining signal to reduce spatial redundancy is adopted. The decoder has motion compensation capability, allowing optional incorporation of this technique in the coder.

2.5 Audio channel

Audio is coded according to mode 2 of Recommendation G.722. This is combined with control and indication information and conveyed in one 64 kbit/s timeslot which conforms to Recommendation H.221.

2.6 Data channels

Recommendation H.221 permits part of the 64 kbit/s timeslot carrying the audio to be used for auxiliary data transmission.

Additionally, one of the timeslots normally used for video may be reassigned as a 64 kbit/s data channel. The possibility of further such channels is under study.

2.7 Symmetry of transmission

The codec may be used for bidirectional or unidirectional audiovisual communication.

2.8 Error handling

Under study.

2.9 Propagation delay

Under study.

2.10 Additional facilities

Under study.

3. Source Coder

3.1 Source format

The source coder operates on non-interlaced pictures occurring 30000/1001 (approximately 29.97) times per second. The tolerance on picture frequency is +/-50ppm.

Pictures are coded as luminance and two colour difference components (Y, C_R and C_B). These components and the codes representing their sampled values are as defined in CCIR Recommendation 601.

Black = 16
White = 235
Zero colour difference = 128
Peak colour difference = 16 and 240

These values are nominal ones and the coding algorithm functions with input values of 0 through to 255.

For coding, the luminance sampling structure is 288 lines per picture, 352 pels per line in an orthogonal arrangement. Sampling of each of the two colour difference components is at 144 lines, 176 pels per line, orthogonal. Colour difference samples are sited such that their block boundaries coincide with luminance block boundaries as shown in Figure 2. The picture area covered by these numbers of pels and lines has an aspect ratio of 4:3 and corresponds to the active portion of the local standard video input.

Note: The number of pels per line is compatible with sampling the active portions of the luminance and colour difference signals from 525 or 625

line sources at 6.75 and 3.375 MHz, respectively. These frequencies have a simple relationship to those in CCIR Recommendation 601.

3.2 Video source coding algorithm

The video coding algorithm is shown in generalised form in Figure 3. The main elements are prediction, block transformation, quantisation and classification.

The prediction error (INTER mode) or the input picture (INTRA mode) is subdivided into 8 pel by 8 line blocks which are segmented as transmitted or non-transmitted. The criteria for choice of mode and transmitting a block are not subject to recommendation and may be varied dynamically as part of the data rate control strategy. Transmitted blocks are transformed and resulting coefficients are quantised and variable length coded.

3.2.1 Prediction

The prediction is inter-picture and may be augmented by motion compensation (§3.2.2) and a spatial filter (§3.2.3).

3.2.2 Motion compensation

Motion compensation is optional in the encoder. The decoder will accept one vector for each block of 8 pels by 8 lines. The range of permitted vectors is under study.

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

Motion vectors are restricted such that all pels referenced by them are within the coded picture area.

3.2.3 Loop filter

The prediction process may be modified by a two-dimensional spatial filter which operates on pels within a predicted block.

The filter is separable into one dimensional horizontal and vertical functions. Both are non-recursive with coefficients of 1/4, 1/2, 1/4. At block edges, where one of the taps would fall outside the block, the peripheral pel is used for two taps. Full arithmetic precision is retained with rounding to 8 bit integer values at the 2-D filter output. Values whose fractional part is one half are rounded up.

The filter may be switched on or off on a block by block basis. The method of signalling this is under study.

3.2.4 Transformer

Transmitted blocks are coded with a separable 2-dimensional Discrete Cosine Transform of size 8 by 8. The input to the forward transform and output from the inverse transform have 9 bits. The arithmetic procedures for computing the transforms are under study.

Note: The output from the forward and input to the inverse are likely to be 12 bits.

3.2.5 Quantisation

The number of quantisers, their characteristics and their assignment are under study.

3.2.6 Clipping

To prevent quantisation distortion of transform coefficient amplitudes causing arithmetic overflow in the encoder and decoder loops, clipping functions are inserted. In addition to those in the inverse transform, a clipping function is applied at both encoder and decoder to the reconstructed picture which is formed by summing the prediction and the prediction error as modified by the coding process. This clipper operates on resulting pel values less than 0 or greater than 255, changing them to 0 and 255 respectively.

3.3 Data rate control

Sections where parameters which may be varied to control the rate of generation of coded video data include processing prior to the source coder, the quantiser, block significance criterion and temporal subsampling. The proportions of such measures in the overall control strategy are not subject to recommendation.

When invoked, temporal subsampling is performed by discarding complete pictures. Interpolated pictures are not placed in the picture memory.

3.4 Forced updating

This function is achieved by forcing the use of the INTRA mode of the coding algorithm. The update interval and pattern are under study.

4. Video Multiplex Coder

4.1 Data Structure

Note 1: Unless specified otherwise the most significant bit is transmitted first.
Note 2: Unless specified otherwise Bit 1 is transmitted first.
Note 3: Unless specified otherwise all unused or spare bits are set to '1'.

4.2 Video Multiplex arrangement

4.2.1 Picture Header

The structure of the Picture Header is shown in Figure 4. Picture Headers for dropped pictures are not transmitted.

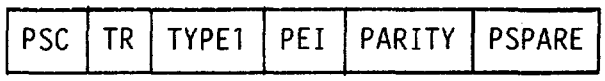


Figure 4/H.12x Structure of Picture Header

Picture Start Code (PSC)

A unique word of 21 bits which cannot be emulated by error free data. Its value is under study.

Temporal Reference (TR)

A five bit number derived using modulo-32 counting of pictures at 29.97 Hz.

Type Information (TYPE1)

Information about the complete picture;

Bit 1 Split screen indicator. '0' off, '1' on.
Bit 2 Document camera. '0' off, '1' on.
Bit 3 Freeze Picture Release. Under study.
Bit 4 Under study. Possible uses include signalling of the use of motion compensation and the method of switching the loop filter.
Bit 5 Number of classes. '0' one, '1' four.
Bits 6 to 12 Under study.

Extra Insertion Information (PEI)

Two bits which signal the presence of the following two optional data fields.

Parity Information (PARITY)

For optional use and present only if the first PEI bit is set to '1'. Eight parity bits each representing odd parity of the aggregate of the corresponding bit planes of the locally decoded PCM values of Y, C_R and C_B in the previous picture period.

Spare Information (PSPARE)

Sixteen bits are present when the second PEI bit is set to '1'. The use of these bits is under study.

4.2.2 Group of Blocks Header

A group of blocks consists of 2k lines of 44 luminance blocks each, k lines of 22 C_R blocks and k lines of 22 C_B blocks. The value of k is under study.

The structure of the Group of Blocks Header is shown in Figure 5. All GOB Headers are transmitted except those in dropped pictures.

Figure 5/H.12x Structure of Group of Blocks Header

Group of Blocks Start Code (GBSC)

A word of 16 bits, 0000 0000 0000 0001.

Group Number (GN)

A m bit number indicating the vertical position of the group of blocks. The value of m is the smallest integer greater than or equal to $\log_2(18/k)$. GN is 1 at the top of the picture.

Note: GBSC plus the following GN is not emulated by error-free video data.

Type Information (TYPE2)

TYPE2 is p bits which give information about all the transmitted blocks in a group of blocks. The value of p is under study.

Bit 1 When set to '1' indicates that all the transmitted blocks in the GOB are coded in INTRA mode and without block addressing data.
Bits 2 to p Spare, under study.

Quantiser Information (QUANT1)

A j bit codeword which indicates the blocks in the group of blocks where QUANT2 codewords are present. These blocks, their codewords and the value of j are under study.

Whether QUANT1 is in the GOB Header or the Picture Header is under study.

Extra Insertion Information (GEI)

Under study.

Group of Blocks Global Motion Vector (GGMV)

Under study.

Spare Information (GSPARE)

Under study.

4.2.3 Block data alignment

The structure of the data for n transmitted blocks is shown in Figure 6. The values of n and the order are under study. Elements are omitted when not required.

Figure 6/H.12x Data structure of transmitted block

Block Address (BA)

A Variable Length Code-word indicating the position of n blocks within a group of blocks. VLC codewords using a combination of relative and absolute addressing are under study.

The transmission order and addressing of blocks are under study.

When bit 1 of TYPE2 is '1' BA is not included and up to 132k blocks beginning with and continuing in the above transmission order are transmitted before the next GOB Header.

Block Type Information (TYPE3)

Variable length codewords indicating the types of blocks and which data elements are present. Block types and VLC codewords are under study.

Quantiser (QUANT2)

A codeword of up to q bits signifying the table(s) used to quantise transform coefficients. The value of q and the codewords are under study. QUANT2 is present in the first transmitted block after the position indicated by QUANT1.

Classification Index (CLASS)

CLASS is present if bit 5 of TYPE1 is set to '1' and indicates which of the four available transmission sequence orders is used for luminance block coefficients. If bit 5 of TYPE1 is set to '0' then luminance block coefficients are transmitted in the default sequence order.

Chrominance block coefficients are transmitted in one sequence order.

The CLASS codewords and sequence orders are under study.

Motion Vector Data (MVD)

Calculation of the vector data is under study.

When the vector data is zero, this is signalled by TYPE3 and MVD is not present.

When the vector data is non-zero, MVD is present consisting of a variable length codeword for the horizontal component followed by a variable length codeword for the vertical component.

Variable length coding of the vector components is under study.

Transform Coefficients (TCOEFF)

The quantised transform coefficients are sequentially transmitted according to the sequence defined by CLASS. The DC component is always first. Coefficients after the last non-zero one are not transmitted.

The coding method and tables are under study.

End of Block Marker (EOB)

Use of and codeword for EOB are under study. An EOB without any transform coefficients for a block is allowed.

4.3 Multipoint considerations

4.3.1 Freeze picture request

Causes the decoder to freeze its received picture until a picture freeze

release signal is received. The transmission method for this control signal is under study.

4.3.2 Fast update request

Causes the encoder to empty its transmission buffer and encode its next picture in INTRA mode with coding parameters such as to avoid buffer overflow. The transmission method for this control signal is under study.

4.3.3 Data continuity

The protocol adopted for ensuring continuity of data channels in a switched multipoint connection is handled by the message channel. Under study.

5. Video Data Buffering

The size of the transmission buffer at the encoder and its relationship to the transmission rate are under study.

Transmission buffer overflow and underflow are not permitted. Measures to prevent underflow are under study.

6. Transmission coder

6.1 Bit rate

The net bit rate including audio and optional data channels is an integer multiple of 384 kbit/s up to and including 1920 kbit/s.

The source and stability of the encoder output clock are under study.

6.2 Video clock justification

Video clock justification is not provided.

6.3 Frame structure

6.3.1 Frame structure for 384-2048kbit/s channel

The frame structure is defined in Recommendation H.222.

6.3.2 Bit assignment in application channel

Under study.

6.3.3 Timeslot positioning

According to Recommendation I.431.

6.4 Audio coding

Recommendation G.722 56/48 kbit/s audio, 0/8 kbit/s data and 8 kbit/s service channel in the first timeslot.

The delay of the encoded audio relative to the encoded video at the channel output is under study.

6.5 Data transmission

One or more timeslots may be allocated as data channels of 64 kbit/s each. The first channel uses the fourth timeslot.

Positioning of the other channels, and possible restrictions on availability at lower overall bit rates are under study. The BAS codes used to signal that these data channels are in use are specified in Recommendation H.221.

6.6 Error handling

Under study.

6.7 Encryption

Under study.

6.8 Bit Sequence Independence Restrictions

Under study.

6.9 Network interface

Access at the primary rate is with vacated timeslots as per Recommendation I.431.

For 1544 kbit/s interfaces the default H0 channel is timeslots 1 to 6.

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

Interfaces using ISDN basic accesses are under study. (Recommendation I.420)

END

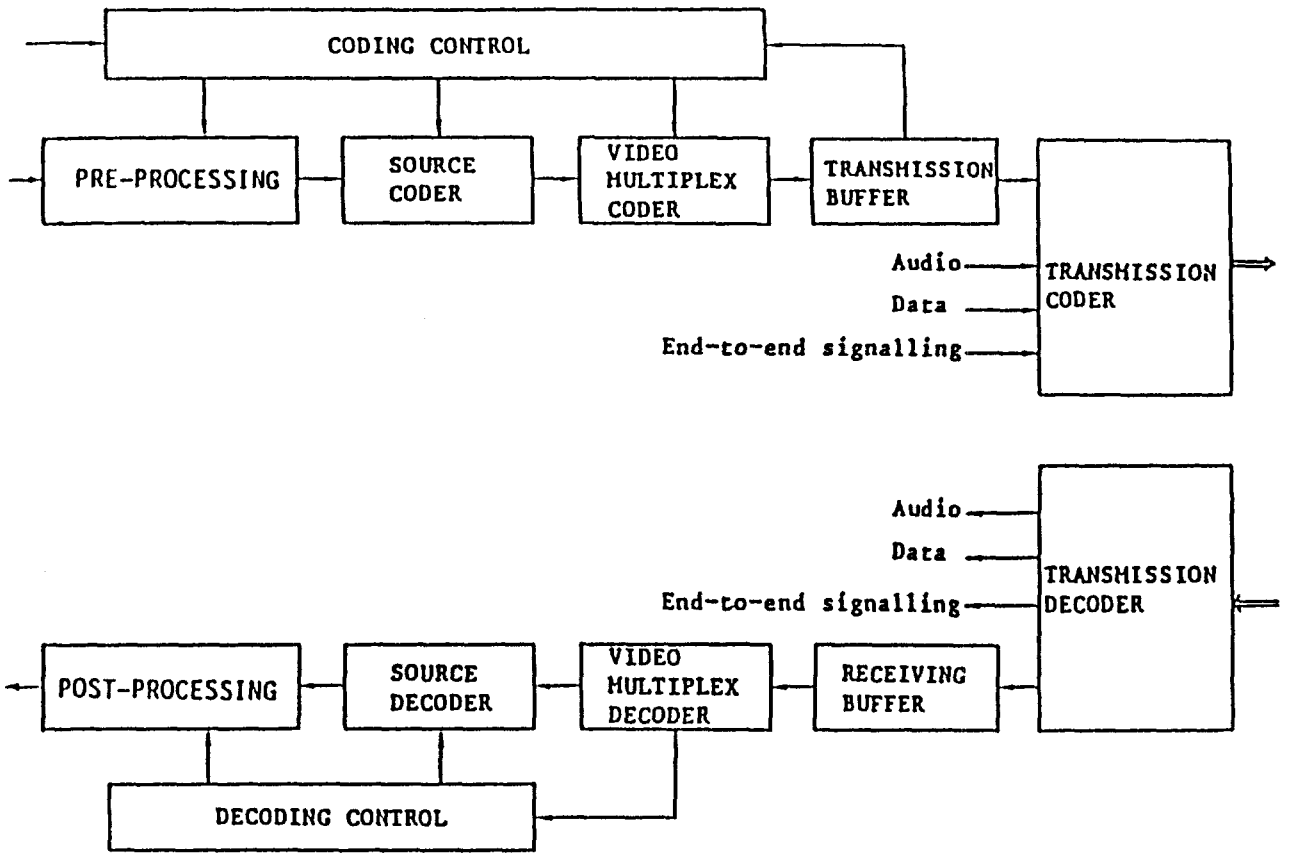


Figure 1/H.12x Outline of block diagram of the codec

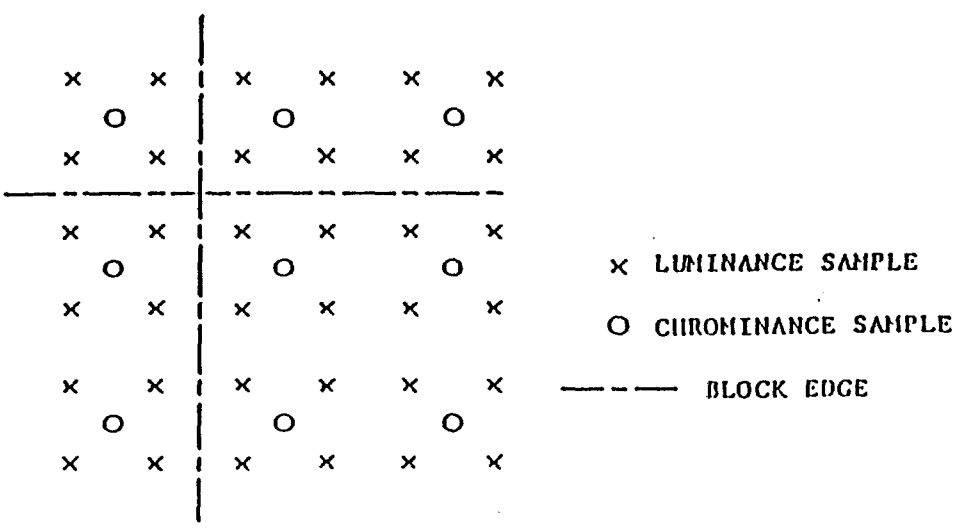
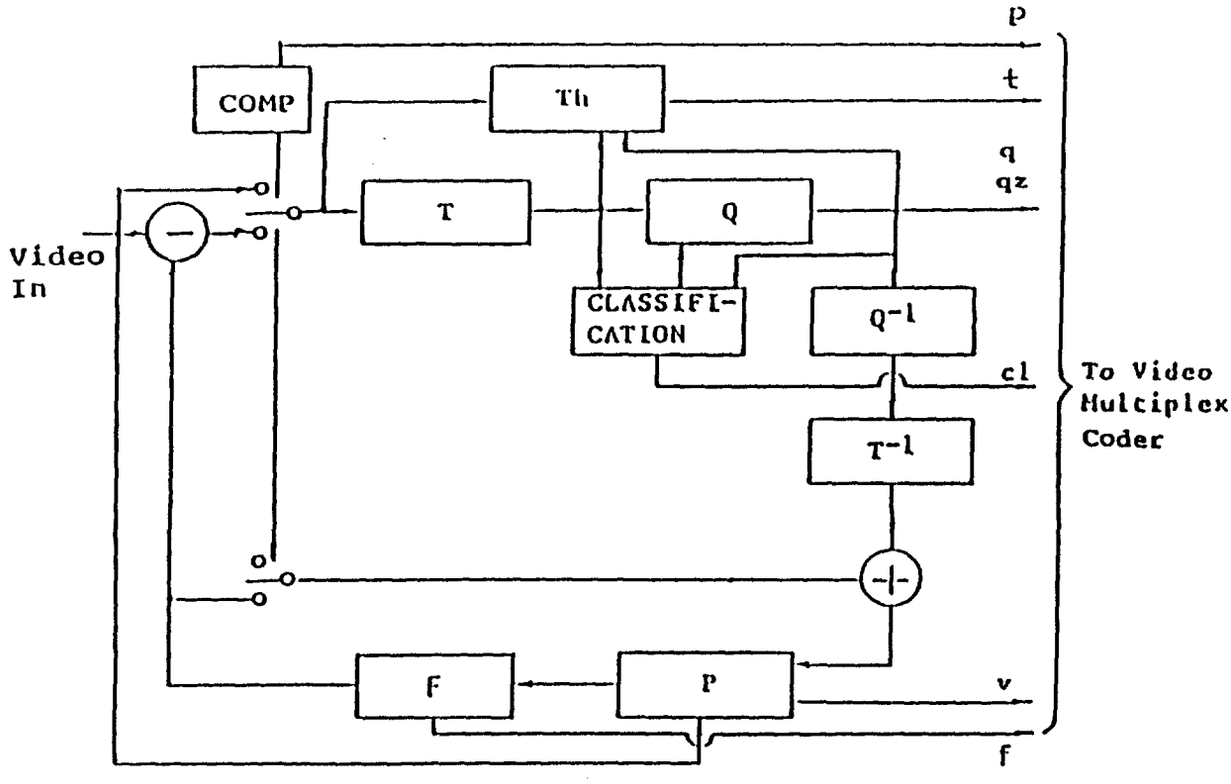


Figure 2/H.12x Positioning of luminance and chrominance samples



COMP Comparator for intra/inter
Th Threshold
T Transform
Q Quantizer
P Picture memory with motion compensated variable delay
F Loop filter

p Flag for intra/inter
t Flag for transmitted or not
q Quantizing index for transform coefficietns
qz Quantizer indication
v Motion vector
cl Classification index
f Switching on/off of the loop filter

Figure 3/H.12x Video coding algorithm

SECTIONS OF H.12x WHERE A COMMON VIEW HAS NOT BEEN REACHED

3.2.2 Motion compensation

On the subject of the range of the motion vectors:

European countries proposed that the range of horizontal and vertical vectors should be +/- 16.

Japan proposed +/- 15 for horizontal components and +/- 7 for vertical.

On the subject of applicability of motion vectors:

Japan proposed that motion compensation be applied only to luminance blocks.

The European countries proposed that the possibility of using vectors on the colour signals in addition to the luminance be kept under study.

REPORT ON ALGORITHM STUDY ACTION GROUP

Participants : F, UK, FRG, N, Japan, USA, I, S, NL
It was agreed that:

* Reference model 5 as described in doc 339 should be adopted as sucessor of Reference Model 4. Figures concerning simulation results are revealed in doc #344 for comparison.

* Improvements on Reference Model 5 are presented on a split-screen basis (see figure 1) i.e. on the left side of the screen RM 5 is displayed, and on the righ side the improvement or method.

* Each improvement or method should be accompanied by a detailed description in order to give the members of the specialists group the opportunity to cross-check the proposed improvement or method.

* Pre- and Post processing are EXCLUDED when presenting the results.

* The test sequence to be used are:

  - Claire (originator CNET)
  - Miss america (originator DIS)
  - Salesman (originator DIS)
  - SWING (originator FI/FTZ)
  - Blue Jacket (originator NTT,NEC, KDD, Fujitsu)

* The minimum frame rate will be at least 10 Hz.

MISCELLANEOUS

All the originators were requested to provide the start position of the displaying window.

AT&T is so kind to provide the meeting with information concerning evaluation criterions.

DIS is so kind to provide the meeting in the case of sufficient advanced hardware a sequence of 15 min.

CLI and PictureTel will investigate the possibility of information disclosure (see point for presentation).

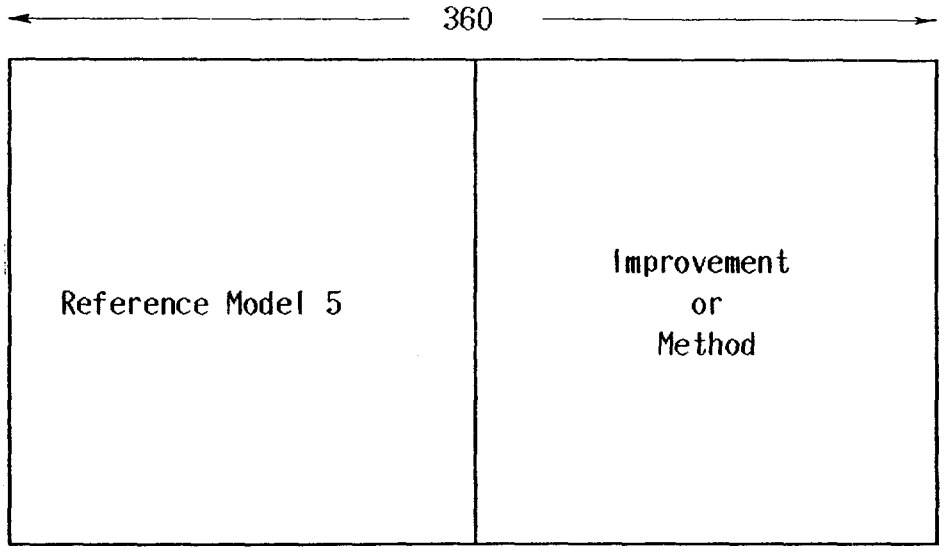
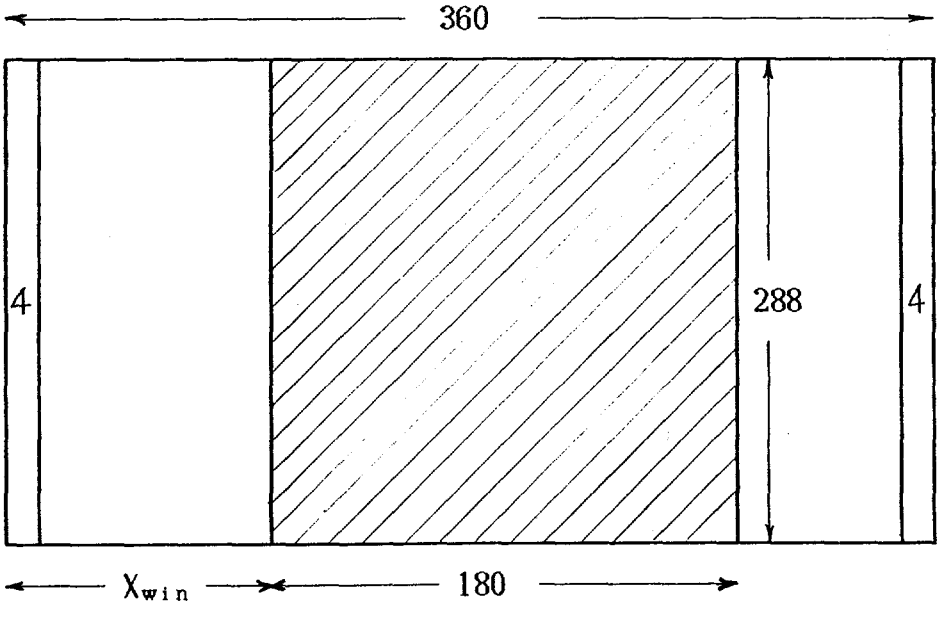


Figure 1



X_{win} = Horizontal position of window

Figure 2

Annex 7 to Doc. #346R

OUTLINE DRAFT RECOMMENDATION H.12y

CODEC FOR AUDIOVISUAL SERVICES AT mx64 kbit/s

1. Scope
2. Brief Specification
3. Source Coder
   - 3.1 Source format

     a. Full CIF: 288 lines x 360 pels x 29.97 Hz*
     b. Format smaller than full CIF: 1/4 CIF,
     144 lines x 180 pels x 29.97 Hz*

     * This temporal aspect is for further consideration.

     Note: The possibility of a third format of 4/9(2/3x2/3)CIF is under study.

     Interconnection between different formats is on the basis of format b.

   - 3.2 Video source coding algorithm
   - 3.3 Data rate control
   - 3.4 Forced updating
4. Video multiplex coder
   - 4.1 Data structure
   - 4.2 Video multiplex arrangement
   - 4.3 Multipoint considerations
5. Video Data Buffering
6. Transmission Coder
   - 6.1 Bit rate
   - 6.2 Video clock justification
   - 6.3 Frame structure

     As per Recommendation H.221.

   - 6.4 Audio coding
   - 6.5 Data transmission
   - 6.6 Error handling
   - 6.7 Encryption
   - 6.8 Network interface

END