CCITT SGXV
Working Party XV/1
Experts Group for ATM Video Coding

Document AVC-355R
October 1, 1992

SOURCE : CHAIRMAN OF THE EXPERTS GROUP FOR ATM VIDEO CODING
TITLE : REPORT OF THE THE EIGHTH MEETING IN TARRYTOWN (Sep.28 - Oct.1, 1992)
Purpose: Report

PART I GENERAL

The eighth meeting of the Experts Group was held at Tarrytown Hilton Inn in Tarrytown, USA, at the kind invitation of David Sarnoff Research Center and IBM. It consisted of two parts;

- CCITT sole sessions : Sep. 28,
- Joint sessions with ISO/IEC JTC1/SC29/WG11 (MPEG): Sep. 29 - Oct. 1.

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

At the opening of the joint sessions, Dr. C. Gonzales made a welcoming address on behalf of the hosting organization. The Experts Group appreciated the support of the hosting organizations for their providing meeting facilities, D1 tape demonstration equipment and secretarial services.

Before starting discussions, we confirmed the following objectives of this meeting;

Sole sessions

- To prepare for the joint sessions in issues of CCITT's concern, particularly low delay, cell loss resilience, compatibility
- To prepare for the Ipswich meeting

Joint sessions

- To elaborate Test Model
- To prepare for the London meeting

PART II SOLE SESSIONS

1. Date

September 28, 13:00 - 19:00

2. Documentation (TD-2)

For the CCITT sole sessions, 34 AVC-numbered documents and 5 temporary documents were made available as in Annex 1.

3. Tape demonstration

The meeting reviewed a number of D1 tape demonstrations as listed in Annex 2.

4. Low delay mode

4.1 Prediction (AVC-327,339,340; TD-5)

One of the main issues regarding the low delay mode is to clarify what predictions are required for the low delay mode since the prediction is a primary source of the coding efficiency improvement in this mode.

Three documents reported experimental results on improvements obtained by "intelligent" predictions such as FAMC, SVMC, DUAL and DUAL'. Mr. Yukitake provided a summary table (TD-5) listing SNR improvements, indicating that those new intelligent predictions bring significant improvements (more than 0.5 dB) in some sequences such as Flower Garden, Mobile & Calendar, Cheer Leaders, compared to Field/Frame adaptive MC. It is a common understanding of the meeting, however, that SNR evaluation alone may mislead us. Particularly, different coding controls produce different pictures without direct correspondence to SNR values.

After having reviewed demonstration tapes, Chairman asked the participants of their impression. Roughly 1/4 of them voted for having observed impressive improvements and 3/4 for not.

The matter was left for impression of the larger population and consideration in the context of the general prediction scheme in the joint sessions.

Chairman raised a related question whether this additional prediction mode can be optional in the encoder and mandatory in the decoder as H.261 motion compensation is. This question is generalized into what modes are inside the "maximum core" and what are outside. The members are requested to study how to apply the generic coding (H.26X) to our communication terminals and feedback to the structure of coding standard.

It was pointed out that in communication applications, two way channel may allow negotiations of the operational mode between two terminals at the start of call. It was pointed out at the same time that B-ISDN applications will need one way distribution in the form of broadcast or multicast.

4.2 Steady state delay (AVC-327,329)

Intra slice/column is compared to intra picture in the two documents. The meeting concluded that the former is clearly favorable in delay performance at the same coding efficiency.

Leaky prediction has been identified as an alternative to smooth out the number of bits generated throughout sequences, but there was no input in this respect.

4.3 Scene change handling (AVC-327,328,330)

AVC-328 showed that if the transmission buffer has a capacity equivalent to twice the frame time (buffering delay = 1 frame time), graceful scene changes are achievable without picture skipping.

AVC-327 showed that picture skipping can maintain low delay during the steady state though some temporal degradation is observed at a scene change.

The meeting concluded that the standard should allow occasional picture in the low delay mode.

AVC-330 discussed the following items to include picture skipping in the standard;

- Buffering (modification of VBV specifications)
- Picture header (to be sent with a flag for skipping)
- Temporal reference (to indicate source picture number)

The meeting accepted the proposals and agreed to put them forward to the consideration of the joint sessions.

4.4 Modified TM specifications (AVC-335, 348)

It has been found that according to the current rate control for the low delay mode the number of bits decreases as time passes. AVC-335 reports that there were two solutions experimented, and they equally worked well.

The meeting concluded that this can be left to the choice of Low Delay Adhoc group, chaired by Mr. T. Yukitake, which may take into account other factors, if any.

AVC-348 provided clarifications for the low delay mode experiment; handling of the first picture for the intra picture case and skipped pictures. The meeting supported this proposal, leaving the skipping order issue for further experiments.

5. Cell loss resilience

5.1 Overview (AVC-351)

Mr. M. Biggar, Chairman of Adhoc group for "ATM, Packet Loss and General Error Resilience", provided a progress report covering possible error resilience techniques and their characteristics.

Our objective is to answer the question of what elements should be included for cell loss resilience in the standard.

5.2 Spatial and temporal localization (AVC-332)

AVC-332 classified cell loss resilient techniques into spatial localization and temporal localization. The meeting agreed to the view that both should be provided by the standard for the single layer coding. Spatial localization techniques are for reducing the corrupted area in a picture, and temporal localization techniques are for reducing the time the corrupted pictures are displayed.

The two localizations may similarly be mapped to the layered coding, but needs further consideration.

5.3 Spatial localization (AVC-332, 333)

Both documents presented the following techniques are effective for spatial localization;

- concealment, and
- structured packing.

The latter requires definition of new syntactic element(s) for Absolute

MacroBlock (AMB) which has an absolute macroblock address and MQUANT, and resets motion vector coding. The choice is left for further study. It should be also studied whether short slice be applicable for the same purpose and how the multimedia multiplex structure gives impact on the applicability of this technique.

5.4 Temporal localization (AVC-332, 333)

Both documents compared

- intra slice or intra picture, and
- leaky prediction

as a method for temporal localization. Though either technique has been proved effective, leaky prediction is concluded a little better as far as recovery from the cell loss is concerned.

It has been demonstrated that by use of both spatial and temporal localization techniques, disturbance due to cell loss of 10E-3 can be controlled to an acceptable extent.

5.5 Leaky prediction (AVC-331, 332, 333, 349, 350; AVC-329)

Various experiments were carried out for the leaky prediction defined in Core Experiment No.6. From coding efficiency point of view, the leaky prediction has been found almost competitive against the perfect prediction with cyclic intra refreshing.

The leaky prediction, however, has two inherent problems. One is residual error due to limited arithmetic accuracy in the prediction loop (limit cycle). AVC-332, 349 provided analysis of the problem, and furthermore AVC-349 demonstrated this problem by coding a still picture and gave a recommended solution by use of a pseudo-random auxiliary signal.

The other problem is noisy background as pointed out in AVC-331, 333. This is caused by the fact that the leaky prediction output converges to 128 as time passes but recovers when a prediction error exceeds a quantizer threshold. A solution for this problem is not yet found.

With respect to channel hopping, the leaky prediction has been confirmed to provide graceful recovery.

Chairman raised whether leaky prediction makes intra refresh to cope with the IDCT mismatch unnecessary. This awaits further study.

5.6 Layered coding (AVC-345)

AVC-345 provided a simulation guideline for layered coding experiments to study its cell loss resilience property. The meeting supported this guideline.

Mr. Dunstan indicated Australia will present experimental results at the next meeting.

5.7 Generic coding and adaptation to media (AVC-351)

H.26X is intended to be generic across a wide range of applications. On the other hand error resilience should be media specific; e.g. ATM networks require cell loss resilience, STM networks require random/burst error resilience.

We should provide answers to the following questions;

- At what level, bitstreams can be interchangeable across applications?
- What side information is required for interchanging bitstreams?

Chairman suggested a case study for transporting picture data retrieved from a disk to a communication terminal through an ATM network . It is felt that an illustrative model will help us.

FEC needs study whether it be applied at the separate level of audio, video, data, etc. or at the multiplexed level. It is envisaged that different transport media need different FECs.

6. H.261/H.26X compatibility (AVC-334,341,347,352)

AVC-334,341 provided experimental results to compare the two compatible coding structures;

- prediction from the base layer, and
- prediction of the prediction error.

From these experimental results as well as from consideration for flexibility of the structure (i.e. relation between the upper layer picture and base layer one is not fixed as far as upconversion methods for the base layer picture is defined), the meeting decided to adopt the prediction from the base layer. This structure allows HDTV/TV compatibility, smooth PAN/SCAN etc. where exact 2:1 relationship of resolution does not appear. Mr. Parke stressed that upconversion can be covered by a generic method.

Other experimental results in AVC-334,347,352 confirmed that embedded coding outperforms simulcast at the total bit rate of 4 Mbit/s.

Chairman requested members to consider whether this functionality be in or outside the "maximum core", in other words, whether all the H.26X coder/decoders should have this functionality. The notion of "flexible layering" has been reminded in this respect.

7. Scalability (AVC-343,344,346,353)

Experimental results on the following topics were shortly reviewed;

- frequency pyramid,
- scalable side information,
- interlace-in-interlace extraction.
- frequency scanning

We did not prepare specific comments of the Experts Group for this topic.

8. Prediction (AVC-337,338; AVC-339; AVC-329,331,333,349,350)

The meeting shortly reviewed available documents for various prediction methods. The questions is what prediction modes are appropriate for the generic standard. This is left for the consideration of the joint sessions (see Section 4.1).

9. Quantization

There was no input from the Experts Group members on this topic.

10. Video clock recovery (AVC-336)

AVC-336 proposed provision of 8 bit field in the picture header for video clock recovery. There were comments that there are cases where a common network clock is not be available at the transmitter and receiver, and other cases where pre-recorded materials are retrieved, thus the common clock can not be used at the time of coding.

The meeting suspended discussion until the next meeting, and decided that AVC-336 be submitted to the joint sessions as an informational document.

11. PSTN videophone (AVC-342)

European countries expressed that it is appropriate to develop standards for the PSTN videotelephony complementing ISDN ones and suggested guidelines. Japan stated that they will submit a paper to the WPXV/1 meeting in November, addressing interworking between PSTN and ISDN videophones. Mr. Schaphorst informed the meeting that T1A1 is establishing a related project subject to the last approval.

The meeting considered what should we recommend to Working Party XV/1 as a continuation of the New Jersey meeting in July. Chairman proposed to recommend organizing a rapporteurs (experts) group for this task, covering both short term and long term study items. This matter will be finalized at the next meeting in Ipswich.

12. Work plan (TD-3)

Mr. Biggar provided comments requesting to set up clear objectives of the Experts Group. Since the framework for H.26X is being formed, it is the time to review and clarify our future work plan. Members are requested to consider this toward the next meeting.

13. Status report (TD-4)

Members are requested to review the draft and return comments to Chairman by October 10.

14. Preparation for the joint sessions

14.1 Documents

All the contributions to this meeting were put forward to the joint sessions except AVC-342. See also Section 10.

14.2 Reporters

The meeting appointed reporters of the joint sessions as follows;

Adhoc Group	Adhoc Chair	CCITT EG Reporter
Overall	D. LeGall	S. Okubo
Low delay	T. Yukitake	T. Yukitake
Cell loss resilience	S. Dunstan	S. Dunstan
Compatibility	A. Puri	I. Parke
Scalability	E. Viscito	O. Poncin
Prediction	H. Watanabe	H. Watanabe
Quantization	N. Wells	G. Eude

Their roles are;

- to reflect the discussion of the sole sessions for consideration of the joint sessions
- to provide a report containing conclusions, discussed items, action points, etc. by Oct. 9 at the latest.

15. Others

1) Editor for the common text

The meeting appreciated that Mr. M. Biggar had volunteered to be an Editor representing the Experts Group.

2) Next meeting

- Ipswich meeting of the Experts Group (Oct. 28 - 30)
- IVS Technical Session in Ipswich (Oct. 26 - 27)

PART III JOINT SESSIONS WITH MPEG

1. Date

September 28 - October 1

2. Documentation

The list of documents considered during the joint sessions is attached as Annex 3.

3. Video Test Model (S. Okubo)

3.1 Work plan

The following was confirmed toward freezing technical specifications at the Sydney meeting in March 1993;

- New technical possibilities will no longer be accepted after the london meeting and the members should focus instead on resolving and refining the core experiments

- In the spirit of convergence process, proposals that are not actively worked on should be dropped from the list of core experiments and resolved by taking the older alternative.

- As many as possible of the members are requested to devote their energy to solve the outstanding problems, e.g. scalability and compatibility. Current activities should be reviewed at a joint Video/Requirements session in London to identify those areas that need more work.

- An additional WG11 meeting will be held during January 25-29, 1993 in Torino.

3.2 Patent statements

Members are requested to submit their patent statements relevant to the current phase of work at an early occasion, if possible by the London meeting.

3.3 Profile

We initiated to study how to structure the generic standard. "Profile" has been thought to assist this work, but it needs clear definition for our particular case. A summary of discussion on this topic is contained in Annex 4. Use of the word "profile" should be avoided until clarification is obtained.

3.4 Liaison toward CCIR

Since CCIR TG11/4 is going to meet during October 13-15, and it will deal with comments on the requirements listing, it was decided to send a provisional updated version of the integrated requirements listing which is now being worked out by Adhoc Group established in Rio.

4. **Low delay** (T. Yukitake)

4.1 General

The group met a couple of times in this week, one is joint meeting with prediction ad-hoc group and another is for itself.

4.2 Prediction modes for low delay mode (Core experiment No.1; Joint meeting with prediction group)

- What prediction modes are appropriate for the low delay mode? We need special prediction modes (S-FAMC,SVMC, Dual') or not?

1) Among the three coding structures preferable for low delay profile, field structure with M=2 has the best performance and frame structure with M=1 is the next and field structure with M=1 the last. The coding performance can be improved at the cost of basic delay.

2) In terms of SNR, the introduction of special predictions can improve the coding efficiency. It is very desirable for low delay, because we can save the basic delay and hardware cost to get the better image quality. For example, the SNR of field structure M=2 with fi/fr prediction is as that of frame structure M=1 with special prediction. By introducing the special prediction, we do not need to use the bi-directional prediction and it makes the hardware simple and delay low.

There is a discussion about the necessity of special prediction mode, and we can not get the firm conclusion. However, the feeling of the meeting seems to be positive.

3) These three special predictions seem to similarly solve the same problems. Therefore we should have an effort to unify these predictions or select one of these. To unity them or select one prediction, a new core experiment (L10 defined in the prediction ad-hoc group) is carried out towards the London meeting.

4.3 Intra picture vs Intra slice/column (Core Experiment No.2)

1) The coding efficiency of intra slice and intra picture is almost the same in terms of SNR.

2) The delay of intra slice is more than 10 times shorter than that of intra picture.

3) The coding efficiency of intra slice and intra column is almost the same in terms of SNR.

4) In case of low delay profile, the low delay ad-hoc group recommends to use intra slice or column in stead of intra picture.

4.4 Introduction of skipped pictures (Core experiment No.3)

1) Picture skipping is a useful technique to keep the delay law at the cost of the temporal smoothness just after the scene change. The introduction of skipped picture expands the freedom of encoder control, so the low delay ad-hoc group recommends to introduce the picture skipping.

2) In order to keep low delay except for transient periods, the decoder should identify dropped pictures so that it does not take time to display them. We have two choices ;

- to send only their picture headers
- to send no data at all

In case of the picture header transmission, we need 1 bit indication in the picture header for distinction between the case of picture skipping and the case where a picture is coded with almost no coded data.

We discussed these two methods, and we concluded that we did not have a strong opinion to choose one of two. This item is still open now. Some discussions are needed at the London meeting.

3) In MPEG-1, TR is defined as indicating the display picture order. In case of low delay profile, it should be defined as indication of the source picture order, because some pictures are not displayed. It is noted that if all pictures are displayed, the source picture order is equivalent to the display picture order.

4) To allow picture skipping, VBV specifications were modified.

4.4 New experiments

Core experiment No.1 concerning prediction modes is re-defined as L10. A new experiment, which is concerned with the skipped picture order in field structure with M=2, is defined.

5. Cell loss resilience (S. Dunstan)

5.1 Introduction

The second meeting of the Ad-hoc Group on ATM, Packet Loss and General Error Resilience was held at Tarrytown in September/October.

5.2 Transmission media

One of the Ad-hoc group's aims is to recognize at an early stage what syntax changes, if any, might be required to support MPEG over different transmission media. Which media are of interest need to be recognised, along with their error characteristics. Currently only B-ISDN ATM is being investigated and mention made of satellite transmission.

5.3 Inputs

There were five inputs relevant to the Tarrytown meeting, concerning

- leaky prediction and cell loss error recovery,
- consideration of layered coding cell loss resilience.

1) Error localisation

Two documents classified the localization of the effects of cell loss as

- spatial localization
- temporal localization

Spatial localization relates to the number of macro blocks before resynchronization of the decoder occurs. Two possibilities to control this are;

- resynchronize on the next slice start code
- resynchronize at the next macro block having an absolute address (structured packing)

Results have shown that structured packing results in only a small change in coding efficiency. However, future work is required to investigate;

- syntax changes required to implement new macro block types having an absolute address,
- implementation of structured packing. It is not clear how alignment of the structured packing format could be achieved in for example AAL type 1.
- the impact of the System Layer upon structured packing method

Temporal localization relates to the number of frames that an error is allowed to propagate. Two possibilities of control are;

- intra frame or intra slice,
- leaky predictor.

Two documents conclude that leaky prediction provides superior subjective performance to that of intra slice, though further refinement is required to improve the image quality in stationary conditions.

The performance of leaky prediction as a method of temporal localization of errors should be further tested.

2) Layered coding

A document discussed the design of an experiment to allow a fair comparison between the cell loss resilience performance of a layered coder to that of a single layer coder. Possible relationships between the two cases, when both layers of the layered coder were subject to cell loss, were presented.

5.4 Generic coding and media adaptation

The MPEG-2 standard provides for generic coding. Specific media may require specific adaptation, e.g. prioritised encoding of the MPEG bitstream. Transfer of an MPEG bitstream across different media should avoid coding and recoding of the bitstream.

A system model is required to identify demarcation between generic coding and media specific adaptation.

The impact of the System Layer should be considered. For example the audio and video signals may require different adaptation. Also the layers of a scalable coder may require separate virtual channels on B-ISDN.

5.5 Frequency scanning

It is noted that the frequency scanning method may have desirable cell loss resilience properties, because of the ordering of the DCT coefficients within the slice. The application of priority to different parts of the slice components is simple in this method.

Further work should be done to investigate the error performance of frequency scanning methods.

5.6 Future work

The following are proposed as core experiments;

- cell loss resilience of frequency scanning method,
- evaluation of leaky prediction as a temporal error localization method, where sophisticated concealment techniques are used.

Results of the following work will be presented at the London meeting;

- structured packing versus slice resynchronization with concealment,
- refinement of leaky prediction method,
- performance of layered coding versus single layer coding under conditions of cell loss.

5.7 Request for information

Information is requested from the appropriate bodies concerning the following;

- error characteristics of satellite transmission,
- impact of Systems Layer on media adaptation,
- other media of interest.

6. **Compatibility** (I. Parke)

The relevant documents in this group were MPEG92/ 421,430,458,462,464,465,485,492,495,506,509.

The discussion focused on three topics, review of core experiment results, improvements to compatible prediction modes and interlace-to-interlace conversions.

The core experiment on comparison of prediction of prediction error and prediction from the base picture was reported in documents 92/430, 92/458 and 92/492. These showed that prediction from the base picture was more efficient though in the case of field structured pictures the difference was small. The group decided to adopt the prediction from the base picture as the compatible prediction method.

Improvements to the compatible prediction modes are documented in 92/485, 92/495 and 92/506. The compatible prediction mode was improved by weighting with the normal TM2 prediction. The weighting used in 92/485 and 92/495 was a simple averaging. Document 92/506 described a more general structure with more weightings. The technique has been given the name 'spatial-temporal weighted prediction'. The spatial prediction is from Mpeg-1 and the temporal prediction is the TM2 motion compensated predictions. There is a proposed core experiment to study suitable weightings.

The group agreed to unify the syntax for the compatible experiments. This has led to a proposal for a new set of macroblock type tables for when a macroblock is coded compatibly. There is also a 2 bit weight code associated with compatible macroblocks.

There was discussion on other syntax issues such as pan vectors, windowing and upconversion tables. The group was unable to resolve these issues. An interim solution was to define a code word with 3 fields that indicated the compatible standard, the picture format and the subsampling ratios. The group will continue to work on improving this.

On interlace to interlace compatible coding. Columbia University presented results on motion compensated deinterlacing. They had tried using the low resolution vectors for this. Their conclusion was that this did not give acceptable results and were now looking to use the high resolution vectors.

A further technique on interlace to interlace conversion is described in 92/509. Here a 'spatial-temporal' technique is used without motion compensation. This gave better results than the existing technique documented in TM2. A core experiment has been proposed to compare the technique of 92/509 with that in TM2. Further solutions are also requested.

The group has much to do. The core experiments proposed are to improve the compatible prediction modes and to find a good solution for interlace-to-interlace compatibility.

7. Scalability (O. Poncin)

7.1 Status Report

Ten core experiments on scalability had been defined during the Rio meeting. They are contained in the appendix I of TM2. The status of each of them was reviewed at Tarrytown.

Among those 10 core experiments:

- one (I.2) was completed and led to a new core experiment proposal (I.12),
- one (I.5) was not yet started,
- the remaining 8 core experiments are in progress, results of them are expected by the London meeting.

The main topics which the scalability ad-hoc group is dealing with are;

- to extract an interlaced downsampled signal from an interlaced source,
- to reduce the drift effect in the lower resolution layers,
- to improve the coding efficiency in as many layers as possible,
- to adapt the amount and the accuracy of the transmitted information layer by layer,
- to define a rate control layer by layer,
- to compare block scanning coding and frequency scanning coding.

The prediction, scalability and compatibility ad-hoc groups met jointly on Wednesday morning (from 11 to 13). Two recommendations were issued by the scalability group.

The cell loss resilience, scalability and compatibility ad-hoc groups met jointly on Thursday morning (from 9 to 11). The frequency scanning technique was presented to ATM people. This technique was found promising for cell loss resilience; a new core experiment on that topic was decided. Other people expressed their wish to achieve the same goal by keeping the usual block scanning technique.

7.2 Recommendations

The scalability ad-hoc group recommends that MPEG/Video consider;

1) the proposal for core experiment I.11 concerning the comparison of several codec options which represent different method of frequency domain scalable coding,

2) the proposal for core experiment I.12 concerning the use of adaptive inter-scale prediction in a situation where lower layer rate is controlled.

The scalability ad-hoc group recommends that MPEG/field-frame prediction ad-hoc group;

1) study the use of special prediction modes in scalable systems

2) clarify the level at which motion compensation modes will be specified or adapted.

7.3 Statement on convergence in scalability

The requirement for very high quality in all layers of a scalable system may conflict with the desire for low implementation complexity. This has led to the investigation of two classes of solutions. The possibility of an unified solution is still being considered.

8. Prediction (H. Watanabe)

8.1 Introduction

Many documents and video tapes were presented in the field/frame prediction adhoc group at the Tarrytown meeting. This report describes the results of discussions and conclusions of the adhoc group.

8.2 Document list related to the prediction group at the Tarrytown meeting

422	Watanabe	Meeting announcement, agenda and call for contributions
432	Yukitake	Simulation results on prediction and DCT mode coupling for S-FAMC
433	Yukitake	Simulation results on S-FAMC and Dual'
434	Noguchi	Results of some prediction experiments
437	Nagata	Results of Core Experiments on Simplified FAMC, Dual-Prime and
438	Nagata	Test Model Simplification
440	Takahashi	Simulation Results on Global Motion Compensation Core Experiment
441	Sugiyama	Results of Prediction Experiments
442	Sugiyama	Results of Leaky Prediction Experiments
448	Yagasaki	Simulation results on TM2 core experiment of prediction mode No.8
453	Nakajima	Results of prediction core experiments on TM2 (SFAMC, SVMC, DUAL')
456	Odaka	Simulation results on prediction core experiment No.3 (Dual-prime) in TM2
459	Kameyama	Comparison of Prediction Modes and Simplified Test Model
461	Yukitake	Clarification of Appendix H in TM2
463	Yu	Results of core experiments H2.1 and L7
467	Nishikawa	Core experiments: TM prediction methods and simulation results
471	CCITT/J	Coding efficiency of leaky prediction
493	Reibman	Leaky prediction: Eliminating the limit cycle
494	Reibman	Leaky prediction: Experiments results
499	Corset	Additional tests on FAMC compensation mode
505	Madec	Results on leaky prediction experiments
510	Watanabe	Intermediate report of the adhoc group on field/frame prediction experiment
511	Paik	A proposal for switching the coding mode (Intra/Inter) on a block basis
512	Paik	A proposal for specification of 8x8 motion vectors
518	Savatier	Simulation results on prediction modes
519	Savatier	Reference fields for forward prediction of p-fields
520	Savatier	Correction to test model
521	Savatier	Frame prediction in field-pictures
523	Wong	TM2 Errors
536	Koster	Tarrytown recommendations of test model editing

8.3 Allocation of the contribution

8.3.1 Core experiments

L.1 Simplified FAMC;	431, 432, 433, 434, 437, 459, 499
L.2 SVMC;	453, 518
L.3 Dual';	433, 456, 459, 467
L.4 Global MC;	440
L.5 Leaky prediction 1;	442
L.6 Leaky prediction 2;	471, 493, 494, 505
L.7 Reverse order prediction;	463,
L.8 Simplification of Test Model;	448, 459, 467
L.9 16x8 MB;	441

L.10 Special prediction mode	Annex 5.3
L.11 8x8 Motion vectors;	512
L.12 8x8 Intra/Inter	511

8.3.2 Others

Document #422, 510, 519, 520, 521, 523

8.4 Circulated e-mails after 8/15

L.1 Simplified FAMC; Koster(8/26),Savatier(8/26,8/28,9/1),Yukitake(9/1)
L.2 SVMC; Yukitake(8/21),Savatier(8/28)
L.8 Reverse order prediction; Anastassiou (9/23)
L.9 Simplification of Test Model; Yagasaki(9/22)

8.5 Discussion

8.5.1 Special Predictions (Simplified FAMC, SVMC, Dual')

- Joint meeting with Low delay group (See Annex 5.1)
- Three candidates seem to have similar improvement, effective to constant vertical motion
- Difference from field/frame adaptive is not definitely significant by tape-viewing
- Purpose is recognized to supply high performance mode for low delay application
- Agreed to merge three into single mode
- Prepare new core experiment (L10, see Annex 5.3)
- We will choose one among five sub-candidates in L10 core experiments in terms of the subjective quality and hardware complexity in the future meeting. Simulation at M=1 has a priority. Test should be carried out for field-picture and frame-picture.

8.5.2 Global MC

- Tape demonstration by Matsushita
- 0.2 to 0.3 dB improvement than I-P structure in I-picture
- needs more experimenter

8.5.3 Leaky prediction 1

- Tape demonstration by JVC
- Role of this prediction is a loop filter, different from error resilience purpose
- needs more experimenter

8.5.4 Leaky prediction 2

- Tape demonstration by AT&T, CCETT
- Limit cycle problem is solved by AT&T proposal
- Intra slice has advantages than leaky prediction in terms of coding gain
- Delay between leaky prediction and intra slice is not compared
- Ask ATM group how to decide to include it into TM2

8.5.5 Reverse order prediction

- Tape demonstration by Columbia Univ.
- Field coding with Dual gives sufficient quality
- Dual can be replaced by SVMC or Dual'

- Needs more experimenters

8.5.6 TM Simplification

- Tape demonstration by SONY.
- Coupling of MC mode and DCT mode gives simplicity
- Information to define "decoder's level" (TM2, p.124) according to the hardware complexity

8.5.7 16x8 MB

- Tape demonstration by JVC.
- Needs unified syntax in TM2
- See Annex 5.2.

8.5.8 Specification of 8x8 motion vectors

- New core experiment (L11) (MPEG92/512)
- Decision rule; first 16x16Inter or 8x8Inter, next Inter or Intra, both a posteriori decision
- A posteriori decision needs more discussion at the London meeting
- Necessary information for field coding will be provided by Mr. Paik (wpaik@gi.com)

8.5.9 8x8 Inter/Intra decision

- New core experiment (L12) (MPEG92/511)
- Current intra macroblock can be partly inter-coded
- Intra_block_pattern is used to specify intra block in a macroblock at the macroblock layer

8.5.10Answer to TM2 editorial group

- See Annex 5.2.

8.5.11Joint meeting with Compatibility and Frequency Scalability group

1) Prediction and compatibility

- Compatibility, Spatial Scalability group may use either field-picture or frame-picture
- 16x16 / 16x8 motion compensation block is desirable for field coding
- Field/frame adaptive prediction can be used for frame-picture case
- Needs complete syntax for 16x8 macroblock in TM2

2) Prediction and frequency scalability

- Frequency scalability group use frame-picture, and field-frame adaptive prediction.
- New high performance prediction needs to be checked whether it fits several layers or not
- A level of the layer that determines prediction modes should be notified to the scalability group
- Current level is macroblock layer

8.6 Summary

- Core Exp. L1, 2, 3 (FAMC, SVMC, Dual') should be merged to a new core experiment (L10).

- Core Exp. L4, 5, 7 (Global MC, Leaky prediction 1, Reverse order prediction) needs another experimenter.
- Core Exp. L11, L12 (8x8 motion vector, 8x8 intra/inter) are new core experiments.
- Core Exp. L6 (Leaky prediction 2) should be treated in ATM group.
- We fixed some ambiguity of TM, but still need more elegant syntax.
- For Core Exp. 10, members are requested to consider a matching feature with scalability and compatibility.

8.7 Recommendations of Frame/Field prediction adhoc group

- Core Exp. L1, 2, 3 (FAMC, SVMC, Dual') should be merged to a new core experiment (L10). The adhoc group will choose one among five sub-candidates in L10 core experiment in terms of the subjective quality and hardware complexity in the future meeting. Simulation at M=1 has a priority. Test should be carried out for field-picture and frame-picture. Members are requested to consider a matching feature with scalability and compatibility.

- Core Exp. L4, 5, 7 (Global MC, Leaky prediction 1, Reverse order prediction) needs another experimenter.

- Core Exp. L11, L12 (8x8 motion vector, 8x8 intra/inter) are identified as new core experiments.

- Core Exp. L6 (Leaky prediction 2) should be treated in ATM group.

- The adhoc group will fix ambiguities of the description for prediction in TM.

- No more core experiment will be accepted after the London meeting.

The current core experiments are;

L.4 Global MC;
L.5 Leaky prediction 1;
L.6 Leaky prediction 2;
L.7 Reverse order prediction;
L.8 Simplification of Test Model;
L.9 16x8 MB;
L.10 Special prediction mode
L.11 8x8 Motion vectors;
L.12 8x8 Intra/Inter;

9. Quantization (G. Eude)

The Ad-hoc group on quantisation had several sessions during its meeting in Tarrytown and 18 documents were presented. The results of the different core experiments have first been reported and then compared in order to propose modified or new core experiments for the London meeting. Accompanying video tapes were viewed.

The group also discussed on the way to decide to include or to discard new techniques. One proposal was that new technique should satisfy at least one of the following features;

- visibility better quality,
- bit saving around 10%,
- less hardware,

- syntax for generic standard.

The discussions during the meeting showed that not many proposals meet this rule(!). Further considerations seem needed.

The technical topics were the followings:

1) New proposal "vector quantisation"

In Doc MPEG/92/525, a "vector quantisation" method is proposed to transmit the transformed coefficients. A vector is chosen to best fit the quantized coefficients with a "cost function". The coefficients which belong to the selected vector are 1-D variable length encoded.

It has been pointed out that the matching method could be adapted to perceptual criteria. A new core experiment has been defined. As this technique implies dropping of coefficients, TM2 will be compared by using the same dropping.

2) VLC proposal

The effectiveness of an alternative VLC for the INTRA coded pictures has been demonstrated - with MQUANT less than 8, the bit saving is about 10% for the INTRA coded pictures (MPEG/92/452).

A second new proposal (MPEG/92/427), correlated to the "vector quantisation" coefficients coding, consists in using switched 1-D VLCs at the slice level. A core experiment will be defined.

Modified UVLC, as proposed in MPEG/92/504 has also high efficiency in I pictures. The other advantages of this technique seems to be: frequency sclalability, SNR scalability and self-adaptability. A higher hardware cost has to be checked by the implementation group.

In MPEG/92/450 it is concluded that CBP and "first coefficient trick" are not necessary (the gain is less than 0.1 dB for all the sequences).

3) Range extension and precision

- INTRA DC: it has been agreed that some applications need to increase the precision of the intra DC to 9 bits. Requirement for more than 9-bit precision seems related to more than 8-bit input signals (papers for discussions are requested).

- transmitted coefficients (MPEG/92/449): the results presented show that the range of +/-256 is exceeded on all sequences with MQUANT=1 or 2. Effect on bitrate is very small (MPEG/92/435). Current TM2 solution was accepted to fix this issue.

- MQUANT: the ad_hoc group reached a consensus to recommend an alternative 5-bit nonlinear law for the control of quantization stepsize (as described in MPEG/92/508). The law is fixed, simple in definition and selected at the picture layer (hardware impact should be checked).

4) Scanning (MPEG/92/435, 436, 460 & 480)

Considering the improvements obtained by adaptively scanning according to DCT mode (max in SNR +0.3 dB which was impossible to see on the pictures), the ad-hoc group recommended to drop this mode from the TM.

Zigzag/vertical scan adaptation experiments showed some improvements particularly for field-based coding. It has been concluded that this mode needs a suitable "a priori" decision criteria and a strong support to be considered further.

5) Non 8x8 DCT block coding

- DCT blocksize adaptation: the results obtained by using adaptive 8x8/8x1 DCT selection give up to 0.5 dB improvement (MPEG/92/435, 444 and 445). It has been agreed that 8x4 DCT adaptation, which does not work as well as 8x1 DCT, will not be considered further. Comparisons with non-DCT modes are needed on new sequences including vertical and horizontal scrolling text (as it can be seen at the end of the movies).

- Non transform coding: the two proposals NTC1 and NTC2 (MPEG/92/451 & 478) gave improvement in SNR and in subjective picture quality as well (in particular in moving text). The main issues are the prediction method and the quantisation (linear or not). A refinement of the existing core experiment was defined in order to be able to reach only one compromise solution. For this core experiment new sequences including text will be used. In order to be sure that "new tools" are really needed, the performances must be compared with the use of MQUANT (and/or BQUANT) control based on the same "edge-detection" criteria.

6) quantization

BQUANT: Results does not show improvement by using BQUANT and more results are requested.

Weighting matrix: there was no result on this topic. New matrix descriptions and selection criterion will be provided by AT&T (new inputs are needed to keep this option in TM).

7) List of "new" or "modified" core

- vector quantization:TCE, CNET, AT&T, GI, PHILIPS, Matsushita
- VLC adaptation on MB basis: AT&T, GI, Sony, TCE
- modified UVLC: HHI, Belgacom, UCL, Siemens
- non-transform coding: Sony, Matsushita, AT&T, JVC, University Hannover
- MQUANT control: GCT, Sony, AT&T, University Hannover, CNET.

END

Participants of the eighth meeting of
Experts Group for ATM Video Coding
(28 September - 1 October 1992, Tarrytown)

Australia	J. Arnold	University of New South Wales	
	S. Dunstan	Siemens Ltd	(CM)
	T. Sikora	Monash University	
Belgium	Mr. O. Poncin	Belgacom	CM
USA	Mr. B.G. Haskell	AT&T Bell Labs	
	Mr. S. Kumar	Wiltel	
	Mr. A. Luthra	Tektronix	
	Mr. D. Hein	VideoTelecom	
	Mr. D. Raychaudhuri	David Sarnoff	
	Mr. R. Schaphorst	DIS	CM
	Mr. A. Tabatabai	Bellcore	CM
	Mr. X. Yuan	PictureTel	
France	Mr. G. Eude	CNET	
	Mr. J. Guichard	CNET	CM
Japan	Mr. K. Asai	Mitsubishi	
	Mr. W. Fujikawa	Matsushita	
	Mr. Y. Nakajima	KDD	(CM)
	Mr. S. Nogaki	NEC	
	Mr. S. Okubo	NTT	Chairman
	Mr. K. Sakai	Fujitsu	
	Mr. H. Watanabe	NTT	(CM)
	Mr. T Yukitake	Matsushita Communication	
Norway	Mr. H. Sandgrind	NTA	CM
Netherlands	Mr. A. Koster	PTT Research	(CM)
UK	Mr. I. Parke	BT	(CM)

CM: Coordinating Member
(CM): Substitute for CM

Documents for the Tarrytwon Meeting
(Sep. 28 - Oct. 1, 1992)

Normal Documents

[ ] indicates MPEG92/???.

AVC-317R Report of the seventh meeting in New Jersey and Rio de Janeiro - July 1992, Part 1 (Chairman)
AVC-318R Report of the seventh meeting in New Jersey and Rio de Janeiro - July 1992, Part 2 (Chairman)

These two documents report the outcome of the previous meeting in New Jersey and Rio de Janeiro.

AVC-319 to be considered in Ipswich
AVC-320 to be considered in Ipswich
AVC-321 to be considered in Ipswich
AVC-322 to be considered in Ipswich

AVC-323 [N0245] Test Model 2 (Test Model Editing Committee)
AVC-326 [413] TM2 Erratum (Test Model Editing Committee)

These two documents give specifications for the Test Model algorithm and all agreed experiments.

AVC-324 [229rev] Information on requirements for MPEG-2 video (Requirements Group)
AVC-325 [230rev] Guide for the video work, Rio revision (Requirements Group)

These two documents are the outcome obtained in Rio de Janeiro regarding video requirements.

AVC-327 [431] Results of low delay core experiment (Japan)

Low delay core experiments are carried out with TM1, mainly adopting FAMC as prediction mode and Frame M=1 as image structure. And the first picture is treated as an intra picture. Through these experiments, it has been confirmed that the techniques of forced intra slice and picture skipping are useful in terms of delay time.

AVC-328 [469] Scene change handling without picture skipping in low delay mode (Japan)

In order to reduce buffer delay, forced intra slices is a very efficient technique. When scene change takes place, however, temporal or spatial picture quality of the first picture in the new scene is degraded. Picture quality degradation of the coding 'without picture skipping' is considered. The simulation based on TM1 shows that bit allocation for a I-picture influences the picture quality of the I-picture and following P-pictures. However, the influence fades out in about 30ms (10frames). In the worst cases, visual degradation is detected. It is recommended in the low delay coding system without picture skipping that target bits for I-picture should be larger than twice of target bits for P-pictures.

AVC-329 [455]   Intra slice/column and leaky prediction (Japan)

Intra Slice and Intra Column methods are compared, and only a minor difference was found. Leaky prediction method is also studied, and SNR degradation curve for various leaky factor values is obtained using TM2. From the view point of SNR, LF=1-1/16 is concluded as the minimum for sufficient picture quality.

AVC-330 [470]   Buffering for low delay mode (Japan)

This contribution analyzes the buffering problem for the low delay mode (M=1 frame coding and M=2 field coding) where occasional picture dropping takes place to cope with scene changes or other forced update situations. An extension of the existing VBV specifications is proposed to keep the steady state low delay operation by momentarily sacrificing temporal reproduction of the coded pictures. It is also proposed that the headers of dropped pictures are sent with a flag indicating picture dropping and that Temporal Reference (TR) is numbered according to the source picture.

AVC-331 [471]   Coding efficiency of leaky prediction (Japan)

Coding efficiency of the leaky prediction as defined in Core experiment No.6 of TM2 has been measured by an open loop experiment with fixed step sizes. Data for the information generation and macroblock type distribution are provided against the leak factor. Coding loss due to leaky prediction is competitive with the use of periodic intra pictures. Coded pictures are also observed. It is concluded that if we can find a means to remove artifacts peculiar to leaky prediction, such as busy or blocky backgrounds, it will contribute to making the coding algorithm simple and robust, hence generic.

AVC-332 [472]   Cell-loss compensation scheme (Japan)

A cell-loss compensation scheme, which consists of leaky prediction, structured packing and concealment at the decoder, was introduced at the previous meetings. In this paper, the effects of each element, especially leaky prediction, are evaluated in detail through computer simulation for TM2 IP mode and low delay mode. The results indicate that three elements work cooperatively and are effective for cell loss. It is proposed that the standard should provide both the spatial and temporal domain compensation schemes. It is also pointed out that the 8 bit arithmetic accuracy for the leaky prediction can produce residual error of +/- (2**n-1) at maximum for LF=1-1/2**n.

AVC-333 [473]   Experiments on cell loss resilience (Japan)

This document discusses two techniques which should be considered to realize cell loss resilience, i.e.

1) Spatial localization of lost macro block
2) Temporal localization of degraded image

For temporal localization, intra picture, intra slice and leaky prediction have been compared. Leaky prediction gives the most favorable impression for the recovery of degraded image, however we have to improve the image quality at the stationary state.

AVC-334 [458]   Simulation results of compatibility core experiment (Japan)

In this document, prediction from the base layer mode and prediction for the prediction error mode were compared again within the TM2 framework. The same tendency as that of the previous document (MPEG 92/257) was obtained. As a conclusion, this document describes that the prediction from the base layer with a switchable compatible type for each field is better than the prediction for the prediction error as a candidate for the compatible mode from the coding efficiency point of view.

AVC-335 [474] Proposal of TM2 rate control modification for low delay mode (Japan)

The current TM2 rate control for the low delay mode makes the number of bits per picture decreases as time passes by, thus SNR is degraded toward the end of sequence. This document provides two alternative solutions for this problem based on experiments.

AVC-336 [475] Picture header modification for source clock recovery (Japan)

Modification of picture header to include a new field for video source clock recovery is proposed. Source clock frequency measured with a network clock is transmitted using this field.

AVC-337 [432] Simulation results on prediction and DCT mode coupling for S-FAMC (Japan)

This document addresses the S-FAMC prediction and DCT mode coupling for inter macro block, which is related to core experiment NO.8 on prediction mode. S-FAMC with frame DCT and with field/frame DCT have been compared and it is concluded that S-FAMC needs the field/frame DCT for rapid motion sequences.

AVC-338 [433] Simulation results on S-FAMC and Dual' (Japan)

This document addresses Simplified FAMC (S-FAMC) and Dual' prediction, which are core experiment NO.1 and No.3 on prediction modes. S-FAMC defined in TM2 has no interpolation mode for B-picture. However this exclusion of interpolation suffers from loss of coding efficiency. This document proposes the introduction of frame base averaged macro block as interpolation mode of S-FAMC. This introduction makes the coding efficiency high as almost same as original FAMC with keeping memory band width same as field or frame prediction. The field/dual' simulation at 4Mbit/s is also carried out and it is concluded that dual' improves the coding efficiency slightly from field/frame prediction but less than S-FAMC.

AVC-339 [456] Simulation results on prediction core experiment (Dual-prime, SFAMC) in TM2 (Toshiba)

Performance of the dual prime prediction was evaluated comparing with that of SFAMC. Simulations results shows that both DUAL' and SFAMC similarly improve coding efficiency if they are combined with Frame and or Field. As a conclusion, the dual prime prediction mode is proposed for its high prediction efficiency and its hardware implementation simplicity.

AVC-340 [454] Results of low delay core experiments on TM2 - Among prediction modes (KDD)

Several prediction schemes are simulated on the low delay mode with M=1, N=15, concluding that;

1) SFAMC with Frame or Field MC gives about 1dB gain in FG, however the gain is limited to about 0.2dB in other sequences.

2) SVMC is effective for both the slow and fast movement and is the best among tested prediction modes in most of all sequences even if its motion

calculation load is about 60% of that of Frame/Field MC.
3) DUAL' with Frame and/or Field MC gives better performance than SFAMC with Frame or Field MC in most of all sequences, but less than SVMC especially in BC and CL sequences.

AVC-341 [430] Results on the comparison of error prediction versus reconstructed signal prediction (PTT Research)

Based on the experimental results showing that there is a small difference (o.o to 0.2 dB) in favor of the prediction from the base layer, it is concluded that the selection should be made according to complexity or functionality.

AVC-342 - Short term PSTN videotelephone standardisation (UK, France, Italy, Denmark, FRG, Belgium, The Netherlands, Sweden)

Several suggestions are mode toward developing standards, which complement the existing ISDN ones, in a rather short period of time (draft proposal by summer 1993).

AVC-343 [488] Results of Core Experiments I.4 - Scalable Side Information (Aus. UVC consortium)

This document discusses results obtained within the framework of Core Experiment "Scalable Side Information" whose purpose is to investigate the possibility of reducing the side information on the lowest layer scale_2 in scalable coding schemes. Four coding architectures were experimented in comparison to a single layer TM2 implementation. Results show that a target bit rate of 750 kbit/s at the lowest layer is realizable using scalable side information and appropriate scaling of quantization between the layers.

AVC-344 [489] A Frequency Pyramid Architecture with Improved Coding Efficiency (Aus. UVC consortium)

A new coding structure is introduced to improve the coding efficiency and to solve the drift problem in frequency scalability schemes. This structure makes use of the prediction from the lower scales by using "path_Y". Experimental results show that this structure improves the coding efficiency of the scalable coder. The following is proposed;
- this encoder/decoder structure be investigated as a follow-up experiment for Core Experiment I.2 and I.3,
- different bit rate distributions be investigated,
- an 8 bit word be transmitted in the sequence header to indicate the structure of the decoder which is necessary in decoding the bitstream.

AVC-345 [490] Considerations on ATM cell loss experiments (Aus. UVC consortium)

The comparison of the cell loss resilient properties of a spatial layered coded to those of a single layered codec requires careful consideration if a fair comparison is to be made. This document raises important issues and provides Table 1, listing parameters and their values, for spatial layered codec cell loss resilience experiment.

AVC-346 [491] Results of Core Experiment I.1 - Interlace-in-Interlace extraction (Aus. UVC consortium)

This documents reports on the results of Core Experiment for the interlace-in-

interlace extraction. Comparisons are made with a simple upper 4x4 extraction technique both in terms of PSNR and bit rate in each layer. In addition, the effect of using simple extraction on images coded in field mode and adaptive field/frame mode are presented. It is concluded that the Method A extraction technique provides the highest quality reconstructed scale-4 image sequence. It is pointed out that any method of extraction results in a very high rate but only fair quality for the scale-4 service, suggesting the use of a coder and decoder with more than one coding loop.

AVC-347 [492] Result of the compatibility experiment 1c : H.261 SIF based prediction of Prediction Error vs Prediction of Input in Frame structure pictures (CNET-FRANCE TELECOM)

The base layer is coded by H.261 (RM8). The results show that the compatible coding using the prediction of input outperforms the simulcast.

AVC-348 [461] Clarification of Appendix H in TM2 (T. Yukitake)

Handling the first picture for "intra picture" case and skipped pictures is proposed to clarify the low delay mode experiment.

AVC-349 [493] Leaky prediction: Eliminating the limit cycle (AT&T)

The limit cycle problem of the leaky prediction which is caused by finite precision of the digital filter is described. The following two solutions to eliminate the limit cycle are provided;

- error spectrum shaping which generates the auxiliary signal using the n least significant bits from the previous truncated signal,
- using a pseudo-random auxiliary signal

Based on experimental results showing faster convergence, the second solution is recommended.

AVC-350 [494] Leaky prediction: Experimental results (AT&T)

Leaky prediction was experimented using the second solution described in AVC-349. For five sequences, values of the leak factor which give equivalent quality compared to the case using cyclic I pictures are obtained as 7/8 - 31/32. Then, cell losses are applied to conclude that the leaky prediction gives less visually objectionable pictures.

AVC-351 [484] Progress Report for Ad-hoc Group on ATM, Packet Loss and General Error Resilience (Chair, MPEG Ad-hoc Group on ATM, Packet Loss and General Error)

An overview is given for the current status of the work, listing various approaches and action points .

AVC-352 [495] TM2 compatibility experiments (BT)

A two-layered coding scheme is experimented where coded SIF base layer pictures are upsampled and used as a prediction for the CCIR601 layer. A new prediction using the average of upsampled SIF and locally decoded CCIR601 is also experimented. It is concluded that the two layer coding scheme can improve the SNR picture quality over the simulcast approach, and that the new compatibility prediction mode gives subjectively better pictures and warrants further study.

AVC-353 [503]   Core Experiment 1.8 results (BELGACOM, UCL)

The aim of Core Experiment 1.8 is to compare the efficiency of the frequency scanning method versus the block scanning one. This document collects the first results obtained. It is concluded that;

- at this stage of experimentation, both techniques present similar efficiency,
- the higher the bit rate, the higher the efficiency of frequency scanning technique,
- further studies are required to adapt frequency scanning technique to the context of TM2.

AVC-354 [498]   Simulation results of the basic TM2 (CNET)

This document presents TM2 simulation results for four sequences with statistical data and processed pictures.

Temporary Documents

TD-1 Agenda for the eighth meeting in Tarrytown (Chairman)
TD-2 Available documents (Chairman)
TD-3 Review of the Experts Group work plan (Chairman)
TD-4 Draft status report - Issue 3 (Chairman)
TD-5 Coding structure vs delay in terms of SNR (T. Yukitake)

END

List of Tape Demonstrations
(28 September 1992, Tarrytown)

No	Organization	Topics	Tape	Doc.
a.	Matsushita	- Low delay	D-60	AVC-327
b.	NEC	- Low delay coding	D-60	AVC-329
c.	KDD	- Frame/field vs SVMC at M=1	D-60	AVC-340
d.	Matsushita Comm.	- Cell loss resilience	D-60	AVC-333
		- Coupling of prediction and DCT for SFAMC	D-60	AVC-337
		- SFAMC and DUAL'	D-60	AVC-338
e.	NTT	- Leaky prediction	D-60	AVC-331
f.	Fujitsu	- Cell loss compensation	D-60	AVC-322
g.	Toshiba	- Compatible coding	D-60	AVC-334
h.	AT&T	- Leaky prediction: eliminate limit cycle	D-60	AVC-349
		- Leaky prediction results	D-60	AVC-350
i.	CNET	- TM2 basic mode	D-50	AVC-354
		- H.261 compatibility	D-50	AVC-347
j.	Monash Univ.	- Scalability: scalable side information	D-50	AVC-343
		- Frequency pyramid architecture	D-50	AVC-344
k.	BT	- TM2 compatibility experiments	D-50	AVC-352

List of documents for the Tarrytown joint sessions

Note: Some of the documents have also AVC-numbers. See Annex 1 to this report.

MPEG92/		
411	Chiariglione	Meeting notice of ad-hoc group on Video Test Model
412	Audio group	List of requirements for MPEG-2 Audio
413	Koster	TM2 Erratum
414	Day	Joint technical session on Integrated Video Services (IVS)
415	CCITT	Report of the fifth meeting of the experts group for ATM video coding in Stockholm/Haifa
416	CCITT	Report of the seventh meeting of the experts group for ATM video coding in New Jersey and Rio de Janeiro
417	Brannon	Indication of copyright
418	JTC1	Resolutions of the ISO/IEC JTC1 Plenary meeting, 30 June - 03 July 1992
419	Holtzmann et al	Bus: Demonstration employing new test footage
420	Hentinen	Letter to Convenor
421	Puri	Meeting announcement, agenda and call for contributions
422	Watanabe	Meeting announcement, agenda and call for contributions
423	Koster	Meeting announcement, agenda and call for contributions
424	Yukitake	Meeting announcement, agenda and call for contributions
425	Biggar	Meeting announcement, agenda and call for contributions
426	Viscito	Meeting announcement, agenda and call for contributions
427	Wells	Meeting announcement, agenda and call for contributions
428	Haskell	Meeting announcement, agenda and call for contributions
429	JTC1	DIS 11576 "Procedures for the registration of algorithms for the lossless compression of data"
430	Koster	Results on the comparison of error prediction versus reconstructed signal prediction
431	Yukitake	Results of low delay core experiments
432	Yukitake	Simulation results on prediction and DCT mode coupling for S-FAMC
433	Yukitake	Simulation results on S-FAMC and Dual'
434	Noguchi et al	Results of some prediction experiments
435	Noguchi et al	Results of some quantization experiments
436	Sugaya	Simulation results on scanning
437	Nagata	Results of Core Experiments on Simplified FAMC, Dual-Prime and Test Model Simplification
438	MIT	Result of core experiments slice vs macroblock
439	Hanamura	Simulation results of core experiments on frequency scalability
440	Takahashi et al.	Simulation Results on Global Motion Compensation Core Experiment
441	Sugiyama	Results of Prediction Experiments
442	Sugiyama	Results of Leaky Prediction Experiments
443	Sugiyama	Results of Quantisation Experiments
444	Sugiyama	Results of Non-8x8DCT Experiments
445	Nakasu	Non-8x8 DCT Experiment
446	Tahara et al.	Results of core experiment - Interlace-in-Interlace Extraction
447	Tahara et al.	Results of core experiment - Encoder with drift correction layer
448	Yagasaki et al	Simulation results on TM2 core experiment of prediction mode No.8
449	Yagasaki et al	Simulation results on TM2 core experiments of quantisation
450	Yagasaki et al	Simulation results on the recommendations from Implimentation Gp

451 Yagasaki et al Proposal of Adaptive DCT/NTC coding
452 Yagasaki et al Introduction of the second VLC table
453 Nakajima Results of prediction core experiments on TM2 (SFAMC,SVMC,DUAL')"
454 Nakajima Results of low delay core experiments on TM2 (Among predction modes)
455 Nogaki Experiment results on low delay coding
456 Odaka et al. Simulation results on prediction core experiment No.3 (Dual-prime) in TM2
457 Odaka Simulation results on quantisation experiment Q.1 (Scanning) in TM2
458 Ueno et al. Simulation results on compatibility core experiment No.1(a) in TM2
459 Kameyama et al Comparison of Prediction Modes and Simplified Test Model
460 Chinen et al. Independent Scanning Results & B-scale Results

461 Yukitake Clarification of Appendix H in TM2
462 Koster Overlapping windows in TV and HDTV compatibility
463 Yu et al. Results of core experiments H2.1 and L7
464 Chiang et al. Results of compatibility experiment G4 using motion compensated up/down sampling
465 Chaing et al. Results of compatibility experiment G2
466 Mokry et al. Insight into the problem of error drift in frequency scalability
467 Nishikawa et al Core experiments: TM prediction methods and simulation results
468 CCITT/J Experiments on cell loss resilience
469 CCITT/J Scene change handling without picture skipping in low delay mode
470 CCITT/J Buffering for low delay mode

471 CCITT/J Coding efficiency of leaky prediction
472 CCITT/J Cell-loss compensation scheme
473 CCITT/J Experiments on cell loss resilience
474 CCITT/J Proposal of TM2 rate control modification for low delay mode
475 CCITT/J Picture header modification for source clock recovery
476 USA NB Contribution on construction of MPEG 2 video syntax
477 Reader et al. Contribution on MPEG 2 video requirements
478 Kogure et al. An adaptive DCT/Non DCT core experiment
479 Kogure et al. Interlace-in-interlace frequency scalability core experiment and improvement
480 Kogure et al. Report on independent scanning mode core experiment

481 Kogure et al. A preliminary report on DPCM lossless coding
482 Iwahashi Tape demonstration of motion compensation technique on 2:1 scaled-down pictures
483 Iwahashi Evaluation of transformation error caused by frequency scaling decoder
484 Biggar Progress report for Ad-hoc group on ATM, Packet Loss and General Error Resilience
485 Koster Improvement of the reconstructed signal prediction compatible coding scheme
486 Viscito Core experiment I.2: Pyramid encoder improvements
487 Liu et al. Result of core experiment on slice vs. MB rate control
488 Aus. UVC cons. Results of Core Experiments I.4 - Scalable Side Information
489 Aus. UVC cons. A Frequency Pyramid Architecture with Improved Coding Efficiency
490 Aus. UVC cons. Considerations on ATM cell loss experiments

491 Aus. UVC cons. Results of Core Experiment I.1 - Interlace-in-Interlace extraction
492 CNET - FT Result of the compatibility experiment 1c: H.261 SIF based prediction of Prediction Error vs Prediction of Input in Frame structure pictures

493	AT&T	Leaky prediction: Eliminating the limit cycle
494	AT&T	Leaky prediction: Experimental results
495	BT	TM2 compatibility experiments
496	Haskell	WD 11172-4 Conformance testing
497	Nocture	Results on TM2 from the VADIS European collaboration
498	CNET-FT	Simulation results of the basic TM2
499	LEP	Additional tests on FAMC compensation mode
500	U. Hannover	Experiments on quantisation and simplification of MPEG2-TM2
501	DTB	Results of scalability core experiments I1 to I3
502	LER	Frequency scanning using ACVLC: new simulation results
503	Poncin et al	Core experiment I8 results
504	HHI	Some new results from frequency scanning using MUVLC
505	Madec	Results on leaky prediction experiments
506	Puri et al.	Flexible spatial scalability syntax and weighted spatio-temporal prediction
507	Puri et al.	Efficient frequency scalability and adaptive interlace-interlace extraction
508	Puri et al.	Report on quantization core experiments and VLC proposal
509	Wang et al.	Spatio-temporal adaptive interlace-interlace extraction
510	Watanabe	Intermediate report of the ad-hoc group on field-frame prediction experiments
511	Paik	A proposal for switching the coding mode (inter or intra) on a block basis
512	Paik	A proposal for specification of 8x8 motion vectors
513	Yukitake et al.	Hardware implementation of SFAMC motion estimation
514	Gonzales et al.	Preliminary results of core experiment I.7
515	Viscito	Tarrytown contr ibution list, core experiment summary, and meeting agenda
516	Daewoo	Core experiment for low delay coding
517	Hyundai	Simulation results of core experiment for scene change handling in low delay
518	Savatier	Simulation results on prediction modes
519	Savatier	Reference fields for forward prediction of P-fields
520	Savatier	Corrections to test model
521	Savatier	Frame prediction in filed-pictures
522	Okubo	Integrated requirements listing, second draft
523	Wong	TM2 errors
524	CLI et al.	Recommendations for a broadcast profile for MPEG-2
525	Haskell et al.	Vector quantization
526	Haskell et al.	Comparison studies between run length coding and vector quantisation
527	Haskell et al.	Adaptive switching of coefficient VLC's
528	CableLabs	Application profile/requirements for the genric coding method of moving picture images from the perspective of cable television distribution in North America
529	JTC1/SGFS	DTR 10000-2.3 (E) - Information Technology - Framework and Taxonomy of International Standardized Profiles - Part 2: Taxonomy of Profiles
530	Wells	Ad-hoc group on quantization etc.: Summary of papers received before Tarrytown meeting of the Video Group
531	Aravind et al	Results on switching of quantizer matrices
532	Sun et al.	A proposal for increased error resilience
533	CCIR	Change in venue - Task group 11/4 meeting, 13-15 October 1992 - Washington, D.C.

534	JTC1	Annex G to the JTC1 directives
535	Koster	Proposal for TM2 Rev. 2
536	Koster	Tarrytown recommendations of the Ad-hoc group on Test Model Editing
537	Puri	Tarrytown recommendations of the Ad-hoc group on compatibility
538	Watanabe	Tarrytown recommendations of the Ad-hoc group on Field/Frame prediction
539	Yukitake	Tarrytown recommendations of the Ad-hoc group on low delay mode
540	Dunstan	Tarrytown recommendations of the Ad-hoc group on ATM
541	Viscito	Tarrytown recommendations of the Ad-hoc group on scalability
542	Wells	Tarrytown recommendations of the Ad-hoc group on quantisation
543	Chiariglione	Tarrytown recommendations of the Ad-hoc group on Video Test Model
544	Haskell	Tarrytown recommendations of the Ad-hoc group on conformance testing
545	Watanabe	Report of the Frame/Field prediction ad-hoc group at the Tarrytown meeting
546	Yukitake	Report of the Low delay mode ad-hoc group at the Tarrytown meeting
547	Chiariglione	Report of the Video Test Model ad-hoc group at the Tarrytown meeting

END

Report of the discussion on "Profile"

Adhoc Group on Video Test Model had a session on Wednesday to discuss how to structure the generic standard MPEG-2, paying attention to the notion of "profile". Before the discussion, MPEG Convenor presented the background, referring to the related decisions to date;

- Our policy is to take the "toolkit with maximum core" approach to MPEG-2.
- Performance and functional requirements from various applications have been collected and now being integrated into a single document covering system, video and audio.
- The generic standard and specific applications may be liaised by use of the "profile" notion given in ISO/IEC Technical Report DTR 10000. Its terminology is as attached.
- "Profile" is currently used rather liberally, thus clarification and common understanding of the members are necessary.

Activities on video requirements so far have identified features to be developed, and these have been reflected in various experiments of the Test Model. If these experiments are successfully carried out, we will be able to provide sufficient tools in the new MPEG-2 standard to meet those requirements from a wide range of applications.

The issue is that one specific application may use a particular combination of features and parameters conforming to MPEG-2 for its optimization and another specific application may choose another combination also conforming to MPEG-2 for the same reason; possible examples mentioned were use/non-use of B-pictures, new predictions (such as intelligent prediction, leaky prediction). The result may be that two MPEG-2 conforming equipment cannot talk each other (i.e. incompatibility inside MPEG-2). Since implementation technology will evolve, there may happen interoperability problems between different generation equipment.

The policy of "toolkit with maximum core" is intended as a compromise between meeting as many requirements as possible (i.e. generic) and keeping interoperability as widely as possible. The boundary of the core should be clarified at some point of time (hopefully in November).

Toward this direction, application of "profile" may help us. Dr. LeGall provided the attached diagram as a basis of further elaboration, to which a particular application (or group of applications) may be mapped. Another of his suggestions is to define a couple of "target profiles" similar to the "constrained bitstream" in MPEG-1, according to this diagram.

One of the suggestions raised during the discussion was to work out a matrix relating functionalities provided in the TM to specific or group of applications. The purpose of this is to clarify application attributes, maximum core, interoperability and interface issues. There was a concern expressed against this approach, however, that there may be a risk to get involved in business aspects for which this engineering group is not responsible.

Members are requested to contribute to clarifying this issue and the way to reflect the outcome in the standard.

```
                        -----------
                       | The Works |
                       |-----------|
                       |     |     |
        |--------------|---| | |---|------------|
        |               ---|---|---             |
        |                  |   |                |
 ---------------     --------------     -------------
| Compatibility |   |   No Frill   |   | Scalability |
|   Extension   |   |     Core     |   |  Extension  |
|               |   |    Syntax    |   |             |
 ---------------     --------------     -------------
                           |    ---------
                           |---| Normal |
                           |    ---------
                           |    ---------
                           |---|Low Delay|
                           |    ---------
```

Figure   Link between features and syntactic subset

* * * *

Definitions given in DTR 10000

International Standardised Profile

An internationally agreed-to, harmonized document which identifies a standard or group of of standards, together with options and parameters, necessary to accomplish a function or set of functions.

Profile

A set of one or more base standards, and, where applicable, the identification of chosen classes, subsets, options and parameters of those base standards, necessary for accomplishing a particular function.

Profiles provide a basis for the development of uniform, internationally recognized, conformance tests.

Note: An International Standardised Profile includes the specifications of one or more Profiles

Base Standard

An approved International Standard or CCITT Recommendations which is used in the definition of a Profile.

END

Frame/field prediction Adhoc Group

Annex 5.1 Frame/Field Prediction, Low Delay Adhoc Group Joint Meeting Report

Title: Discussion on Core Experiments L1, L2, L3, H1

Several prediction techniques which accommodate interlace structure are proposed as core experiments L1, L2, L3. These are (1) FAMC, (2) SVMC, (3) Dual'. They are also listed up as a special mode to improve picture quality for low delay core experiment (H1). Simulation results of these prediction modes are demonstrated, and they are recognized as the technique having the same level of performance in terms of SNR and picture quality by tape viewing.

1. Tape demonstration

TI	FAMC, Dual', TM simplification
Toshiba	FAMC, Dual'
Sharp	FAMC
Mitsubishi	Dual'
LEP	FAMC
KDD	SVMC, FAMC, Dual'
Matsushita	FAMC
TCE	SVMC
GCT	FAMC,

2. Discussion

1) Purpose: Purpose of the special prediction mode is mainly for the low delay application to improve picture quality especially at M=1. It is theoretically effective when a sequence has slow vertical motion with a constant velocity.

2) Necessity: Majority of the group can distinguish the difference of the picture quality between special prediction modes and field/frame adaptive prediction. However, some of members could not recognize significant differences. Thus, an agreement of the necessity of this mode is not unanimous.

3. Conclusion

1) Evaluation: Differences of the coding efficiency between FAMC, SVMC and Dual' is very small. The agreement to select the best among three was not obtained.

2) Action towards the London meeting: Two adhoc groups agreed with the idea to merge three candidates (FAMC, SVMC, Dual') into the new single core experiment by mixing up their good features. The decision of adaptation in the TM syntax will be based on the recognition of its significance at the London meeting.

Annex 5.2 Resolution of the issues raised by Test model editing group and MPEG92/523

1. Inconsistent use of horizontal_f and forward_f_code

"forward_f_code" and "backward_f_code" indicate "horizontal_f", therefore, they are interpreted as "forward_f_hori_code" and "backward_f_hori_code". "forward_f_vert_code" and "backward_f_vert_code" should exist in the picture layer after "extension start code". The number of the maximum f_codes is 4. (andria@nyquist.bellcore.com)

2. The motion estimation and compensation range for field-picture

The horizontal motion compensation range is half of frame-pictures per field. The vertical range is quarter of the value written in TM2 page21.

3. The definition of the search pseudo c-code in SVMC

Modifications are supplied by Mr.Nakajima (KDD, nakajima@spg.elb.kddlabs.co.jp).

4. Clarification of the use of SVMC and FAMC in field picture

Modifications are supplied by Mr.Nakajima (KDD, nakajima@spg.elb.kddlabs.co.jp), and Mr.Yukitake (Matsushita, yukiatke@adl.mci.mei.co.jp).

5. Clarification of the use of 16x8 field macroblock

We use "field_motion _type" in the macroblock layer to specify 16x8 motion compensation block. If "field_motion_type" is "10", 18x8 motion compensation is used instead of SFAMC.

"field_motion_type"

code	prediction type	motion_vector_count	mv_format
10	16x8	2	field

"sub_MB_type" should be added just after "backward_reference_field". This is 1 bit flag, uimsbf.
sub_MB_type=="0" indicates the condition written in Core Exp. L9 3(1).
sub_MB_type=="1" indicates the condition written in Core Exp. L9 3(2).

6. Default values of "forward_reference_fields" and "backward_reference_fields"

Default values are set to "11" in both cases, which means Field 1 and Field 2 are used for prediction.

7.Syntax on progressive material

Frame structure, frame prediction may be used for progressive material. However it does not prohibit to use frame field adaptive or field structure.

8. "temporal_reference" is needed to obtain frame/field distance for decoder. Its unit is "Frame".

9. Ambiguity in the definition of the noMC mode in field-pictures (p-field pictures only)

If two previous fields are allowed as reference fields, the noMC mode does not specify which of the reference fields is to be used for the prediction. Thus, the noMc mode shall refer to the reference field of the same parity as the target field. In the case a (0,0) MV is to be used with the reference field of the opposite parity, the noMC mode cannot be used. The (0,0) MV must be explicitly transmitted (after the appropriate selection bit).

Annex 5.3 <u>Special Prediction Modes (Ver. 2)</u>

1. Definitions

SVMC3 is SVMC without FAMC. The latest document is used (TM-2 Erratum), with the modifications and corrections described in this document.

2. Temporal Scaling of the Motion Vector

Scaling of the motion vector is done in the same manner for all the special prediction modes (FAMC, SVMC3 and DUAL-PRIME).

The transmitted motion vector (x, y) corresponds to a prediction from same-parity field.

The horizontal coordinate is in 1/2-pel units. The vertical coordinate is in 1/2-pel units or 1/4-pel units (depending on the mode and of the experiment performed).

If the same parity reference frame is at a distance of 2*k fields from the predicted field, the coordinates (x', y') of the "scaled-motion-vector" used for accessing the different-parity field is computed as follows:

$x' = (x * m * K) // 32$ (x and x' are integers)
$y' = ((y * m * K) // 32) + e$ (y and y' are integers)

$K = 16 // k$ (k is integer)

m = field-distance between the predicted field and the different-parity-field (m is integer and can be negative).

The "e" is an adjustment necessary to reflect the vertical shift between the lines of field 1 and field 2. To give an example, line 1 of field 2 is in fact located 1/2 line under line 1 of field 1.

If vertical unit is 1/4-pel, "e" is defined as follows:

e = -2 if the reference field corresponding to the scaled vector is field 2
e = +2 if the reference field corresponding to the scaled vector is field 1

If vertical unit is 1/2-pel, "e" is defined as follows:

e = -1 if the reference field corresponding to the scaled vector is field 2
e = +1 if the reference field corresponding to the scaled vector is field 1

3. Reference Fields for SVMC3 and DUAL-PRIME

The reference fields used for SVMC3 and DUAL-PRIME are not always contiguous in time. Those modes can now be used in all cases of Field-structure P-Pictures.

When SVMC3 or DUAL-PRIME is used in the second P-Field of a P-Picture, the first P-Field is used as a reference (different-parity) field.

SVMC3 and DUAL-PRIME can be used with reversed order prediction of P-Fields (in this case, m = -1).

4. Decision for Field-based Prediction

In order to take advantage of the various special prediction modes, the decision rule must be modified for Field-based prediction.

It has been noted by various members that quality is improved by choosing Field-based prediction less often, to the benefit of another special prediction mode, particularly in B-Pictures.

For example, even in cases where Field-based prediction has an MSE slightly better than any of the other prediction modes, it may cost a significant overhead to transmit two field-vectors (four in B-Frames).

Until further improvement, we propose to use the following decision rule in core experiments involving one of the special prediction modes:

- Field-based chosen
    if MSE_field + 8 < MSE_best_of_other_modes in B-pictures
    if MSE_field < MSE_best_of_other_modes in P-pictures

where MSE = Mean Square Error PER PEL of predicted MB

5. Concise Specification of SVMC3

The transmitted motion vector is scaled with the specified rule to obtain motion vectors origination from each of the reference field, and pointing to the predicted field. When the reference field and the predicted field are of same parity, the motion vector is used directly (no scaling is necessary).

5.1 Forward Prediction

5.1.1 Forward Prediction of the pels of Field 1 (16Hx8V)

The coordinates (x'1 , y'1) of the scaled motion vector are computed as specified, with m = m1.

A 16Hx8V prediction block is obtained from reference field 1 with the vector originating from this field Vertical interpolation is 1/4-pel linear interpolation. Horizontal interpolation is 1/2-pel as usual. Like in the "usual" case, horizontal and vertical interpolation are done in a single step, involving only one division (in this case by 8). An example is given in Figure 1 and Figure 2.

A 16Hx8V prediction block is obtained from reference field 2 with the vector originating from this field. Vertical interpolation is 1/4-pel linear interpolation. Horizontal interpolation is 1/2-pel as usual.

The selection of the prediction is done according to the SVMC3 type:

- Near-field: The prediction block used is the one corresponding to the reference field closest to the predicted field (in time axis).

- Same-parity: The prediction block used is the one corresponding to the reference field of same parity as the predicted field.

- Dual: The prediction block used is obtained by averaging the two prediction blocks from field 1 and field 2. The averaging is done like in "Interpolation-mode" in B-Pictures.

5.1.2 Prediction of the pels of Field 2 (16Hx8V)

The coordinates (x'2 , y'2) of the scaled motion vector are computed as specified, with m = m2.

For the rest, the prediction is done like in 5.1.1.

5.2 Backward Prediction

The forward rule is simply transposed.

5.3 Averaged Prediction in B-Pictures

5.3.1 Averaged Prediction of the pels of Field 1 (16Hx8V)

The predictor blocks for forward prediction are computed as in sections 5.1.1 and 5.1.2.

The predictor blocks for backward prediction are computed as in 5.2.

The selection of the prediction is done according to the SVMC3 type:

- Near-near: The prediction block used is obtained by averaging the prediction blocks from the closest forward and backward reference fields (in time axis).

- Same-near: The prediction block used is obtained by averaging the prediction block from the same parity forward reference field and the prediction block from the closest backward reference field (in time axis).

- Near-same: The prediction block used is obtained by averaging the prediction block from the same parity backward reference field and the prediction block from the closest forward reference field (in time axis).

- Same-same: The prediction block used is obtained by averaging the prediction block from the same parity forward reference field and the prediction block from the same parity backward reference field.

Note that the four SVMC3 averaged modes and the SVMC3 dual mode are extremely similar. Only the choice of the two reference fields differs. The averaging is always done like in "Interpolation-mode" in B-Pictures.

The other SVMC3 modes are equivalent to field-based prediction with 1/4 vertical accuracy.

5.4. Chrominance

The motion vector used for chrominance is obtained from the luminance SVMC3 motion vector with precisely the same rule as in the case of field-based prediction (for 4:2:0 : divide each coordinate by 2 as described section 5.2.2.1. of TM-2). The rules of prediction are same as for luminance.

5.5. Motion estimation of SVMC3

Search is done by a local refinement around several candidate motion vectors resulting of a first search. The candidate motion vector can be the result of a full-pel accuracy search. The local search covers 5Vx5H = 25 motion vectors and is done on reconstructed. For each of those, all the candidate SVMC3 prediction blocks must be evaluated. For local search, the vertical step is 1/4-pel, and the horizontal step is 1/2-pel.

In the case of Frame-Pictures, the candidate motion vectors used as starting point of local search are:

- The Frame motion (result of Frame-based search).

- The Field motion vector (result of Field-based search) from the closest reference field to the predicted field of same parity. In this case, the vertical cordinate must be multiplied by two to have 1/4-pel vertical field accuracy.

If forward: from field 2 to field 2
If backward: from field 1 to field 1

- Optionally, the other field motion vectors (scaled appropriately) could be used as candidate motion vectors.

In the case of Field-Pictures, the candidate motion vectors used as starting point of local search are the field motion vectors (result of Field-based search), scaled to the field-distance corresponding to same-parity. In this case, the vertical coordinate of the field vectors must be multiplied by two to have a candidate motion vector with 1/4-pel vertical field accuracy.

6. Concise Specification of DUAL-PRIME

In DUAL-PRIME prediction, single motion vector like that of SVMC3 with 1/2 pixel precision and one very small differential motion vector called DMV is transmitted per macroblock. To obtain motion vectors originating from each of the reference field, and pointing to the predicted field of the different parity, the transmitted motion vector is scaled and DMV is added as follows:

x' = ((x * m * K) // 32) + dmv_horizontal (x and x' are integers)
y' = ((y * m * K) // 32) + dmv_vertical + e (y and y' are integers)

The variables (x, y), (x', y'), K, m, and e have been already defined above, and vertical unit is 1/2-pel.

dmv_horizontal and dmv_vertical are horizontal and vertical components of DMV with 1/2 pixel precision. These values are restricted within the range from -1 to +1. Note that the same DMV is used for the two scaled motion vectors in the frame picture as illustrated in Figure 3.

When the reference field and the predicted field are of same parity, the motion vector is used directly (no scaling is necessary, and the addition of DMV is not necessary).

6.1 Forward Prediction

6.1.1 Forward Prediction of the pels of Field 1 (16Hx8V):

The coordinates (x'1 , y'1) of the scaled motion vector are computed as specified, with m = m1.

A 16Hx8V prediction block is obtained from reference field 1 with the vector originating from this field. Both horizontal and vertical interpolation is 1/2-pel linear interpolation as usual in the field motion vector. Like in the "usual" case, horizontal and vertical interpolation is done in a single step.

A 16Hx8V prediction block is obtained from reference field 2 with the vector originating from this field. Both vertical and horizontal interpolation is 1/2-pel linear interpolation as usual.

The prediction block used is obtained by averaging the two prediction blocks from field 1 and field 2. The averaging is done like in "Interpolation-mode" in B-Pictures.

6.1.2 Prediction of the pels of Field 2 (16Hx8V):

The coordinates (x'2 , y'2) of the scaled motion vector are computed as specified, with m = m2.

For the rest, the prediction is done like in 6.1.1.

6.2 Backward Prediction

The forward rule is simply transposed.

6.3 Prediction mode in B-Pictures

The averaging mode is inhibited in DUAL-PRIME. Only the forward/backward prediction is used in B-pictures.

6.4. Chrominance

From DUAL-PRIME motion vector, four field motion vectors for luminance from the reference field 1/field 2 to the predicted field 1/field 2 can be obtained. Corresponding four chrominance vectors are obtained with precisely the same rule as in the case of field-base prediction (for 4:2:0; divide each coordinate by 2 as described section 5.2.2.1. of TM-2). The rules of prediction are same as for luminance.

6.5. Motion estimation of DUAL-PRIME

The motion estimation of DUAL-PRIME is carried out by the following two steps.

The first step is to obtain four candidate motion vectors as follows. First, four field motion vectors with half-pel accuracy from the reference field 1/field 2 to the predicted field 1/field 2 are searched by the normal field motion vector search method defined in TM2, except that original pictures are used in half-pel refinement. Then, these vectors are appropriately scaled, if the parity of the predicted field is opposite to that of the reference field.

The second step is to evaluate prediction errors using possible combinations of four candidate motion vectors obtained by the first step, and 3Vx3H = 9 candidates for DMV using local decoded pictures, and to select the best combination of the motion vector and DMV.

7. Core Experiment L-10

L-10.1. Frame + Field + DUAL-PRIME

L-10.2. Frame + Field + SVMC3

L-10.3. Frame + Field + SVMC3-1/2-pel

Same as L-10.2, except that all motion vectors involved are only 1/2-pel vertical accuracy. The motion vector transmitted is field-type, as in DUAL-PRIME. However the rule for selecting the PMV's is same as for SVMC3, i.e., the same as for Frame-prediction. Scaling of the PMV vertical coordinate is done like for field vectors.

L-10.4Frame + Field + DUAL-PRIME + SVMC3

L-10.5Frame + Field + DUAL-PRIME + (SVMC3 - dual)

The dual mode of SVMC3 is not used, since DUAL-PRIME may replace it advantageously. However, no significant hardware simplification is expected by implementation of l-10.5 vs. L-10.4. Among the modes in the above core experiment, mode selection is decided by MSE.

Annex 5.4 Selection Process of 8x8 motion vector for L-11

1. For each block within a macroblock, count the number of bits (n16) required for interframe coding using 16x16 motion vector and the number of bits (n8) required for interframe coding using 8x8 motion vector. "n8" should include the overhead required for the 8x8 motion vector. "n16" does not include the overhead for the 16x16 motion vector. Choose 8x8 motion vector if "n8" is smaller than "n16".

2. Add all the bits ("n8" or "n16" depending on the decision) for all block in the macroblock, and add the overhead for the 16x16 motion vector. This number is "nDPCM". Compare the number (nDPCM) and the number of bits required for intraframe coding of the macroblock (nPCM). Choose interframe coding if "nDPCM" is smaller than "nPCM".

END

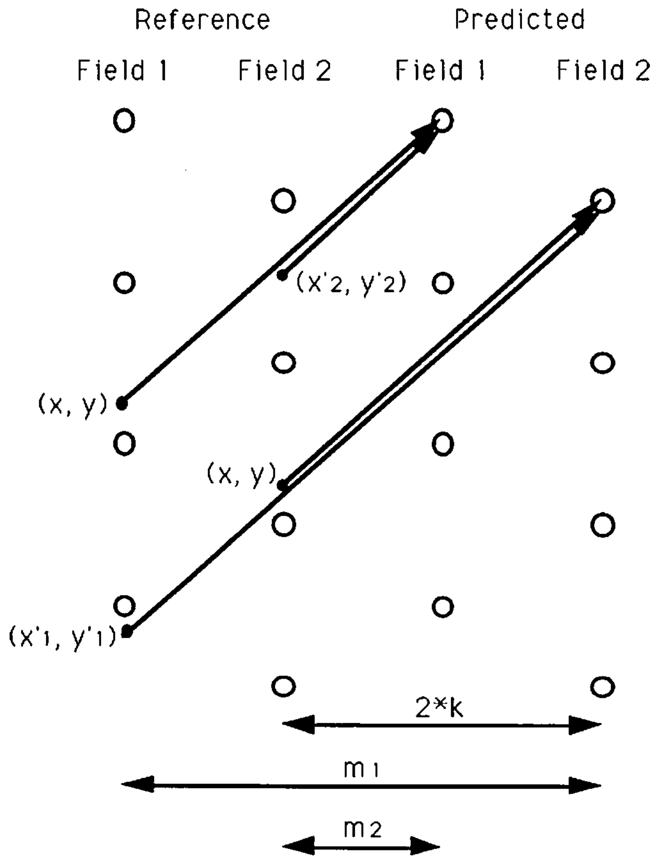

Figure 1 SVMC

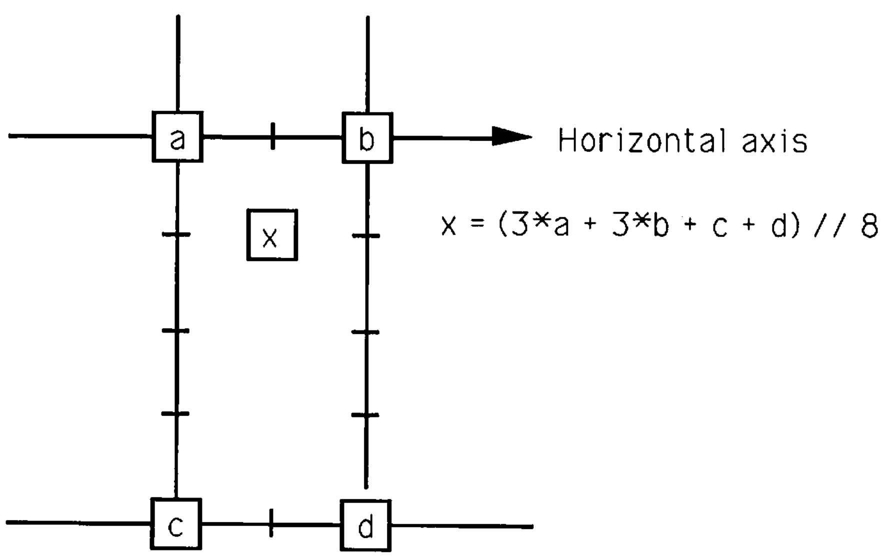

Figure 2 Interpolation at 1/2 H, 1/4 V

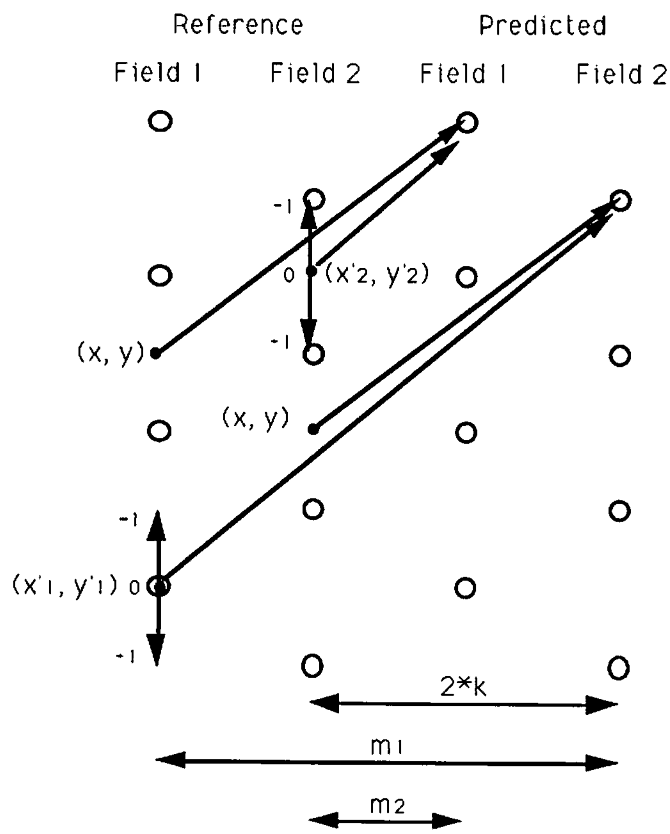

Figure 3 DUAL-PRIME