Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG
(ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6)
3rd Meeting: Fairfax, Virginia, USA, 6-10 May, 2002

Document: JVT-C001d3
Filename: JVT-C001d3.doc

Title: **(Draft) Report of Fairfax Meeting**

Status: (Draft) Output Document Approved by JVT]

Purpose: Report

Author(s) or Contact(s):

Gary Sullivan,	Rapp. | Chair	Email:	garysull@microsoft.com
Ajay Luthra,	Assoc. Rapp. | Co-Chair		wiegand@hhi.de
Thomas Wiegand,	Assoc. Rapp | Co-Chair		aluthra@motorola.com

Source: JVT

1. Introduction

The Joint Video Team (JVT) of ITU-T SG16 Q.6 (VCEG) and ISO/IEC JTC1/SC29/WG11 (MPEG) held meetings during the period of May 6-10 in Fairfax, Virginia, USA. These meetings were held under MPEG auspices and were chaired by the JVT Rapporteur | Chair Gary Sullivan and the JVT Associate Rapporteurs | Co-Chairs Ajay Luthra and Thomas Wiegand.

2. Administrative Issues

2.1. General administrative documents

JVT-C000 **[Sullivan+] List of Documents**
JVT-C001 **[Sullivan+] Report of Fairfax Meeting (#3)**
JVT-C002 **[Sullivan+] Report of Geneva Meeting (#2) Approved.**
JVT-C003 **[Sullivan+] List of Fairfax Participants**
JVT-C004 **[Sullivan+] List of JVT Experts**
JVT-C005 **[Sullivan] AHG Report: JVT Project Management**

2.2. Administrative Ad Hoc Group Reports

JVT-C006 **[Wiegand+] AHG Report: WD & JM Doc & S/W Editing**
JVT-C015 **[Lim+] AHG Report: NAL and High-Level Syntax**
JVT-C039 **[Wiegand+] Editor's latest draft of JVT WD2 design**

The group agreed that the WD2 rev8 editor's input was the best basis for drafting the CD, plus specific actions as noted.

2.3. Inputs through MPEG

M8338 **USNB Comments**: Actions taken are consistent with the USNB requests.
No informative annexes will be included in the draft at this stage. We will continue to work on these as a longer-term effort.

M8407 **Draft CFP for fast motion estimation.** JVT prefers to focus on normative text finalization at this time, rather than to start a process for evaluation of fast example encoding methods.

JVT-C154 **[Sullivan, Microsoft] Summary of JVT-Related Actions 3/02 MPEG Mtng**
A report of actions taken at previous MPEG meeting where JVT was not present.

Procedure to develop profiles & levels: Met with MPEG to discuss this.

Nokia Exhibit 2024, p.1
Snap/Hisense v. Nokia
IPR2025-01440

Proposed guidelines for carriage: [N4714] Replied to MPEG that strong collaborative work took place and we believe these issues to be essentially resolved by the CD draft.

The JVT committed to providing configuration files for each defined profile for using the JVT reference s/w.

JVT-C156 **[MPEG] MPEG Requirements for AVC Codec [N4672p]**
Requirements for AVC codec [N4672]. Understanding is that this input does not require change to current JVT video-level requirements. The JVT reported to MPEG that the 4:2:2 requirement is unfulfilled in the CD.

JVT-C157 **[MPEG] MPEG Communication to JVT [N4673n]**
Regarding communication to JVT: The JVT reported to MPEG that we believe these issues to have been properly addressed and followed in the submitted CD.

JVT-C160 **[MPEG] MPEG Guidelines Workplan for Complexity Analysis of AVC [N4571n]**
The JVT met with MPEG to discuss complexity analysis and decided to establish an AHG on complexity analysis with co-chairs Jan Bormans and Mike Horowitz using one reflector for both WG11 and JVT.

2.4. Inputs through VCEG

JVT-C153 **[Sullivan, Microsoft] Summary of JVT-Related Actions 2/02 SG16 Mtng**
This input contribution was appreciated. The JVT work ongoing is consistent with the remarks provided by SG16.

VCEG input comments:

- There should be a minimum number of profiles (definitely not more than three)
- The profiles should be designed as something "to be proud of" in terms of coding efficiency
- Suggestion of "B+2" level of memory capacity in all profiles.

2.5. Group adoption of "5%/10% Rule of Thumb"

The group adopted the following guideline for coding efficiency proposal evaluations at this meeting:

Proposed changes are not to be adopted unless the change results in average BD-Rate savings of at least approximately 5% average across group-agreed common conditions (quantization range agreed to be realistic, test sequences agreed to be realistic), and 10% peak.

Very minor changes (like adding one or two sentences to the draft) or (better) bug-fix/clean-up work may also be considered.

Complexity reduction proposals require different consideration.

Participants should expect these threshold values to go up at the next meeting.

Exceptions are to be made only with very strong group consensus.

2.6. IPR issues

JVT-C110 **[Sullivan, Microsoft] JVT IPR Status Report**
This document contained a report on the status of IPR issues in the JVT project on a best-effort basis, pursuant to the IPR policy and goals described in the JVT Terms of Reference (VCEG-O54 / WG11-N4400).

JVT-C085 **[Kogure+, Matsushita] IPR WG establishment for JVT**
This document requested to establish an IPR working group in JVT to conduct expert investigations of the validity and potential infringement of IPR related to the potential royalty free baseline profile and to discuss possible solutions to reach that objective.

JVT-C123 **[van der Meer, Philips] JVT Licensing issue**
The potential JVT licensing situation was discussed in this contribution with an expression of a need to address IPR issues. Doubt was expressed regarding whether the royalty-free baseline profile goal of the JVT project could be

Nokia Exhibit 2024, p.2
Snap/Hisense v. Nokia
IPR2025-01440

realistically achieved. Concern was expressed in particular regarding the potential applicability of patents relating to MPEG-2 video.

JVT-C124 **[van der Meer, Philips] Information on JVT patents**
This contribution reported that a significant percentage of patents relating to MPEG-2 may be essential to the JVT baseline and that some of the owners of these patents may not actively participate in JVT. The JVT was suggested to make sure that potential patent holders beyond JVT are approached to provide commitments on licensing conditions of IPR they may hold that may be essential elements of JVT technology, including on the JVT baseline. It was suggested to approach in particular potential patent holders that participated in developments on video coding technology in ITU and MPEG since about 1985.

JVT-C149 **[Yagasaki, Sony] Licensing issues on JVT**
This contribution reported that patent rights of some technologies in the JVT codec may belong to companies and organizations that may have declared licensing terms under RAND terms and conditions for other standardization projects and that these patent holders may not necessarily be members of JVT. The contrubtion requested to establish a working group consisting of "IPR experts", to discuss a licensing model for the JVT codec.

JVT-C150 **[Lindbergh, Polycom] Support for JVT Royalty Free Baseline**
This contribution from a large number of co-authoring companies (Apple Computer, British Telecom, Broadcom, Cisco Systems, Conexant, Deutsche Telekom, FastVDO, Nokia, Polycom, Radvision, Sand Video, Siemens, Sun Microsystems, Tandberg, Telenor, Teles AG, Texas Instruments, UBVideo, VCON, Videolocus, ViXS, and VWeb) expressed strong support for the royalty-free baseline profile goal of the JVT project. The contribution expressed the view that the royalty-free baseline profile goal is critically important to the success of the JVT standard in the marketplace, and that this goal is practical and workable.

Response to these contributions:

The Baseline Profile as defined in the Joint CD Fairfax output document does not include any currently-reported 2.2 (or 3.2 referring to unknown or 2.2 status) IPR in the adopted list. The editor was asked to double-check the adoption list and identify whether any concerns exist regarding this status assessment.

The JVT supported the concept of a joint MPEG+VCEG call for information on IPR status to address the need for outreach to potential patent holders and to ensure no infringement of their patent rights in the JVT design.

No change to JVT goals or policy was made in response to these contributions.

3. Motion Compensation

3.1. "Improved MB Prediction" Modes

JVT-C119 **[Adachi, NTT DoCoMo] Complexity Analysis of Improved MB Prediction Modes**
This document reported the results of the complexity analysis at the decoder and encoder side, focusing on the impact of different MB coding strategies for motion estimation. In particular it compared the performance and the complexity of VCEG-017 and VCEG-022. It proposed the adaptation of the so-called "improved MB prediction modes" as a feature of higher profiles.

Intel Xeon Processor 1.7 GHz and 1024MB of memory used for test. Decoder complexity was reported to be similar. Encoder complexity was reported to between 5% and 14% higher than VCEG-O17.

For encoder operation, scalable mode selection is proposed. The 7 shapes of WD1 are proposed for baseline while the complete set of 19 shapes is proposed for a higher complexity profile.

Group comments: increased number of MB modes causes increased amount of code maybe leading to cache misses. May require 4x4 MC for all MB modes.

Some people pointed out that "the above statement about cache misses is true for the WD2 model as well."

Adding Intra provides 0.8 % additional bit rate savings.

Nokia Exhibit 2024, p.3
Snap/Hisense v. Nokia
IPR2025-01440

Some people sympathetically were disposed to the concept, but for motion estimation in hardware, complexity maybe increased.

Some people remarked: no clear judgment of code size / complexity relationship possible.

Compression quality relative to JWD2? Reported to be expected to be 3-4% better.

3.2. Motion Interpolation

JVT-C014 **[Johansen+] AHG Report: Motion Interpolation**
Significant activity was reported. 3 contributions for adaptivity and 3 for complexity were noted.

JVT-C037 **[Bossen, NTT DoCoMo] Full 16-bit implementation of 1/4-pel motion compensation**
The document provided information about an implementation of quarter-pel motion compensation compliant with JWD2 could be implemented that requires 16-bit registers only.

It was suggested to turn this into an informative Annex.

It was also reported that when using a 2-D kernel using only 14-bit registers should suffice.

For DSPs, a shift-based implementation often does not make a difference while for ASICs it could.

Group suggested to draft for consideration short informative annex. However, no informative annexes were included in the CD output draft due to scheduling issues. The group plans to pursue the drafting of such informative content on a slower schedule than that of the main text.

JVT-C040 **[Miyamoto+] Adaptive Motion Interpolation on MB-basis**

Using a set of pre-calculated filter coefficients, the proposed encoder selects a filter number and sends the selection to the decoder on a macroblock basis.

PSNR: Reported Up to 0.5 dB, 5% BD-PSNR on Mobile, average 2% overall, additional information provided

Complexity: Higher than current ¼ pel method, but less than 1/8 pel.

Common test conditions: yes

Perceptual quality: No demo. Proponent impression very small but visible gain.

Verification: No.

Software: Used JM1.4, not yet made available to JVT.

JVT-C052 **[Sato+, Sony] Adaptive MC Interpolation Filter for Complexity Reduction**
Currently, a 6-tap filter is used for 1/4-pel MC and 8-tap filter is used for 1/8-pel MC, which requires high decoder computational complexity and memory bandwidth especially for SDTV and HDTV bitstreams. Various solutions were proposed:
Solution 1: FIR1 for larger MC blocks than 8x8; FIR2 for smaller MC blocks
Solution 2: P-picture …FIR1 for larger MC blocks than 8x8; FIR2 for smaller MC blocks
B-picture …FIR2 for all MC blocks
Solution 3: P-picture …FIR1 for larger MC blocks than 8x8; FIR2 for smaller MC blocks
B-picture …FIR2 for larger MC blocks than 8x8; FIR3 for smaller MC blocks
FIR1 {1, -5, 20, 20, -5, 1} // 32
FIR2 {-1, 5, 5, -1} // 8
FIR3 {1, 1} //2
Subset of test set used. Very little losses reported.

Verification: code will be available.

Proposed: Solution 1 and 3.

Nokia Exhibit 2024, p.4
Snap/Hisense v. Nokia
IPR2025-01440

Motion Search? Currently 3 different sub-pel filtered areas are used in motion estimation. The impact of using 1 filter for ME instead of three and then compensating with the right one is not known.

No subjective tests conducted.

Roughly 1% loss. Reported roughly 40% less complex for P pictures in decoder – complexity increase in encoder.

JVT-C057 **[Fuldseth, Tandberg] Evaluation of Adaptive Interpolation Filter for Real-time Applications**
The use of an adaptive interpolation filter has been evaluated with focus on real-time applications. In particular, various methods have been investigated to reduce the encoder complexity. It was reported that the author had not been able to find low-complexity implementations that justifies the inclusion of an adaptive interpolation filter in the standard.
A codebook of candidate filters is used.
Different filters in (0,1/2), (1/2,0), and (1/2,1/2) positions respectively.
Basic algorithm gives about 0%-5%.
Using a codebook of 128 filters that was generated manually reduced gain by 0-1%.
No explicit results provided.

Info doc. Tried to achieve complexity reduction, but found loss of performance – provider does not think much complexity reduction likely to be achievable.

JVT-C059 **[Wedi, Hannover U.] New Results on Adaptive Interpolation Filter**
In this proposal an adaptive interpolation scheme was presented. This interpolation scheme is based on filter coefficients that are adapted once per frame to the non-stationary statistical properties of the video signal. The filter-coefficients are coded with 12 bits (requiring 32 bit registers) and transmitted in the slice header. Due to the adaptive interpolation filter a coding gain up to 0.8 dB PSNR was reported.

BD-Rate savings for a set of CIF sequences is 0.8-8% and for one sequence (Waterfall) the gains were up to 18%.

Common conditions 4.9% average for 1 reference frame, 3.2% for 5 reference frames. 8% peak for 6-tap, 12% for 8-tap. On other tested sequences the gain is higher.

Subjective results ok.
Verification: software of older version is available and the presented version will also be made available.
Results for B frames not available.
Group: Gains are impressive. Simple encoder implementation requested in order to show graceful degradation.
Feature adds some flexibility for encoder optimization.
Puts a potential burden on hardware decoder.

Expressed idea for possible adoption at the next meeting in non interactive profile subject to successful tests with 5 bit (instead of 12-bit) quantization and 16 bit implementation and results for B frames.

Further work planned.

JVT-C104 **[Boyce, Thomson] 4 Tap Motion Interpolation Filter**
This contribution proposed inclusion of a 4-tap 1/4 pel motion interpolation filter in the standard, selectable by profile. Experimental results were presented comparing 4-tap and 6-tap performance, with and without the use of B frames and multiple reference frames.

The average loss is about 6% with a maximum of about 12%.

Subjective problems are reported with 4 tap filtering.

JVT-C105 **[Benzler, Bosch] Block Boundary Mirroring for Motion Interpolation**
For motion compensation, the subpel values have to be generated by an FIR interpolation filter. For a 4x4 block, this means to access a block of 9x9=81 pel instead of the 5x5=25 pels which are needed for bilinear interpolation. An efficient method for reducing the amount of reference memory access was proposed. A similar method is found in the MPEG-4 Visual standard.

Nokia Exhibit 2024, p.5
Snap/Hisense v. Nokia
IPR2025-01440

Response to these contributions:

Continue Interpolation AHG with mandate to study implementation/complexity and performance aspects of sub-pel MC.

3.3. B Frames

JVT-C044 **[Gu, Vweb] Introducing Direct Mode P-picture (DP) to reduce coding complexity**
A new Direct Mode P-picture type was introduced to reduce coding complexity, in which the motion vector and mode information could be derived directly from previous normal P-picture without motion estimation and mode decision process. The decrease in residual coding efficiency was reported to be offset by the savings in motion vector and mode information overhead coding.

Proposing a syntax change. Use this on every other picture – not sent mode and MV information, which is a lot of the bits at low bit rate.

Comment: Can get same complexity reduction without syntax change – Response by proponent: This has coding efficiency improvement relative to the no-syntax-change approach (results not provided).

Subjective: not shown

Considerations:
- How low bit rate?
- Encoder complexity reduction amount?
- Error resilience?
- Complexity reduction at decoder implementation dependent.
- Concern raised for memory consumption.

JVT-C076 **[Tian+] Coding of Faded Scene Transitions**
A so-called "overlay" coding technique for faded scene transition coding is presented. Overlay coding is based on independent coding of the source sequences of the scene transition and run-time composition of the fade. This paper analyzes possibilities to preprocess the source sequences and finds that pixel-amplitude weighting methods can further improve the performance of simple overlay coding. Random access is possible with overlay coding.

Complexity can be controlled versus level assignment but the increase was considered to be difficult to estimate.

The gain of the proposal is significant but its relationship to the complexity increase remains unclear. Generality of composition approach may not be given because of Gamma correction.

JVT-C120 **[Jeon, LG Electronics] B picture coding for sequence with repeating scene changes**
When there are multiple repeating scene changes in the sequence and the co-located MB in the subsequent picture has a forward motion vector from a long-term picture belonging to the same scene group, the long-term picture is used as reference picture for B picture coding and MVf and MVb of direct mode are approximated to the forward MV of the collocated MB and 0, respectively.

Considerable gains for the affected B pictures.

The group adopted a change to the MV calculation feature, modifying the proposed method per below.

Send numbers N1, N2 and D at picture level
MVf = N1*MV/D
MVb = N2*MV/D
UVLC code each of these

Concern was expressed regarding use of division operator. Alteration to (N1*MV+offset)>>K form, possibly with zero offset is for further study.

Post-meeting remarks: Significant loss of compression quality was reported to result from the adopted change by Byeong-Moon Jeon. Further study is expected.

Nokia Exhibit 2024, p.6
Snap/Hisense v. Nokia
IPR2025-01440

JVT-C121 **[Jeon, LG Electronics] New syntax for Bi-directional mode in MH pictures**
When bi-directional mode in a bi-predictive picture is used, which combination is employed has to be specified to separate two index parameters, ref_idx_fwd and ref_idx_bwd. This document proposed introduction of a new syntax element "Multi-Hypothesis mode Indicator" into macroblock layer of bi-predictive pictures to indicate direction information for bi-directional mode.

Group response:
The nomenclature of "bidirectional picture" in the draft was changed to "bi-predictive" picture for enhanced clarity of the intent to decouple the decoding design from prior assumptions of the temporal direction of bi-predictive motion compensation.

The draft was double-checked to try to ensure clarity of the relationship between time and motion compensated prediction operation, and further work on the subject of that clarification is expected.

JVT-C127 **[Tourapis, Microsoft] MV Prediction in B frames**
In this proposal, the performance of the motion vector predictor within B frames, with regards to the B Direct mode was investigated. It was reported that the existing JVT standard does not specify how such should be performed, whereas in the current software Direct Mode motion vectors are not used for the prediction. Instead it was proposed that B Direct mode motion vectors are also used within the prediction, thus improving B frame coding efficiency up to 4.5%. Scheme also increases motion vector correlation for B frames within the encoded stream which is quite important for error concealment purposes.

Results with CABAC are missing.

Verification: software was provided.

No subjective results.

Group action: Change appeared to be a minor bug-fix/clean-up issue and unlikely to cause problems. Proposal adopted.

JVT-C128 **[Tourapis, Microsoft] Direct Prediction in P and B frames**
In this contribution was proposed a new Inter macroblock mode, which is an extension of the Direct Mode used within B frames. This new mode was reported to effectively exploit temporal correlation of motion information within a sequence, thus effectively improving performance. The implementation was reported to be relatively simple, with little complexity added on the encoder/decoder. The combination of this mode with the predictor skip mode (MotionCopy JVT-C027) was reported to achieve considerable bit rate reduction (up to 10.7% - average 3.85%). Further extensions of this mode, with the addition of submodes and the inclusion of multihypothesis prediction, were also described, and the new mode was reported to appear to also be very promising for a future implementation within B frames.

The average gain on top of JVT-C027 is 1.65% with a maximum of 5%. A problem in error prone environments was mentioned and the gain appeared not to be worth the syntax change.

3.4. Multi-Frame Motion Prediction

JVT-C010 **[Schlockermann]AHG Report: Multiframe Motion Pred.**
Active AHG activity was reported. In the new definition of the direct mode, it was reportedly not clear which motion vector to use. The storing in display order may have an error resilience impact.

JVT-C047 **[Kadono+, Matsushita] Result on Multi-frame interpolative prediction with modified syntax**
The results on MFIP (JVT-C066) were reported. MFIP shows average bit saving 14-28% for fading scene. By downloading weighting coefficients, the average bit savings sometimes exceeded 50%. Results of JVT-C067 and JVT-C103 are verified.

Noted: Multiplying by 2 requires clipping.

JVT-C049 **[Borgwardt, VideoTele.com] Multi-picture Buffer Semantics for Interlaced Coding**
Three proposals related to the Multi-picture Buffer.

Nokia Exhibit 2024, p.7
Snap/Hisense v. Nokia
IPR2025-01440

Proposed semantics for handling the mix of frames and fields in the Multi-picture Buffer: picture number (PN) is incremented by 1 for each frame.

Proposal to de-couple RPSLI and RPBT in the syntax.
Proposal to move Multi-picture Buffer re-ordering and management commands from the slice layer to the picture parameter set.

Note: assumption that the picture is stored in Adaptive Memory Control
Note: Add codewords to indicate reference buffer resampling from first slice and codeword on how to store the picture.
Note: field/frame adaptive operation analysis required.

JVT-C066 **[Kikuchi+, Toshiba] Multi-frame interpolative prediction with modified syntax**
The multi-frame interpolative prediction (MFIP) originally proposed in JVT-B075 exploits a temporal-interpolation on multiple reference frames. A modified syntax was proposed as an extension to B-picture that enables switching between the two default MFIP coefficients with no overhead. The coding gain was reportedly significant as shown in JVT-C047, JVT-C067 and JVT-C103.

Two approaches are investigated: 1) weight ½,½, 2) weight 2, -1 for fading.
Problem with > vs. >=.

Response by group:

Adopt JVT-C067 and JVT-C103 with revised syntax considering JVT-C077 into non-baseline with final decision left to Profiles and Levels discussion.

JVT-C067 **[Kikuchi+, Toshiba] Experiment result on multi-frame interpolative prediction with default coefficients**
This contribution was to report the experiment result on multi-frame interpolative prediction (MFI) proposed in JVT-B075. Automatically works with Lagrangian mode selection.

PSNR improvement are reported up to 1.7dB with simple default coefficients (1/2,1/2,0) and (2,-1,0).

For non-fading sequences, the gains with 2 reference pictures are small.

Verification ok.

JVT-C077 **[Hannuksela, Nokia] Generalized B/MH-Picture Averaging**
A generalized formula for pixel amplitude scaling according to temporal distance is presented. Under certain constraints, the formula equals to weighted averaging of B-pictures (Q15-K44) and interpolative motion compensation (JVT-B075). In addition, explicit signaling of pixel amplitude scaling factors is proposed.

Adopted as noted elsewhere.

JVT-C103 **[Kikuchi+, Toshiba] Interpolation coefficient adaptation in multi-frame interpolative prediction**
The coding performance of the multi-frame interpolative prediction (MFIP) with interpolation coefficient adaptation was evaluated. Simulation results show extremely high improvement of up to 5.8dB

Subjective results showed a significant gain.

2 multiplications and 1 shift, 1 addition, and clipping instead of 2 shifts.

Verification ok.

3.5. Motion Vector Coding

JVT-C041 **[Suzuki, Hitachi] Core Experiment Results on Adaptive MV Coding**
Adaptive MV coding sub-samples the MV table to more efficiently transmit MVD. Syntax extension of MB modes as well as the 8x8 block indication needed to signal accuracy. This contribution provides a report of core experiment

Nokia Exhibit 2024, p.8
Snap/Hisense v. Nokia
IPR2025-01440

results on adaptive motion vector coding and the revised text for the CD. The proposed scheme was reported to provide general improvements on average between 2-3% BDBR with similar performance for most cases with 0.27-4.5% for higher complexity with B frames and 0.15-3.2%.

Group: Good trade-off between complexity and gain, but not enough support to adopt the feature.
Coding efficiency gain is too small.

JVT-C051
[Sato+, Sony Corp] Result of Core Experiment on Adaptive MV Coding
The CE result proposed in JVT-B061 was presented and JVT-C041 was verified.

JVT-C097 **[Nakagawa, Fujitsu] Context-based motion vector coding with 2-Dimensional VLC**
The probability density function of differential motion vector values is strongly dependent on the motion of the local area in a frame. In this contribution, a context-based motion vector coding scheme was evaluated. The scheme utilizes statistics of the motion vectors estimated from the motion in the surrounding area and switches the VLCs for encoding the motion vectors adaptively. In addition, a 2-Dimensional vector coding was introduced to improve the coding efficiency of the scheme. Average gain of 1.3% with up to 4% for Forman.

Further work and discussion in relation to entropy coding is encouraged.

3.6. Global Motion Vector Coding and Global Motion Compensation

JVT-C011 **[Kimata+] AHG Report: GMVC**
Evaluation of GMVC/GMC and so-called "Motion Copy" reported – results were confirmed. The subject has been very extensively studied.

JVT-C021 **[Sun & Lei] Experimental results of GMVC and "Motion Copy"**
The GMVC technique uses a parametric motion model (perspective or affine) transmitted at the slice header. In the MB layer, the conventional copy mode for an MB is replaced by a GMVC_COPY mode as when GMVC_flag is ON. An additional GMVC_BLOCK mode is defined as a sub-partition mode for the 8x8 (split) MB mode when GMVC_flag is ON.

The bit rate savings are up to 20% at QP=28 and Foreman, CIF. The average bit rate savings for the test set are 2.05%. For B frame coding, the maximum bit rate saving is 4.8% with and average of 1.45%.

JVT-C022 **[Sun] Additional "zero-motion" subblock for "Motion Copy"**
The performance is slightly improved against JVT-027 with the option to skip also 8x8 blocks.

JVT-C027 **[Lainema, Nokia] Skip mode motion compensation**
This document proposed a simple modification to the Skip mode motion compensation. A motion vector is generated similarly to the prediction motion vector for the 16x16 macroblock mode and used in motion compensation. That is, no syntax changes were proposed, but the way Skip mode is interpreted is modified. Compression efficiency improvements are similar to those of significantly more complex GMC/CMVC methods. The new rule is that zero motion vector is used if:

- If a Macroblock immediately above or to the left is not available (that is, is out of the picture or belongs to a different slice).
- If a Macroblock or block immediately above or to the left that is used in motion vector prediction has a zero motion vector and uses the latest picture as reference in motion compensation.

Otherwise the predictor is used for skip-mode motion compensation prediction.

Bit rate reductions are only at the very low bit rate end up to 18% at QP=28 and Foreman, CIF. The average bit rate savings for the test set are 2.2%. For B frame coding, the maximum bit rate saving is 5.4% with and average of 0.35%. The group felt that the performance vs. complexity tradeoff is reasonable.

Group action: Adopted.

JVT-C043 **[Kimata, NTT] GMVC and GMC with B-picture**
GMC and GMVC which combined with "Motion copy" are proposed. Experiments with both P-picture and B-picture showed improvement of coding efficiency up to 9.68% BD-bit rate savings reported especially in sequence

Nokia Exhibit 2024, p.9
Snap/Hisense v. Nokia
IPR2025-01440

with zoom. The maximum (best setting) average bit rate saving over the test set was 3.36% and 4.52% over a set of sequences with global motion. Maximum bit rate saving against JVT-C027 was 3.5% with an average of 2%. The group was not clear about whether the performance vs. complexity tradeoff is reasonable. Comparing JVT-C027 and JVT-C043, the group felt that the performance vs. complexity tradeoff is not reasonable. Further improvements on top of JVT-C027 or reductions in complexity would be required.

4. Deblocking Filter

JVT-C007 **[List] AHG Report: Deblocking**
Significant activity in the area of deblocking filter work was noted.

JVT-C035 **[Le Maguet+, Philips] A Low Complexity deblocking for JVT**
This contribution succeeds to the contribution JVT-B037. A low complexity deblocking algorithm was proposed, referred to FADA (for Fast Deblocking Algorithm) in the rest of the document. This new algorithm was implemented in the encoding loop and compared to the JML1.9 deblocking in terms of both quality and complexity. The contribution reported that FADA was equivalent to the JML1.9 deblocking in terms of perceived visual quality. Complexity gain was reported between 1 and 2.2 depending on the target platform.

Presentation: 1) 4x4 classification based on luma into categories: homogeneous, row, column, texture, 2) filter selection: no filter, one, two, or three samples each side, 3) filtering

Same filter applied for each four-sample edge.

Encoder indicates on picture-by-picture basis whether "same MV with zero residual" edges are filtered or not. Demo of artifact on frame of Tempête. Another example verbally: blocking artifact present in one frame with subsequent zero motion vectors can produce persistent blocking artifacts. (Comment: Encoder could possibly detect and correct in some other way.)

Complexity analysis of 1) CPU cycles, 2) min, max, abs type instructions, 2) parallelization. Branch treated same as add (but asserted that the proposal might get better relative complexity if more penalty for branch).

Comment: Worst-case versus average complexity? Do we have consensus on the complexity analysis?

Cycle count depends on weighting of instruction types: according to the presented analysis from 1.0 to 1.5 complexity ratio on "conventional platforms" (non-SIMD), 2.2 for SIMD/hardware with parallelization. (Proponent estimating deblocking as 40% of decoder, such that factor of 2.0 would be a 20% total complexity benefit).

PSNR common conditions & JM 1.9: (perception emphasized) average loss BD-PSNR 0.5 dB (0.1 to 0.7), BD-Rate: average -10 (-3 to -14).

Subjective results using approx ITU-R Rec. BT.500 DSIS method 16 people, using common conditions subjective recommended results plus some other sequences, summarized as essentially the same overall quality. (A smaller test indicated no problem with interlace)

Verification: Software provided to AHG a month ago. UB Video and Nokia have tried it (see section on UB Video document for relevant comments)

JVT Group Response

Needs more analysis. At this stage no agreement that this indeed reduces complexity for all architectures.

Needs more work :
- Memory size, ASIC architectures, Pentium Based architecture
- Needs more evidence that complexity is indeed reduced
- Drop in SNR and Bit Rate is a concern.

After Demo: No consensus to accept it. Suggest conducting more experiment on wider classes of video sequences.

Nokia Exhibit 2024, p.10
Snap/Hisense v. Nokia
IPR2025-01440

JVT-C056 **[Conklin, RealNetworks] Dithered 5-tap Filter for Inloop Deblocking**
This document proposed a change to the 5-tap filter used in in-loop deblocking. By using a pseudo-random addition factor in the filter definition, a small (+/- 3/8 LSB) dither was applied to the filtered image. This dithering effect was reported to improve visual quality for two reasons:
1. Providing effective smoothing between blocks that differ by a single luminance level
2. Visually "breaking-up" the block structure that can remain even after filtering
While the proposed scheme was reported to improve visual quality, the effect on bit rate and PSNR were reportedly negligible. To further improve visual quality, additional results were presented using a recursive version of the 5-tap filter.

In current JVT design has 5 tap filter now used for smooth intra 16x16 regions. Can still see blockiness on some displays. Particularly annoying in smooth areas.

Change ordering of which edges are processed first. (Complexity impact? Maybe cache issue)

Dither is in offset of result of filtering process before right shift, to break up the structure of the artifacts.

Complexity: Enables filtering with sample difference of 1. Adds a small table look-up (2 tables of size 16 nibbles) during filtering. Does the change of order of which edges done first affect complexity? Note: This only affects intra 16x16 filtering on edges classified as smooth.

Recursive idea – some further visual benefit. Use results of filtering of samples closer to edge when filtering samples farther from edge. Carries the dither range out further to create softer appearance. Complexity: creates an additional data flow order dependency, otherwise nothing.

Possible concern expressed over injection of high-frequency noise into scene. However, noted that the dither magnitude is only 3/8 of 1 LSB before rounding, also noted that subsequent motion compensation would be likely to blur out the high frequencies.

Still-picture demo projected without enhancement of artifacts (a number of examples are in document): Subtle but visible improvement. Natural and artificial examples.

Tests with I period of 15 – did not do BD measurements (yet). Very very small differences in tests.

Complexity based on ref s/w execution time on PC negligible diff. Not a really meaningful measurement.

Verification: S/W provided to ad-hoc group roughly a month ago.

Discussion:

Need to see demo for Moving Pictures – for two sequences shown quality was little better (smoother but visually better).

Improves the sequences with large flat areas.

Need to compare against Post Filtering

May add some burden to low-end processors that exist today. Future?

JVT Group Conclusion

Conditional acceptance – need to do more tests for lower Qp values and more number of sequences.

JVT-C094 **[Au, UBVideo] Complexity Reduction and Analysis for Deblocking Filter**
This contribution proposed a simplified JM 1.9 deblocking filter that was reported on average 23% faster for implementations on media processors while providing equivalent or better subjective quality. Complexity analysis was performed for both simplified and original filters and subjective testing results were presented. IPR 2.1.

Look at media processor (e.g., TI C64x processor or MMX) savings roughly 23% savings of filtering part of the design (for the TI platform) based on their analysis of the complexity.

Nokia Exhibit 2024, p.11
Snap/Hisense v. Nokia
IPR2025-01440

Default filter use input to filter rather than output of filter to allow parallelization

Remove clipping of delta value in "default" filter (makes the filter stronger, sometimes better, sometimes loss of edges, reportedly usually subj better not to do the clipping).

Simplify the filter switching conditions between strong and default filter for intra.

Dependency graphs shown. Critical path shortened.

Estimated: 48 cycle worst case cut to 20 cycles

Each of the three simplifications reported to give roughly equal complexity reduction (8% each).

Proponent estimate essentially same complexity in this proposal as in Philips proposal.

For C implementation on P-III got ~5% savings.

Common conditions: Overall BD-PSNR: 0.08 BD-PSNR loss, 1.7 % BD-Rate increase.

Demo: Foreman softer on proposed method – perhaps slight preference for proposed method. Tempête also softer foreground. Tough to judge – need better demo conditions. More demos to be seen.

Reported subjective results: Rough equivalence (Slight preference to proposed method).

Software: Avail a week ago to AHG. JM1.9.

Verification: Nokia tried their s/w and found essential equivalence in visual quality (very slight preference if any to JM method). No clear feeling on the complexity estimates.

Comment from proponent re Philips proposal is pretty similar but Philips sometimes not as good (particularly QCIF esp Foreman). Horowitz may be able to comment on complexity.

Comment: foreman qcif for example not so good for Philips proposal.

Comment: Foreman is the most complex one – the most important case for deblocking in the common conditions.

Group action:

To adopt modification 1 (removal of recursiveness in "default" filter, and modification 3 (simplified switching for Bs=4, which is I-macroblock boundary).

Further work to be conducted to see if some parts of C130 can be included.

JVT-C111 **[Foo] Revised Downloadable Thresholds for Loop-filter**
Reportedly improving threshold values for subjective video quality.

Syntax added to specify threshold values for decoding

Proponent indicates current thresholds optimized for CIF picture size. Larger thresholds seem to work better on both smaller and larger picture sizes.

Some examples shown of artifacts alleviated with downloadable method.

Parameter set transmission of threshold differences from default values. In proposal document, these are coded as FLC-coded positive difference from default with limited range of difference. Alternatives would be to VLC encode the difference or allow negative differences.

JVT Group Discussion

Nokia Exhibit 2024, p.12
Snap/Hisense v. Nokia
IPR2025-01440

Need more evidence.

Group likes the concept but no clear consensus how often to download (Slice Layer downloadable?).
Peter List: Fixed number of switchable tables is desirable. How many? TBD. Different Profiles will have option either to adopt it or not.

Group also recommends to re-visit the current alpha and beta tables (Foo, Joch UB Video, Peter L)

Not accepted in this meeting. If further study results look encouraging accept in July meeting but not in baseline.

JVT-C112 **[Foo+, Panasonic] Automatic Selection of Threshold Values for Downloadable Threshold Tables**
A program for automatic selection of threshold values for video sequences is presented here. Different threshold values are reported to optimize the subjective video quality of encoded video. Encoder-only optimization for C111

Choose a number of frames (e.g., 5 frames at beginning of sequence), code them using different threshold values, choose set with highest PSNR. Encode entire sequence using those thresholds. (Noted: Eyes are still better than this for threshold selection.)

JVT-C130 **[Gomila, Thomson] Simplification of the JVT deblocking filter**
This contribution proposed a simplification of the deblocking filter currently in the draft standard. The aim was to reduce the number of line-based operations, which are reported to be by far the most demanding term when evaluating the computational cost of filtering. On average, the proposed method was reported to use on average 47.81% fewer operations than the JM 1.9 deblocking filter, with minimal impact on visual quality and PSNR.

HVS argument for reducing processing in very dark and very bright areas (Y below 48 and above 232).

Link chroma and luma filtering (reduce computations to determine chroma filtering conditions). Lock together so filter chroma if and only if filter luma.

Avoid distinction between stronger and normal filtering – change the stronger F6 filter to F4, change the chroma length-4 filter to F2. F# indicates number of samples that change on either side of edge.

Chroma: (a + 3*b)>>2

Reduce worst case (shorter filter for maximum filter length, lack of distinct chroma decisions)

How much of total filter is the line-based operations? (comment: roughly 2/3)

Question is filtering on 2x2 basis or 4x4 basis in current design? Proponent says 2x2. Need an answer on that.

Question about validity of total cost savings estimates whether these weight the luma and chroma processing cycles properly? Appears correct.

Demo foreman qp=24: showed some chroma artifact in new method, overall roughly equal otherwise. Moving demo: some additional chroma artifact. Silent some increase in artifacts. Mobile qp=24 roughly same quality qp=24. Tempête perhaps some increase in sky blockiness (not everyone saw that).

Some quality degradation allowed.

Proponent comment that with lower QP it's harder to see the difference and the cost savings is more.

Comment: Roughly 30-50% of overall decoder is deblocking filter.

Note Less memory needed for storage of edge values.

Common condition: BD-Rate: 1% BD-PSNR: 0.05 dB penalty. (Measured on luminance only.) 0.3 dB penalty in U and V for qp=16.

JVT Group Discussion:

Nokia Exhibit 2024, p.13
Snap/Hisense v. Nokia
IPR2025-01440

Need more analysis for wider classes of Video
Reduces memory requirements
Concerned about reduction in quality

Group likes some of the ideas but hesitant to accept the whole proposal.
Not accepted at this meeting. Work will start to combine some of the ideas here with JVT-C094 (Thomson – Gomila, TI – Minhua, UB Video – Joch)

5. HRD/VBV Buffering Model

JVT-C020 **[Viscito] AHG Report: Buffering**
Issues noted included how to account for header data, timing, low-delay versus non-low-delay.

JVT-C166 **[Hannuksela] Desc. of Hypothetical Reference Decoder (HRD) and Buffering Verifiers**
Eric Viscito & Miska Hannuksela produced a document for consideration on a definition of HRD for VBR and CBR and low delay mode. Includes method of handling header-level data. CBR can be same as VBR with average rate equal to peak rate. Includes explicit film mode state for 3:2 pull-down.

JVT Group Discussion

Incorporate the HRD Syntax and text discussed during the plenary into the CD in addition to prior HRD text.

Do not include special splicing features.

Adopt except for explicit film mode state for 3:2 pull-down – that issue to be resolved in further study as relates to JVT-C138.

JVT-C046 **[MacInnis, Broadcom] NAL for H.264 with MPEG-2 Systems**
This document proposes the specifics of an MPEG-2 Systems NAL for use with H.264. The proposal is reported to be consistent with the current normative agreements of Draft H.264, and it attempts to harmonize other H.264 concepts and suggestions with existing video industry practice and MPEG-2 Systems and Video specifications. Due to the existing very wide deployment of systems, devices and software which utilize MPEG-2 Systems and MPEG-2 Video, and the resulting opportunity for large scale deployment of H.264 in such systems, the NAL to package the proponent argues that H.264 video in MPEG-2 should conform as closely as possible with the outer layers of MPEG-2 video, subject to the efficiency goals and semantic requirements of H.264, while the MPEG-2 Systems layer is recommended to be unmodified from its original specification, except for the unavoidable assignment of stream ID's and similar minor modifications.

Noted also contains some HRD content.

Comments from the group:

- The proposal resembles JVT-C055 to some extent.
- What is the difference between "Type-2" and "Type-1" slices? Why are they both needed?
- The proposal does not follow the spirit of the current JVT high-level syntax based on parameter sets.
- The proposal is not optimal for cross-network gateway design.
- How is the proposal going to affect profiles?
- It was requested to create a table about the critical requirements of the "TELs" being defined (H.320, ISO media file format, RTP, and MPEG-2 Systems).
- Fundamental difference between RTP-oriented TEL and MPEG-2-oriented TEL: the lack/presence of picture header. Picture-synchronous data could be repeated to ensure its correct reception in error-prone environment.
- There were concerns about the overhead of carrying the current JVT slice header in each slice in error-free environments. Agreed to leave the slice header as it is, because the overhead was considered to be moderate.
- Picture header can be useful to find a complete picture. However, a flag at the beginning of the slice header or a picture start code alone or a bit in the PTIB/NALP_type could indicate similar functionality.
- In MPEG-2, vbv_delay is a part of each picture header. In JVT, vbv_delay could be carried in an SEI message, as it is not crucial if it is received for each picture.

Nokia Exhibit 2024, p.14
Snap/Hisense v. Nokia
IPR2025-01440

A fixed-length slice type in a fixed position of the slice header was proposed.

Also has some slice types in it.

Further discussion: Type 2 and Type 1 are not the same.

Proponents were encouraged to bring the evidence of the benefits of having Type 2 Slices. If evidence is found to be satisfactory during July meeting, it will be accepted.

JVT-C069 **[Hannuksela, Nokia] Levels and HRD**
This contribution proposed processing and storage space level definitions for the JVT codec. A Hypothetical Reference Decoder (HRD) including a low-delay mode was also proposed.

Current working draft assumes clearly-established decoding time.

Proposed: Slice arrives as a chunk into HRD at time corresponding the time of the last bit reception of its content.

Post-decoder buffer of uncompressed pictures holds pictures until no longer needed for reference or display.

Each NAL packet has a hypothetical transmission time (time of the last bit arrival for the packet) at the bit rate.

Note: also some levels definition content

Buffer inspected at base clock tick rate.

NAL packet including its header goes to buffer.

Low-delay mode: picture removed once it has arrived.
Normal mode: removed at its decoding timestamp.

Normal mode: space must be available in post-decoder buffer
Low-delay mode: (in proposal) same – can this cause a problem of decoder getting ahead of the content?

Proposed to send initial post-decoder buffering period.

Must a picture arrive by its presentation time? (e.g., MPEG-2 low-delay mode).

Some syntax into parameter set, some to SEI. Operation points and post-decoder buffer size sent at parameter set level.

Do we need post-decoder buffer size

Maybe not send pictures in decoding order?

"low delay" means not necessarily completely reconstructing the input relative timing (e.g., MPEG-2 low delay or H.263 or H.261 HRD's) at the output. "Non-low-delay" operation means the output of the HRD decoding process will preserve the relative timing of all TR's (PTS's).

Distance between PTS's must be at least the equivalent of the MPI.

Question whether "low delay" rules out the need for picture reordering?

SEI for initial buffering period can send for RAP. Pre-decoder for each operating point and post-decoder as a global parameter.

Requirements for encoder and decoder compliance specified.

Disposition captured in actions noted above and with regard to JVT-C166.

Provide text regarding post decoder buffer for tentative adoption subject to profile restrictions.

Nokia Exhibit 2024, p.15
Snap/Hisense v. Nokia
IPR2025-01440

JVT-C163 **[Fautier, Philips] Concept of average rate.**
This document discussed the concept of 'Average bit rate concept for VBR transmission', which is reported to not exist either in AVC, nor MPEG2.

Further study is needed to determine the impact of this contribution on the design.

6. Transform and Quantization

6.1. Transform coding

JVT-C024 **[Kerofsky, Sharp Labs] Notes on the JVT IDCT**
Summary: Cross verifies the equivalence of IDCT implementations and the dynamic range of calculations

Information to help implementers. Good to document this helpful information and put alternative examples in software.

JVT-C025 **[Kerofsky, Sharp Labs] Modifications to the JVT IDCT**
Summary: Proposes modifications to the normative IDCT and dequantization definitions.

Modification described but stability of current design was preferred.

JVT-C102 **[Kim+] Thresh. Adjust. for Adaptive Use of Double Scan**

minor change from B081 due to the change in qp dynamic range.

Proposed in Geneva – repeated in JM2.0

currently single/double scan triggers on QP = 24.

slice based threshold adjustment.

qp out side of [20,28] the scan choice is obvious.

around qp = 24, one gets up to 4.5% bit rate savings.

Delta(thr) per slice. the delta is taken off of a fixed value of 24 so that this mode can be turned on and off.

We lost a bit of gain from B081 to C102

Encoder choice was made in 3 passes:
simple, double, then did a compare

Note relates to JVT-C117.

No consensus to accept.

<u>JVT Group Conclusion:</u>

Gains only for Intra frames/modes. Overall gain for the entire sequences are very small (reported 4.5% gain is for Intra frames only)

No consensus to accept it. Later, double-scan was removed from the draft after consideration of other contributions.

6.2. Quantization

Nokia Exhibit 2024, p.16
Snap/Hisense v. Nokia
IPR2025-01440

JVT-C023 **[Sun] Lossless Coding and QP Range Selection**

Original document submitted as information only.

Provides results of a lossless coding scheme.

Defer to consideration next meeting.

JVT-C045 **[Gu, Vweb] One bit added to indicate universal quant for the whole picture**
Recommendation to add one bit at picture header to indicate that a universal quant will be used for the whole picture, and that the quant fields in the macroblock layer will not exist.

Parameter set or slice header syntax proposed to be added to support this (if slice header perhaps constant per slice instead of constant per picture).

Note would save one bit per coded macroblock with DCT coefficients.

Estimated savings of 1-2%. Slice level. Bug fix/cleanup category. After discussion and in consideration that the reported bit rate reductions are not for the common test conditions and that no verification was provided as to the quantity of bit rate savings and that the estimated bit rate savings was small, there was no consensus to accept this proposal.

JVT-C053 **[Suzuki+, Sony] Weighting Matrix for JVT codec**

Proposed addition of step size adjustment in frequency-dependent manner providing similar functionality as in MPEG-2 video design. Improved coding efficiency was reported, particularly in perceptual terms.

Implemented in JM 1.7 software.

Offset of QP for each coefficient with clipping for extreme ends of range. Six matrices Y, Cb, Cr for intra and inter sent at picture level (parameter set). May be OK to use same matrix for Cb and Cr (this is the MPEG-2 design).

Compared to current method using converted MPEG-2,4 weighting matrices (0 to 12 for intra, 0 to 6 for inter) – flat for chroma. Chroma used same step size offset relative to luma DC as in current design. Since most luma coeffs use significantly larger step size than the DC term in this weight matrix, this decreases chroma step size relative to average luma step size.

Not compared to extended deadzone encoding-only technique.

Verification: No. Software availability: not yet.

All-intra demo (non-common-conditions sequence) – method proposed looked better. IPPP… (same sequence) – method proposed looked better.

Perceptual: Slight preference about half the time.

Working on testing of an encoder trick to avoid the need for this feature to be supported in decoders. Encoding trick is to use normal size dead-zone for DC coefficient, gradually increasing that deadzone size toward doubling the deadzone size at high frequency.

Complexity impact of table selection was discussed.

Asserted by some that smart encoding can avoid the need for this.

Nokia Exhibit 2024, p.17
Snap/Hisense v. Nokia
IPR2025-01440

Further work requested by next meeting.

JVT-C136 **[Sullivan, Microsoft] Rounding, QP Origin, Dynamic Range, and |f|**
This contribution advocated changes to rounding behavior and to the nominal origin of the quantization parameter for specification and transmission purposes, and proposed a refinement of the dynamic range spec for intra 16x16 mode luma DC transform coefficients. A remark was also provided on potential improvement of R-D performance by making |f| larger in the example encoder.

Proposed item 1. Editorial change of QP. has no effect on bitstream. Note: There's some concern about the low end of the QP range (QP < -10 causing a problem for Transform dynamic range) . Proposed change adopted.

Proposed item 2. QP transmission: parameter level -> delta to the middle,
slice -> delta off of that, mb -> delta off of that.
Agree to initialize at slice level to middle of range. Check current design at MB level and use that with mid-range prediction for slice level.

Proposed item 3. (Intra Plane Pred rounding). Adopt the first of the suggested alternative variants.

Proposed item 4. clean up JVT-B38 eq 9-5 to 9-7 with a simpler offset so that one can use a shift instead of a careful divider. There is a difference in the way rounding takes place when the value is exactly in the middle. there is a further shift down by 5 bits at a later stage making this difference very minor. Adopted.

Proposed item 5. Similar to item 4. Adopted.
Proposed item 6. Luma dc 16x16 data has a dynamic range problem for qp < -6. before one does the down shift in the decoder one might need 17 bits.draft 8 doesn't limit the range, allowing 17 bits before the downshift.
not a common intra: only an issue in intra 16x16.

Proposal suggests considering to require encoder to not send data that would require the decoder to use 17 bits of accuracy (before the downshift)

This way, EVERYTHING would be 16 bits, even intermediate calculations.

Conclusion: Adopt subpoint 2 of item 6, not the rest of item 6.

7. B-Picture related divide by 2 rounding alteration. Not adopted.

8. |f| offset added before the quantizing. Proponent reports that |f| may currently be too small. Proponent thinks increasing this could give us FREE coding gain.

We'd like someone to test this idea. Noted that item 8 is just an information suggestion – no action.

6.3. Intra

JVT-C033 **[Zhou, TI] Scalable Intra Prediction**
In contribution is proposed a complexity-scalable intra prediction scheme for H.26L video coding to reportedly satisfy the different needs of low- and high-end H.26L devices. The scheme is based on a design of a memory scalable probability table that allows the scheme to operate in low-complexity or high-complexity mode. In the low-complexity mode, the scheme only needs to store a sub-table of the probability table and the entropy table used for encoding the selected prediction modes. While in the high-complexity mode the entire probability and the entropy table has to be stored. The high-complexity mode is fully backward compatible to the low-complexity mode. The experiments reported that the

Nokia Exhibit 2024, p.18
Snap/Hisense v. Nokia
IPR2025-01440

proposed scheme provides the comparable coding efficiency with the current nine-mode based and previous six-mode based intra prediction scheme when operating in the high- and low-complexity mode, respectively.

Nine intra prediction modes now in draft, 10x10x9 table and additional 9x9 table.

Previous design had smaller tables (7x7x6 and 36).

Theoretically a nibble in each entry – in practice normally a byte.

If encoder uses a subset of the 9 modes then performance measured to be worse than the former 6-mode method.

Proposal to have both a lower and higher complexity type of operation. Lower complexity mode would be the pre-Geneva design except for the ordering on the tables.

Proposal to constrain the high-complexity mode such that the first six entries in each list selectable from the low-complexity values are from the low complexity table.

Confusion over correctness of improvement of 9-mode method versus 6-mode method should be tested.

Question regarding what training set used to design the tables? Answer: Superset of test set.

Software? 1.9, avail as part of contrib.

Verification: Not yet.

Common conditions? Did not do BD measurements.

Proponent did not use BDBR savings. Instead, overall bit-rate reduction was measured.
Concerns about the memory consumption were raised and the complexity vs. coding efficiency trade-off was questioned.

Further work on reducing complexity is requested in an AHG on INTRA frame coding complexity reduction (Chair Marta Karczewicz).

This is considered as a cleanup and further work is encouraged. No change to WD2 was adopted at this time.

JVT-C118 **[Hoang, Globespan] Improved Chroma Spatial Prediction for Intra Coding**
Subjective tests reported to have shown that chroma blockiness can be present in JM 2.0 coded video material, particularly in the presence of highly saturated chroma. In this document was proposed an improved method of chroma prediction that was reported to increase subjective quality. One variant of the proposal requires a syntax change.

Use intra 16x16 methods from luma on chroma, adjusting the plane prediction for the size of the block.

Common conditions all intra except no frame skip except for skip 5 in Paris.

Visual example shows chroma blockiness can occur.

Reported that it helps to turn off the trick that throws out isolated chroma coefficients.

No chroma prediction mode indicator is supported in the current draft.

Nokia Exhibit 2024, p.19
Snap/Hisense v. Nokia
IPR2025-01440

Basically this uses the 16x16 luma prediction method on the 8x8 chroma.

Measurement method described in document designed to focus on chroma.

Best example: News sequence reduce total bit rate by 3% while increasing chroma PSNR by 0.2 dB keeping luma fidelity constant. Never worse.

Average across sequences: 1.5% reduction with 0.06 dB chroma fidelity gain.

Impression is favorable.
Interest in more test results.

Complexity: Probably not a significant issue.

Cross verification: Not yet

Software: not yet.

Code is available now.
Could those prediction modes be inherited from luma?
Clearly visible for News.
Considered as a cleanup. Further work is encouraged. Consider within INTRA AHG.

JVT-C151 **[Yu+, Divio] New Intra Prediction using Self-Frame MCP**

Using motion compensation for intra prediction. Motion search in previously-decoded area of current picture (prior macroblocks in raster scan order). Where does the (0,0) motion vector indicate prediction from? (right where the decoded MB sits) Note: May be able to improve the coding of the MV values.

I picture is redefined to be a P picture with a single reference picture that is the previous parts of the reference picture in raster-scan order (including the MV prediction). Search range +/-32. Could fill the empty region with 128.

More savings % at lower bit rate. Up to 25% or 1+ dB (qp=30 for foreman).

Bug report for motion search range +/- 48 or 64.

Results appear promising.

Test conditions coded the 1st 30 frames of each sequence. Used common sequences (I frames only).

Up to 14% BD-Rate savings (or 0.8 dB BD-PSNR) – no cases of negative impact. Esp. Foreman, also Akiyo. CIF better than QCIF. Overall around 5%.

Note sequentiality constraint on encoding process.

Specify that reference picture is prior to all deblocking operations.

Takes advantage of motion comp engine power that could go unused otherwise in I pictures.

Results reportedly obtained from decoded bitstreams.

Encourage verification and software availability.

Code a B picture that self-references??!

Nokia Exhibit 2024, p.20
Snap/Hisense v. Nokia
IPR2025-01440

Perceptual?

Further work highly encouraged.

Lack of maturity. Missing subjective confirmation. No support from the group. Various interactions with the loop filter are unclear.

6.4. ABT

JVT-C012 **[Wien] AHG Report: Adt'l Transforms and Quant.**

Noting contributions on the ABT subject.

JVT-C106 **[Benzler, Bosch] Adaptive Block size Transform**
For high quality applications, Adaptive Block size Transform has reported to achieve an improved subjective quality compared with the JM. In combination with other tools that improve coding efficiency for high quality applications, ABT is proposed for the Broadcast Profile in the JVT standard."

JVT-C107 **[Wien, RWTH] 16bit Adaptive Block size Transforms**
In this paper is a revision of the ABT proposal according to the current working draft WD2. The transform design is modified to fit into 16 bits. The new JM 4x4 transform and an 8x8 transform proposed by Gisle Bjontegaard are used. The new design was reported to exhibit less ringing artifacts and superior subjective quality compared to the TML9 proposal. The CABAC encoding mechanism is redesigned and a VLC solution is proposed. The intra prediction is extended to 9 prediction modes as adopted at the last meeting. The accompanying contribution JVT-C108 provides the integration of the ABT proposal into the JM20 software. In this paper, the coding elements needed for ABT coding are defined."

Prediction filtering for I macroblocks (similar to modified scheme adopted for current JVT design in Geneva)

>>2 vs. /4 issue is noted.

In this case "16 bit" transform means that intermediate values may exceed 16 bits (different that interpretation of "16 bit" for current 4x4 JVT transform, which does not require more than 16 bits for intermediate results) before storage.

Note: there is a contrib. from L. Wang on scan order for ABT.

Possible problem with length-8 dynamic range when step size is very small – not sure exact impact (could just turn off for such use).

Adjust deblocking filter to be somewhat stronger filtering for bigger blocks (was not in Geneva proposal).

New VLC encoding shown based on Golomb codes with 00…1 prefix and variations of following length that can be selected.

Software available (update to be provided with bug fixes that should improve quality).

Motion search range small (**+/- 16**), II II II … and IB BB BB PB BB BB PB …

Somewhat less PSNR gain than previous Geneva proposal (due to improved intra perf from last meeting for intra prediction, new macroblock segmentation, improved CABAC).

Nokia Exhibit 2024, p.21
Snap/Hisense v. Nokia
IPR2025-01440

Interlace common conditions set with CABAC CCIR I5BP 1.6% to 9% BD-Rate saving, CCIR III 2% to 6%, HHR I5BP 1.5 to 6%, HHR III -0.2 to 4% savings.

Interlace with VLC not available.

Progressive common conditions with VLC (could provide also with CABAC) IBBP 0.4 to 5% BD-Rate savings, III -0.5 to 3%.

Believe ringing and blocking artifacts are now less. Blocking less due to minor mods to deblocking. Ringing better due to transform design and new macroblock prediction structure.

To show subjective demo.

Software available.

No cross-verification (Geneva proposal was cross-verified, partly independent). This time someone else has run the same software.

JVT-C108 **[Wien+, RWTH] JM2.0 ABT Software**

Software to accompany JVT-C107

JVT-C140 **[Wang] Alternate Coefficient Scanning Patterns for Interlaced ABT Coding**
Adaptive Block Transform (ABT) coding for JVT has been proposed as a method which can improve coding performance for interlaced and high resolution video. In [JVT-B053], "traditional" zigzag scans are used to convert two-dimensional coefficient data into one-dimensional data. This contribution looks at the use of alternate coefficient scanning patterns for interlaced material in order to improve the ABT coding efficiency. Preliminary results with alternate scanning patterns were reported to show bit rate savings of up to about 7%.

Summary of outcome on ABT:
The proposal to add ABT capability appears to be technically sound and the results are sufficient as of JVT-C140.
Adoption of C107 and 140 subject to verification.contingent on Profile discussion result.

7. Source Formats]

JVT-C116 **[Sullivan, Microsoft] A default color space for JVT video**
The current JVT draft does not define an operational color space. One such color space is proposed, and it is proposed to be the default color space definition for JVT video. The proposed color space is consistent with ITU-R BT.709.

Also: terminology change: use “luma” and “chroma” on non-linearly transformed color.

There are some precision issues (3 digits of significance) in the various standards (such as 709)

The proposal is to specify the default as ITU-R BT.709 (601 variant of 709). Others could be dealt with others by signaling in the bitstream. but where? in the parameter set data.

MPEG-4 default spec is in according with this proposal EXCEPT for the matrices (MPEG-4/2 defaults to the HD variant of 709)

<u>JVT Group recommendation:</u>

Nokia Exhibit 2024, p.22
Snap/Hisense v. Nokia
IPR2025-01440

Accept that there is a need for a defined color space. Group decided to change what was proposed to provide the same flexibility as in MPEG-2 and MPEG-4 with no default.

Specified at Sequence Layer.

JVT-C135 **[Sullivan, Microsoft] Pixel Aspect Ratios**
Pixel aspect ratio definition is found in Annex C of the current JVT draft. This proposal advocates support of 5 special values of pixel aspect ratio and also custom arbitrary width and height numbers that shall be relatively prime. The 5 special values are 10:11, 12:11, 40:33, 16:11, and 1:1. This design is the same as found in H.263+ and MPEG-4 visual. The purpose of having special enumerated values is to discourage small variations of similar numbers. The proposal advocates restricting the width and height of the pixel aspect ratio to be relatively prime, and disallowing zero values for these parameters. A list of preferred uses of pixel aspect ratios is provided."

JVT Group Recommendation :

Accept. Information inserted at Sequence Layer along with the "cropping rectangle" information at the Sequence Layer.

Add 3:4, 2:1 and 2:3 variations of the proposed aspect ratios.

Use 8 bits (instead of 4).

Use 0000 0000 to indicate "unknown".

Other specific numbers to be specified? Provide inputs in July or later.

JVT-C137 **[Sullivan, Microsoft] Cropping, Generalized Pan-Scan, Source Formats**
This contribution advocates the adoption of generalized pan-scan support, notes that 4:2:2 is a stated but unfulfilled requirement of the JVT video codec project, suggests an initial design for 4:2:2 support, and discusses some thoughts on 4:4:4, 10-bit and 12-bit support, and support of more than three video components. Most of the comments in this contribution echo those made in D.50 at the November 2000 meeting of ITU-T SG16 - a contribution that received a favorable reception at the time.

Proposal subject 1. Proponent voices support for arbitrary cropping (Annex C) per picture. The cropping rectangle is the basis for the output image.

Proposal subject 2. Let the encoder send an arbitrary number of pan/scan rectangles (which is a rect relative to the cropping rect), not necessarily tied to an display aspect ratio (as in MPEG-2), tagged with IDs that are assigned somewhere up in the system level.

This will not conflict with the system layer. A system level scene composition will tell the decoder what to do with these if any. If the system layer has similar information it can override the information in the video layer.

This information will be supplemental enhancement info, and as such, can simply be dropped on the floor and the decoder will still be conformant.

Proposal subject 3: 4:2:2.

420 is reported to be bizarre for interlace. the chroma samples fall between scan lines spatially and temporally.

Nokia Exhibit 2024, p.23
Snap/Hisense v. Nokia
IPR2025-01440

Requirements doc specifies 422. the proposal suggests a cheap way to incorporate 422, not aimed at optimality.

- cbpc indicates presence of coef in either 4x4 chroma blocks
- don't reduce vertical chroma mv
- increase the QP of chroma relative to luma by 2 (or some other amount).

further investigation is recommended.

Proposal subject 4. Potential Extensions

Not in the requirements doc. Not considered at this time.

Potential extensions on a slower schedule.

4:4:4
4:4:4:4
4:4:4:4:4:4

10-bit, 12-bit… pretty easy to spec it out.

JVT Group Recommendation:

Accept the Cropping and Generalized Pan and Scan information at Picture Layer.

4:2:2 is not mature at this stage. Needs further work.

8. Interlace Coding [TW]

JVT-C013 **[Borgwardt+] AHG Report: Interlace**

JVT-C042 **[Suzuki, Hitachi] Core Experiment Results on Interlace Chroma Phase Shift**
Verification of JVT-C050 is indicated with final check on BD values.

JVT-C050 **[Sato+, Sony] Result of Core Experiment on 4:2:0 Interlace Chroma Phase Shift**
In frame mode, for certain vertical MVs, a chroma phase shift is motion-compensated. Proposal is to spatial interpolate the correct chroma position using the WD2 ¼-pel filter. Problem is somewhat suppressed by RD-opt in MB mode decision. For RD-opt on, the proposal provides PSNR gains between 0-0.5dB corresponding to bit rate reductions of 0-10%. Not the complete test set used. The CE result proposed in JVT-B068 are presented.

Further investigation is encouraged on this bug-fix.

JVT-C138 **[Sullivan, Microsoft] TR's, PN's, Pictures, Frames, and Fields**
This proposal contains a design for the representation of time and the management of multi-picture buffer contents for the JVT codec, including discussion of the use of interlaced frames and fields as well as progressive-scan pictures.

Temporal references should be based on a general "timescale" parameter as found in Annex C of WD2r7 and JVT-B109, or should be based on a 27 Mhz clock (not a 90 kHz clock)
Anybody anticipates a problem with using that clock ?
Recommended for adoption – later modified to support general clock, not just 27 MHz.

Nokia Exhibit 2024, p.24
Snap/Hisense v. Nokia
IPR2025-01440

Dangling fields support.
Recommended for adoption.

Each field in a frame may (or shall) have its own temp ref.
Recommended for adoption. Some details about how it works for a frame needs to be worked out.

Work on pic num for interlaced coding.[revisit for this]
Each field should have its own temporal reference: problem with direct mode prediction … work out a solution and present later.
Simple repetition possible: permit to put it into multiframe buffer. Also have a simple repeat indicator for a field and a frame and a slice? Work out in high-level syntax. No, use the skipped mb run to signal that.

SEI to signal SMPTE compatible time code representations.
Recommended for adoption.

JVT-C139 **[Wang, Motorola] Macroblock adaptive frame/field coding for interlace sequences**
This document presents the results of MB-level adaptive frame/field coding for interlaced video materials, and the performance comparisons with picture level adaptive frame/field coding. The simulation results show that MB level adaptive coding can provide additional gain over picture level adaptive coding for sequences that favor frame coding. It is also suggested that MB level adaptive frame/field coding can be integrated in picture level adaptive coding. MB/picture level adaptive coding guarantees the performance over both MB level and picture level adaptive coding.
BD-BR savings are reported up to 14%.

Work out plan for further work and possible inclusion in July.

9. Performance, Implementation, and Complexity Analysis

A joint session was held with MPEG on complexity issues:
M8347 portability of C code, ANSI C compliance.
Desire to convey the info to the JVT s/w coordinator.
M8378 Complexity analysis "simple", "medium", and "high" complexity settings. Analysis of memory access, speed, etc. Memory footprint is big.

Desire to have prototype control files: Make part of the s/w package.

MPEG's AHG on complexity analysis: jvt-complexity-list@imec.be subscription: owner-jvt-complexity-list@imec.be. Please subscribe.

JVT-C148 **[Joch, UB Video] Demonstration of a Computation-Optimized JVT Codec**
This contribution presents results from a computation-optimized JVT codec that was demonstrated at the meeting. For CIF content on an 800 MHz Pentium III PC, the codec achieves speeds of 49, 137 and 36 frames per second for encoding, decoding, and simultaneous encode and decode, respectively. Subjective testing has reportedly shown that this reduced JVT complexity encoder achieves similar subjective quality at 25% reduced bit rate compared to an RD-optimized H.263+ encoder.

JVT-C016 **[Horowitz] AHG Report: Complexity**
Review of complexity analysis and reduction efforts.

JVT-C115 **[Winger, VideoLocus] Reduced Decoder Peak Bus Bandwidth**
The current decoder is reported to be quite demanding in terms of the worst-case bus bandwidth required to be compliant in all cases. This proposal suggests a solution to reduce the system cost of a compliant decoder by reducing the worst-case bus bandwidth. The proposed solution was reported to have virtually no impact on the rate-distortion performance of the JM1.9 codec for the common test conditions, while reducing the worst case total number of block accesses by half.

Nokia Exhibit 2024, p.25
Snap/Hisense v. Nokia
IPR2025-01440

Conduct verification experiment to verify that loss in quality is not visible. Conclude in July meeting.

JVT-C152 **[Horowitz+, Polycom] JVT/H.26L Decoder Complexity Analysis**
Table-driven method to analyze complexity on different platforms. Compare to H.263 baseline.
Complexity of decoder dependence on encoding design. Memory requirements.
Information Only

JVT-C159 **[MPEG] MPEG Memory Complexity Analysis AVC [N4570n]**
Information Only

10. Carriage and File Format

Discussion and planning:
We will have a common representation of all parameter set data, slice data, and SEI's that are carried in the same way in bitstreams and in packets. Each atomic chunk of such data is referred to as a NAL unit. This includes ensuring no emulation of start code prefixes of the bitstream format (i.e., in the current design this is common definition of EBSP's).

We will have a bitstream format in the joint video standard (e.g. in an annex). This format

- will be capable of conveying temporal reference (relative timing data) in the video bitstream for use in HRD
- will be capable of conveying all data in band.
- will include sufficient information to identify conformance of the video bitstream and this requires sufficient relative timing information.

We will have a packet-oriented interface in the joint video spec.

Conversion between the two formats should be simple. In other words, one part of the spec will describe what data needs to be carried, and

- the bitstream definition part will describe how to put all that into a bitstream
- the packet-oriented definition part will describe how to put this data into packets

All participants are strongly suggested to read the above remarks before waking these issues back up again.

If some data is necessary to decode the Y, Cb, Cr sample arrays, goes in parameter set or the NALU, if not then as a general rule, it goes in SEI.

SEI use is generally intended for things that can get lost. Parameter sets for things that cannot get lost. Parameter sets are persistent for sequence unless changed.

Agreed: Bitstream picture start code suffix will include picture type and store/no-store indicator. I picture contains all I slices, P either I slices or P slice, B either I or P or B slices.

SEI is carried as a separate packet (e.g., with a separate start code or in a separate packet).

NALU is Type code incl. error indication (8b), definition of EBSP's to follow.

Compound packets, if done, are external to joint video spec.

Payloads:

- partition(s) of each type of slice (store/no-store indicator in slice header)
- parameter set update(s)
- SEI's

All currently defined parameters nominally reside at the random access point level.

Also a picture type indicator at the picture level.

Nokia Exhibit 2024, p.26
Snap/Hisense v. Nokia
IPR2025-01440

If there is a lot of picture level data, we will do picture-level parameter sets, if not, we will repeat it at the slice level.

RPSL was designed with an intent to have both picture level RPSL and slice level RPSL.

Some concern was expressed as to gateway operation and the support of packet networks that do not have a per-packet timestamp (e.g., H.324/M). These concerns are expected to be the subject of further study.

One way to send a picture

Picture Header Level EBSP

Slice Header Level #1
Slice Content EBSP #1

Slice Header Level #2
Slice Content EBSP #2

Slice Header Level #3
Slice Content EBSP #3

The other way

Picture Header Level EBSP
Slice Header Level #1
Slice Content EBSP #1

Picture Header Level EBSP
Slice Header Level #2
Slice Content EBSP #2

Picture Header Level EBSP
Slice Header Level #3
Slice Content EBSP #3

Picture Header Level EBSP

Picture Number
Picture Storage Information

Slice Header Level EBSP

First MB in Slice
Number of MBs in Slice
Slice QP
Slice Reference Picture Reordering OR do as in first slice

Clarification

No EBSP separation between slice4 header level and slice content

JVT-C122 **[van der Meer, Philips] JVT Transport over RTP**
This contribution addresses requirements for the RTP NAL in particular in relation to the MPEG-2 NAL. An example provides a global description how these requirements could be implemented.

Nokia Exhibit 2024, p.27
Snap/Hisense v. Nokia
IPR2025-01440

JVT-C143 **[Walker+] File Format for JVT Video based on MP4**
Report as input to file format work of MPEG.

JVT-C147 **[Singer; Apple] Towards storing JVT in an MP4 File**
The MP4 file format has been defined for MPEG-4 content, and as it is an extensible file format, it is natural to ask how JVT sequences might be stored in it. Other contributions explore the possibility of adding structures to the file format, which, though generic, are intended to solve the specific issues that arise in JVT. This document attempts to derive a different approach: to what extent can JVT sequences be stored in MP4 without introducing new structural concepts at the file format level?"
Report as input to file format work of MPEG

JVT-C158 **[MPEG] MPEG Prop. Guidelines Carriage AVC [N4714p]**
Discussed with consideration of issues in joint discussion and planned solution as noted.

11. NAL, High-Level Syntax, and Robustness

JVT-C055 **[Suzuki+, Sony] NAL Syntax and Carriage of JVT Stream over MPEG-2 Systems**
A method for carrying JVT video over MPEG-2 Systems was proposed in Geneva in JVT-B049, B063, B070 and B088. It was agreed that the startcode is needed for bit serial communication environment, like MPEG-2 Systems, and adopted in JM. In this document, detail definitions of NAL for MPEG-2 Systems, based on the study in JVT-B70, are proposed.

Chair's remarks on the presentation:

- MPEG-2 TEL proposed. Sequence header, random access header, picture header, and slice header proposed.
- MPEG-2 definition of an access unit is unchanged: An access unit is a picture.
- Mapping to PES
- Modifications to MPEG-2 Systems: New Stream id and stream type
- New descriptors of MPEG-2 Systems for parameter set descriptor and SEI descriptor.

Comments from JVT group:

- New amendment needed for carriage of MPEG-4 Part 10 in MPEG-2 Systems.
- The proposal makes low-delay gateway design very hard, because sequence, random access, picture, and slice header have to be converted to be parameter sets.
- The proposal does not follow the spirit of the current JVT high-level syntax based on parameter sets.

JVT-C063 **[Suh] NAL Compatible to QoS Controlled Network**
The informational contribution discusses the use of the JVT codec in networks where QoS can be controlled. The group did not recognize any immediate needs to respond to the presented problems in the JVT coding standard. The group thanks the author for the information.

JVT-C064 **[Sullivan, Microsoft] Proposed Alteration of Byte Stream Format**
This contribution presents a modification of the prior JVT-B063 method of formatting JVT data into a byte stream with use of start codes in order to address concerns regarding byte alignment loss. The modification was reported to preserve the advantages of the prior method (low data expansion overhead for start code emulation prevention, low start code overhead, byte-oriented processing for encoding and decoding, and avoidance of conflict with MPEG-2 Systems start codes) and adds reportedly-simple byte alignment recovery capability using only byte-oriented processing. The proposed method is reported to be actually simpler than the one from JVT-B063 that is in the current design.

See notes in JVT-C095 section.

Nokia Exhibit 2024, p.28
Snap/Hisense v. Nokia
IPR2025-01440

JVT-C070 **[Hannuksela, Nokia] Signaling of Timestamps**
This contribution proposes signaling of timestamp clock frequency, temporal references, decoding timestamps, and presentation timestamps."

- Question: Is there a limitation on timestamp SEI message frequency? It is up to a specific TEL to decide.
- The use of timestamp SEI messages is up to a specific TEL.
- The group supports sections 3 and 4 of the proposal (signaling of temporal references).
- Presentation timestamps may not have to be carried in coded video stream, because all targeted systems have another mechanism to transmit them (MPEG-2 Systems, RTP) or they are not needed (H.320). However, compound packets in RTP might make an exception.
- Decoding timestamps are needed at least in RTP transport when gatewaying to MPEG-2 Systems.
- Agreed: SEI packets are used to convey timestamps if needed in the system.
- Decision on carriage of decoding and presentation timestamps is deferred to be dealt with HRD.

See notes in carriage section.

JVT-C078 **[Hannuksela, Nokia] Coding of Parameter Sets**
Proposedsyntax and semantics of parameter set update NAL packets (PUPs). Four parameter set structures are used: independent GOP, picture, slice, and presentation parameter set.

The group decided to have only the independent GOP parameter set. Some of the proposed parameters were considered potentially unnecessary.
Parameters proposed for the presentation parameter set were decided to be carried with the Supplemental Enhancement Information.

Parameter set identifier signaling agreed (section 3.3 of the contribution).

Adopted.

JVT-C080 **[Hannuksela, Nokia] Signaling of Enhanced GOP Information**
The enhanced GOP concept organizes pictures in temporal enhancement layers and sub-sequences. The signaling mechanism is proposed in the contribution can be used for improved picture identification and computational scalability in decoders, for example.

Not a requirement and therefore not high priority at this meeting.

JVT-C087 **[Yoshinori Matsui, Matsushita] Common NAL Packet Structure**
It is proposed that the syntax of the network layer independent NAL packet should be able to use commonly among transport layers including file format such as MPEG-2 TS/PS, RTP and MP4.

The contribution proposes a flexible NAL packet syntax that can contain anything from a partial slice to multiple slices.

Comments from the group:

- The proposal would add 24 bits overhead to a NALP.
- The property of having independently decodable NALPs would be lost (if a NALP carries a fragment of a slice).

The proposal makes the operation of gateways harder

JVT-C095 **[Lindbergh] H.320 NAL for JVT**
The contribution proposes start codes and start code emulations based on JVT-C064. The group supports the proposal. Concerns were expressed on the uneven probability to have long zero-runs in the exp-

Nokia Exhibit 2024, p.29
Snap/Hisense v. Nokia
IPR2025-01440

Golomb code and the proportion of start code emulation bytes due to this fact. The proponent promised to bring hard data on the proportion of start code emulation bytes by the next meeting.

Further investigation is encouraged whether it is possible to have a clear way to distinguish within a bitstream format between the H.26Xs. Adopted with N=1 pending on results rolling in at the next meeting.

JVT-C131 **[Stockhammer, TUM] NAL Packet Segmentation**
A new NAL packet type and an additional NAL header flag are proposed to allow segmentation of long NAL packets in shorter packets. Syntax and Semantics are included in this contribution.
Withdrawn.

JVT-C144 **[Walker+] Generic Adaptation Layer for JVT Video**
The contribution proposes division of what used to be called the "NAL" to the network-independent "Generic Adaptation Layer" (GAL) and to the network-dependent NAL.

These issues are dealt with notes in carriage.

JVT-C018 **[Sullivan+] AHG Report: Film Mode / Timing**

Discussion of timing related issues.
Note: JVT-C046 contains film mode

JVT-C164 **[Hannuksela] Prop. NAL description**
Pict and Slice header parts adopted.

JVT-C030 **[Kadono+, Matsushita] Additional MMCO command for supporting more flexible bitstream switching**
The prediction structure of S-pictures reported to need to be strictly limited to guarantee the uniqueness after bit stream switching. Proposed to introduce a new MMCO command to allow a more flexible picture prediction structure.

When an S picture is received, the multiframe buffer must be reset. S frames must be treated like I pictures.

JVT-C054 **[Suzuki+, Sony] Study of Random Access for JVT**
In this contribution, further study on random access for JVT codec is reported. Splicing and bitstream editing for JVT codec is studied. Advocated that JVT codec should support splicing and bitstream editing at video layer as it is supported by MPEG-2. The syntax and semantics extensions are proposed to support splicing.

NALP types. May consider adopting NALP types as part of JVT-C074 as simplified and revised in conjunction with FMO.

Issues discussed also in C074 remarks.

Discontinuity flag conflicts with C117.

JVT-C072 **[Wang, TUT] Isolated Regions: Motivation, Problems, and Solutions**
A sub-area of the picture is defined called isolated region. Over time the isolated region can grow or shrink. The loop-filter is turned off at the boundary of the isolated region. Various limitations are applied to motion vectors.

Nokia Exhibit 2024, p.30
Snap/Hisense v. Nokia
IPR2025-01440

This contribution proposes syntax and semantics for a concept of isolated regions, a technique that provides a form of gradual random access, error resiliency/recovery, picture in picture functionality, and coding of masked video scene transitions.

Software availability indicated.

More discussion needed. Group considers this to be one possible solution.

JVT-C072 to JVT-C075 treated as a group:

Isolated region defines a boundary around which the loop filter is turned off.

Some ways of supporting this functionality:

1. Combines well with FMO concept.
2. Can also do by turning off loop filter at slice boundaries as in C117.
3. In this contribution, propose to define a set of evolution patterns in the parameter set and put position within the pattern into slice header.

Support JVT-B063 gradual decoder refresh design.

Comparable error resilience to RDO method without its complexity, and better performance if the two techniques are combined.

Note re JVT-C117 idea of slices with indication to turn off filter at boundaries of slice or intra/inter boundaries

Further study recommended.

JVT-C073 **[Wang, TUT] Error-Robust Video Coding Using Isolated Regions**
This contribution proposes an error-robust video coding technique using isolated regions. While improving the error resiliency performance of the JVT codec, it also provides a possibility for perfect gradual random access.

Experimental results report that methods involving isolated regions provide significant gains in terms of decoder PSNR (mobile test conditions). Subjective results were requested.
Treated with C072.

JVT-C074 **[Wang, TUT] Gradual Decoder Refresh Using Isolated Regions**
This contribution proposes a gradual decoder refresh technique using isolated regions. It is reported that it is advantageous to use isolated regions instead of intra pictures to provide random access points in error-prone environments. The corresponding NALP type must be transmitted. Main reason for this method to work is the switched off loop filter.
Treated with C072

No consensus to adopt at this time. Encouraging results and can be revisited at the next meeting.

JVT-C075 **[Wang, TUT] Coding of Masked Scene Transitions Using Isolated Regions**
This contribution proposes to code masked scene transitions using isolated regions. A large number of different mask patterns can reportedly be realized conveniently.
Treated with C072.

JVT-C079 **[Hannuksela, Nokia] Improved Coding of Slice Headers**
Bit-rate savings are reportedly achieved in slice headers by fixed-length coding of parameters, granular slice position and size, and differential coding of QP relative to a default value given in the parameter set.

Nokia Exhibit 2024, p.31
Snap/Hisense v. Nokia
IPR2025-01440

No consensus to adopt.
Treat QP issue with C117.

JVT-C081 **[Hannuksela, Nokia] Sync Pictures**
Coded pictures representing the same picture contents are referred to as sync pictures. The paper proposes slice header syntax and semantics to signal the presence of sync pictures.

Members encouraged to consider for further study.

JVT-C083 **[Hannuksela, Nokia] Signaling of "Clean" Random Access Points**
A 'clean' random access point is such that no data prior to the point is referred to in the decoding process. The contribution proposes identification of random access positions using a specific NAL packet type and a specific picture number update rule in the random access positions."

NALP type and multipicture reset and picture number = 0 proposed and adopt.

JVT-C129 **[Halbach, NTNU] Reduced slice headers and bit error resilience**
"Subslices" are proposed for efficient treatment of residual errors in the code stream after transmission. Employment of these structures and their small-size headers reportedly leads to both bit rate savings and a significant increase of robustness in the decoder. Along with subslices, data partitioning, resynchronization markers, a variable-length non-resynchronizing code as well as a simple error concealment technique are utilized. Subslices, DP and RMs aim at the normative, the variable-length code and EC at the non-normative section."

Bit-error channel resilience. Overhead noted for use of many slices. Divide a slice into "sub-slices" with unique resync mark and use non-resynchronizing VLC. Recommend not using resynchronizing VLC (resync not really used for anything in current design). Distinction between slice and sub-slice is that header data is not repeated. By not using full slice header save header bits.

Response: Some aspects out of date relative to current design. We note that different systems will not carry the data in the form described in the document "as-is." We believe the slice structure design will address the concerns in this contribution. The error conditions and the design assumed in the doc for the way the current design works may not be relevant.

JVT-C141 **[Wenger, Teles] ERPS Use for Commercials in B'cast/Streaming**
Informative.

JVT-C142 **[Walker] Proposal for JVT SEI Messages**
SEI capability is already in the design that can carry all possible data including metadata. Metadata support should therefore be feasible.

JVT-C017 **[Horowitz] AHG Report: Robustness**
Review of input contributions. Noting of scattered slices core experiment work.

JVT-C032r1 **[Zhou, Tsinghua Uni] Experimental Result of Fragile Watermark Based Error Detection Scheme**
A set of error detection schemes using fragile watermark for hybrid codec based video communication were proposed. It reportedly improves the error detection rate and error correct detection rate dramatically. To take the advantages of the watermark, the standardization of the schemes is proposed to be considered.

In mobile environments very low bit rates are used. Therefore header and MV data dominate and not many coefficients are used. No detection on those syntax elements with possible degradations in performance.

Nokia Exhibit 2024, p.32
Snap/Hisense v. Nokia
IPR2025-01440

Following mobile environments test conditions is needed.

Comparison to FEC is missing.

Intra refresh.

Network interaction and particular application should be considered.

JVT-C048 **[Borgwardt, VideoTele.com] Rectangular Slices to Tradeoff Error Resiliency and Coding Efficiency**
Rectangular slices are proposed as more efficient for prediction between macroblocks than the current slice rows. They would reportedly allow an encoder to use smaller slices for error resiliency while still having good coding efficiency.

Bit errors in MPEG-2 are more common than one would expect. Can this be combined with FMO? Appears to be. Have a general syntax on slices.

See C089. FMO is more general. No action.

JVT-C071 **[Wang, TUT] Core-Experiment Results of Scattered Slices**
This contribution provides the core-experiment results of scattered slices, proposed in JVT-B027, as a response to the core experiment description given in JVT-B111r1.

JVT-C046 **slice type part (other aspects in another category)**

JVT-C082 **[Hannuksela, Nokia] Spare Reference Pictures**
Proposed is signaling of entire and partial spare reference pictures with the Supplemental Enhancement Information mechanism. With the help of the proposed signaling, receivers may reportedly avoid unnecessary picture freezing, feedback, and error concealment, which are normally done as a response to missing picture data.

Not a high priority to act on it this time. Consider within AHG.

JVT-C089 **[Wenger, Teles] FMO: Flexible Macroblock Ordering**
This document proposes the introduction of a new tool into the JVT baseline, which allows the coding of macroblocks in an order other than the typical raster scan order. The key application of the proponents for the tool is the implementation of error resilience mechanisms such as Scattered Slices and Slice interleaving, as documented in JVT-C090 and JVT-C091, but due to its flexibility, other applications are reportedly possible.

Contributions 89-93 treated as a group.

Concept previously known as “scattered slices”. Send a map of indices such that a map of assignment of each macroblock to a slice group ID. A slice contains a number of MB in raster-scan order within one slice group. Send map at parameter set level.

Low complexity impact reported.

Cross-verified (by Nokia) in JM1.7 code.

Proposed primarily as error reslience tool.

Common conditions tests for IP/RTP (using loss patterns from internet) – reported consistent improvements (more for some sequences). Not entirely clear how much gain.

Nokia Exhibit 2024, p.33
Snap/Hisense v. Nokia
IPR2025-01440

Further improvement capability with better encoder optimization asserted.

Visual gain shown previously in Geneva.

FMO is adopted into baseline and Main profile with limit of maximum of 8 slice groups and subject to further complexity analysis and verification experiments to be completed by JULY meeting.

JVT-C090 **[Wenger, Teles] Scattered Slices: Simulation Results**
This document reports findings when applying the error resilience tool 'Scattered Slices,' which is based on the Flexible Macroblock Ordering tool as proposed in JVT-C089, to the test conditions defined in the current 'Common Conditions for Wireline IP/UDP/RTP Conversational environment'.
See comments for C089

JVT-C091 **[Wenger, Teles] Slice Interleaving: Simulation Results**
This document discusses the different properties of implementing Slice Interleaving on the VCL and the NAL layer.
See comments for C089.

JVT-C092 **[Wenger, Teles] JM1.7 software with implemented FMO**
See comments for C089.

JVT-C093 **[Wenger, Teles] FMO changes proposed to Working Draft content**
See comments for C089.

JVT-C096 **[Kim, Samsung] Sub-picture coding with the reduced boundary effect**
Proposal of an error robust sub-picture coding method with reduced boundary effect. A frame is divided into the foreground and background sub-pictures, and the foreground is coded in a higher picture quality and transmitted with a better error protection than the background."

Relates to foreground/background region proposal of Geneva.
Most of it can be done with FMO syntax.
Incorporate use within FMO as non-normative issue.

JVT-C117 **[Sullivan, Microsoft] Seven Steps Toward a More Robust Codec Design**
This contribution proposes seven ways of reportedly making the current JVT video codec design more robust. It is proposed to:
1) avoiding worst-case data expansion by adding an intra PCM mode,
Conclusion: discuss in July
2) reducing needless design variation by limiting the number of ways of partitioning the VCL data,
Conclusion: only 2 ways are in the current design: 1 packet per slice or 3 packets per slice. No action needed.
3) achieving basic decodability of intra regions by allowing intra prediction from inter-coded regions to be turned off within slices,
Conclusion: already in draft. No action needed.
4) achieving greater error resilience and random access recovery by allowing deblocking filtering to be turned off at intra/inter boundaries within slices,
Conclusion: further study in July
5) achieving greater error resilience and random access recovery by allowing deblocking filtering to be turned off at slice boundaries,
Conclusion: further study in July
6) adopting "gradual decoder refresh" functionality as previously proposed,
Conclusion: further study in July
7) avoiding a needless coupling of codec technology to assumed input characteristics by letting encoder choice rather than the step size selection determine whether double-scan of coefficients is used.

Nokia Exhibit 2024, p.34
Snap/Hisense v. Nokia
IPR2025-01440

Conclusion: void
double-scan aspect discussed in transform, quant section and recommended for adoption there. Double-scan then removed from design – obviating the need for this proposal to be considered.

Study loop filter switchability at slice/slice group boundaries.

JVT-C125 **[Chen+] Flexible data partitioning for improved robustness performance.**
This proposal describes modifications to the VCL data partitioning syntax and the corresponding NAL packetization process to enable flexible data partitioning, which is advocated as essential to supporting video communication applications over diverse **packet-lossy** networks. The proposed changes are reported to be backward compatible with the current specifications. A partitioning framework that reportedly eliminates decoding complexity overhead and reduces bit rate overhead is also proposed.

Two partitions are proposed.

Good receiption from the group. More results with common conditions are needed. Refer to C117 for restriction on DP flexibility.

JVT-C132 **[Stockhammer, TUM] Independent Data Partitions A and B**
Proposed to make partitions A and B independently decodable by duplicating the slice header and the MB type information in partition B. Mandatory decoder actions are specified which allow encoder trade-off for simplicity and compression efficiency.

No action taken in response.

12. SP/SI frames

JVT-C114 **[Sun, Microsoft] The improved JVT-B097 SP coding scheme**
An altered SP coding scheme is proposed by combining the JVT SP scheme and the JVT-B097 scheme. The reconstructed reference can reportedly be improved 3.64% in average, and the display image can be improved 10.34% in average. Experiment results and software were provided.

Adopted.

JVT-C126 **[Sun; Microsoft] Core experiment results on SP**
The statistic results of the core experiment JVT-B112 are reported in this document. An additional excel file contains all detailed results.

13. Entropy Coding

13.1. VLC Approaches

JVT-C008 **[Bjontegaard] AHG Report: VLC.**
As a result of the AHG work, two documents on adaptive VLC coefficient coding were submitted to the JVT meeting.

JVT-C028 **[Bjontegaard, RealNetworks] Context-adaptive VLC (CVLC) coding of** coefficients
The CVLC is a proposed replacement of the UVLC coding of transform coefficients, and uses context adaptive coding to reportedly achieve a significant bit rate savings over UVLC.

<u>JVT Group Conclusion:</u>

Little better than C088

Nokia Exhibit 2024, p.35
Snap/Hisense v. Nokia
IPR2025-01440

Some of the concerns are

- Most of the gain is at the higher than normal operating range of the bit rates for JVT's current profiles
- 18 bit code length. Can the code length be limited to 16 bits? (nice to have)
- Cross verification (Intel, Real)
- Technically both this and C088 are very close (difficult to be distinguished with any significant differences)
- IPR statements favour C028
- Therefore, based on technical and IPR situation, group recommends to use C028 over C088.
- Request proponents and the group to work on reducing the number of tables, code length, look into possible problem of start code emulation and using structured code

Adopted.

JVT-C086 **[Cheung, TI] Improving coefficients coding for Intra Picture**
This contribution proposes an improvement on Nokia's VLC proposal JVT-B072-revised, providing additional gain in coding efficiency on top of their results.

Provides about 1% improvement over C088 for Intra only.

JVT-C088 **[Karczewicz, Nokia] Adaptive coding of transform coefficients**
This proposal addresses the VLC coding of transform coefficients. The proposed method reportedly provides significant improvements, up to 23% for inter and 30% for intra, compared to the current method of coding of transform coefficients. Software is also provided.

JVT Group Conclusion:

Some of the concerns are

- most of the gain is at the higher than normal operating range of the bit rates of JVT's current profiles

Cross verified by Matsushita (JVT-C145)
See comments for C028

JVT-C145 **[Kondo+, Matsushita] Verification results for adaptive coding of transform coefficients (JVT-C088)**
The verification results for JVT-C088 is reported. It is implemented to the JM1.7 software independently. Reportedly verified that it can reduce large amount of bits compared to UVLC coding."

[Information Only]

13.2. CABAC

JVT-C009 **[Marpe] CABAC Ad Hoc Group Report**
Status of CABAC work noted.

JVT-C029 **[van der Vleuten, Philips] Low-Complexity Arithmetic Coding for CABAC**
Proposed is an implementation of arithmetic coding that is reported to significantly reduce the arithmetic coding implementation complexity, while it reportedly has a negligible influence on the compression efficiency. The proposal is reported to satisfy all the requirements that were set by the CABAC-AHG.

JVT Group Discussion:

Nokia Exhibit 2024, p.36
Snap/Hisense v. Nokia
IPR2025-01440

Not cross-verified completely. Only coding efficiency was cross-verified but complexity was not. For some architecture it is expected to be little better than C061 and for others little worse. Group agrees that technically there is not clear winner between C029 and C061.
Do not have specific measurement or numbers.

Proposal is to accept C061 No opposition except from proponent. Encourage proponents to complete the cross verification. C061 leaves the door open for adoption of CABAC in another profile that can possibly be royalty free.

JVT-C031 **[Kondo+, Matsushita Electric Industrial] Improvements on RUN and LEVEL information coding in CABAC**
Altered methods on (RUN, LEVEL) information coding in CABAC are proposed. The proposed methods are inverse scan of transform coefficients and new transition method of contexts in the LEVEL coding.

JVT Group Discussion:

Gains are smaller than those provided by C060.
Not Accepted
Encourage proponents of C031 and C060 to do further work in the direction of merging some of the ideas while still keeping the 2.2.1 status of C060.

JVT-C038 **[Bossen, NTT DoCoMo] Bounding the complexity of arithmetic decoding**
This document analyzes the worst case complexity of arithmetic decoding and concludes that the worst case complexity can be 40-50 times higher than the average complexity. A modification of the arithmetic coding core is proposed to reduce the worst case complexity.

JVT Group Discussion:
Accepted subject to cross verification, and CABAC not being in Baseline.

JVT-C060 **[Detlev Marpe, HHI] Improved CABAC**
Proposed are two modifications to CABAC, which reportedly further improve coding efficiency in comparison to UVLC. The first modification concerns the coding of transform coefficients. The experimental results report that this modification increases the bit-rate savings of CABAC in comparison to UVLC by additional 1-3 %. The second modification is a change of the macroblock type encoding in B-frames. It introduces the SKIPPED macroblock into the set of B-frame macroblock modes for CABAC as it has been done for the UVLC. This modification unifies the process of macroblock type encoding for both schemes. Substantial coding gains for B-frames are reported for sequences with low motion.

Accepted subject to better text description and software availability.

JVT-C061 **[Detlev Marpe, HHI] Fast Arithmetic Coding for CABAC**
In this document was proposed a new arithmetic coding core engine to speed up the arithmetic coding and decoding processes. Compared to the arithmetic coder in WD2 the proposed arithmetic coder reportedly leads to only minor increases of the bit rate, on average ca. 0,5%, while a speed-up of ca. 24 % of the CABAC decoder (for bit parsing, context modeling and symbol decoding) and of ca. 5 % of the total decoding time can reportedly be achieved.

Accepted. See comments in C029.

JVT-C062 **[Detlev Marpe, HHI] Verification Results for JVT-C029**
See comments re JVT-C029

JVT-C100 **[Haskell] Improved Terminology for CABAC Section of WD**
Information only

Nokia Exhibit 2024, p.37
Snap/Hisense v. Nokia
IPR2025-01440

Text editors are encouraged to take this contribution into account during the editing process.

JVT-C133 **[Stockhammer, TUM] Verification of Improved CABAC**
Verification work.

JVT-C134 **[Stockhammer, TUM] Verification of Fast Arithmetic Coding Engine**
Verification work.

JVT-C162 **Putting a reasonable upper limit on Binarization Expansion**

JVT Group Discussion:

Accepted.

JVT-C098 **New Syntax for Error Resilient Transmission of CABAC Stream**

To be studied further in the adhoc group on CABAC

14. Profiles and Levels

JVT-C019 **[Lindbergh] AHG Report: Profiles, Levels & Applications**
Description of remarks and proposals relating to profiles, levels and applications.

JVT-C026 **[Lindbergh, Polycom] JVT Profiles Proposal**
Proposes specifics for the JVT Baseline and other Profiles

JVT-C036 **[Le Maguet+, Philips] TV profile and levels**
This paper aims at defining a TV profile both for standard and High definition. It also defines two levels for the TV profile, one for standard definition and one for high definition and proposes a list of parameters to be part of the level definition.

JVT-C058 **[McVeigh, Intel] Straw-man proposal for baseline profile**
This document is in response to the request issued in document JVT-B108 for proposals on profile tool set definition. A straw-man proposal for the Baseline Profile is presented, justifications for the recommended tool set were discussed, and open issues were identified to be resolved before final profile definition.

JVT-C068 **[Hannuksela, Nokia] Definition of Profiles**
Two profiles were proposed for the JVT codec and the allocation of the accepted coding tools between these profiles.

JVT-C099 **[Chen, Broadcom] Main Profile for MPEG-4 AVC and H.264**
This is a proposal to define the Main Profile of MPEG-4 AVC (H.264) for applications requiring entertainment quality video, such as digital TV, HD DVD, high-quality streaming video, video-on-demand and personal video recording, etc.. The proposal is based on reported requirements for enhanced TV services. Tools selected for this profile were reported to be evaluated by considering coding performance vs. cost trade-off and also with supporting the RFB principle.

JVT-C109 **[Bäse, Siemens] Profile/Level restrictions for very low end applications**
This document is targeted on the "very low end" side of applications. Therefore only the baseline profile is taken into account. Most of the intended applications so far are granting a certain level of available complexity which is not the case on the very(!) low end side. Here some additional restrictions are necessary while supporting existing applications still.

Nokia Exhibit 2024, p.38
Snap/Hisense v. Nokia
IPR2025-01440

JVT-C113 **[Luthra, Motorola] Baseline Profile for JVT Coding Standard**
A proposal was made to define the Baseline Profile in JVT with the aim of unifying the video conferencing, digital TV (including broadcasting of entertainment video, video-on-demand and personal video recording types of applications) and streaming video worlds under the same profile. In the past, MPEG has focused on defining different profiles for these worlds. However, with the advancement in technology and cost effectiveness of new technology, it is proposed that there is no need to define separate profiles for the applications in those classes.

JVT-C155 **[MPEG] MPEG Proc. to Develop Profiles & Levels [N4671p]**
Description of MPEG process for development of profiles and levels.

JVT-C161 **[Boyce+, Thomson] Mobile Profile Proposal**
Proposal of profile targeted toward low-power mobile devices.

JVT Group Discussion:

Some opinions and remarks expressed:

- Issue: Limit number of levels as well as profiles?
- Concept of levels with sublevels introduces as a way of compromising between minimum number of major configuration points and desire for flexibility and scalability of implementation.
- Leave some TBD's?
- Does current draft allow MV's way off edge of picture?
- If FLC for QUANT in slice header, then set the number of bits at parameter set level?
- Baseline: no 1/8 pel, Yes deblocking, no mixing intra/inter within MB, 4:2:0 only, Interlace at level capable of HHR+, maximum of 15 reference pictures, no CABAC, VLC, No SP, all motion prediction block shapes, all intra prediction modes, no data partitioning, all QP values, B pictures? No.
- Max MV Range: 256
- # Ref pictures at highest resolution: ?
- Main: Same as baseline, plus CABAC, ABT, B pictures, bi-predictive coding of blocks smaller than 8x8 is not allowed (revisit after JVT-C115).
- ABT to be discussed.
- Will have maximum bit rate.
- May want limits on bit rate and number of bits per macroblock.
- The level numbering scheme will have sub-levels using numbers.
- Limit on extreme aspect ratios as defined in Geneva are tentatively accepted and will be the subject of further study.

Conclusions reached are reported in Summary Results section and Profile and Level Conclusions are recorded in C165 (JVT output document).

Adoption of ABT is subject to verification experiments and further complexity analysis to be completed by July meeting.

FMO is accepted in baseline and Main profile with limit of maximum of 8 slice groups and subject to further complexity analysis and verification experiments to be completed by JULY meeting.

15. Test Model [GS]

JVT-C034 **[Chen+] Paraboloid Prediction Motion Comp Search**
Motion estimation search method presented.

Nokia Exhibit 2024, p.39
Snap/Hisense v. Nokia
IPR2025-01440

JVT-C065 **[Hong, ???] Further Improvements on Motion Search Range Decision**
This Contribution is about the optimal decision of search range for fast encoder implementation using neighboring motion vectors. It improved the performance of JVT-B022 which was already presented in the Geneva meeting.

JVT-C084 **[Koichi Takagi, KDDI] Lagrange Multiplier and RD-characteristics**
This contribution verifies Lagrange multiplier used in current JM and proposes better value than the conventional one.

JVT-C101 **[Kim+] High-Compl. Mode Dec. for Error Prone Channel**
This contribution proposed an improvement of the high complexity mode decision for error-prone channel presented in VCEG-N43. It was stated that the drift noise, which is defined as the encoder-decoder mismatch of reference frame, is the major factor that causes degradation of video quality when transmitting compressed video bitstream through error prone channel. Though intra-refreshment is known to be an effective method of reducing the drift noise, it generally requires sacrifice of compression efficiency. This contribution proposed an R-D optimization method considering estimated drift noise to perform optimized intra-refreshment.

JVT-C146 **[Wenger] Prop. Common Cond. for Wireline IP/UDP/RTP**
Not available.

JVT Group Discussion:

Due to a high volume of high-priority work toward definition of the normative standard draft, the group agreed to primarily defer consideration of test model and other such non-normative issues for further study in an ad hoc group and at future meetings.

16. Primary Meeting Results

16.1. Communication Practices, Ad Hoc Groups Established, and AHG Charters

The general JVT reflector can be subscribed to by clicking on “join jvt-experts” at http://mail.imtc.org/cgi-bin/lyris.pl?enter=jvt-experts

Email for the reflector should be sent to jvt-experts@mail.imtc.org.

The subject line of each email message will automatically be prefixed with "[jvt-experts]". Project identification and unsubscribe information will be attached to the bottom of each reflected message.

Our new ftp site recently established for JVT use is ftp://ftp.imtc-files.org/jvt-experts/.
Our prior site containing the files of the first and second JVT meetings was the VCEG site at http://standard.pictel.com/ftp/video-site.
Our new work will be put on the new site, and we will be bringing our prior files from the old site to the new one shortly.

We heartily thank the International Multimedia Telecommunications Consortium for their great generosity in hosting both our email reflector and our ftp site.

Nokia Exhibit 2024, p.40
Snap/Hisense v. Nokia
IPR2025-01440

Many good technical discussions have been taking place on the JVT reflector and the reflectors of the various JVT ad hoc groups.

Below is the current list of interim JVT ad hoc groups established at the Fairfax meeting that report to the Klagenfurt meeting, their chair contact information, and ad hoc reflector lists.

To AHG Chairs:

1) Please check this to make sure it contains correct and complete information for your ad hoc group. (Particularly if the Discussion entry says "contact the chairs/general...")
2) Please always include the three rapporteurs on any ad hoc group discussions.
3) When possible, put an identifying tag in the subject line for discussions of a particular ad hoc group. E.g., "[jvt-interpolation]"

We need to make sure that our process is open for participation by all interested parties. We also hope that anyone working in an area covered by a designated ad hoc group will try to coordinate their work with the activities taking place in that group to the maximal possible extent. Care should especially be taken in areas that affect multiple ad hoc groups.

16.1.1. JVT Project Management

Chairmen: Gary Sullivan (garysull@microsoft.com), Ajay Luthra (aluthra@motorola.com), and Thomas Wiegand (wiegand@hhi.de)

Charter: To further the work on the JVT project as a whole, including project planning, work coordination, and status review.

Discussions: General JVT Reflector

16.1.2. Text Editing and Reference Software Development

Chairman: Thomas Wiegand (wiegand@hhi.de)

Charter: To further the work on the draft text and software implementation of the joint design, including incorporation of modifications as approved by the group, production of the joint committee draft text, collection of comments on the text and software, preparations to facilitate necessary future text modification work, and provision of improved software for group use in future experiments and for eventual approval as standardized reference software.

Discussions: Contact the chairs / General JVT Reflector

16.1.3. Motion Compensation Interpolation

Chairman: Thomas Wedi (wedi@tnt.uni-hannover.de)

Charter: To study the design of the motion compensation interpolation processing in the JVT design, including consideration of the completeness and correctness of the draft specification, the potential for use of adaptive motion interpolation and consideration of the rate-distortion-complexity tradeoffs in motion interpolation design.

Discussions: General JVT Reflector, tagline "[jvt-interpolation]"

16.1.4. Study of Complexity

Chairmen: Michael Horowitz (MHorowitz@Austin.Polycom.com) and Jan Bormans (bormans@imec.be)

Nokia Exhibit 2024, p.41
Snap/Hisense v. Nokia
IPR2025-01440

Charter:

- To study the implementation complexity of the AVC/JVT Codec
- To formulate strategies to minimize the computational and implementation AVC/JVT codec complexity
- To provide guidelines for the complexity studies of the AVC/JVT Codec
- To collect meaningful AVC/JVT Codec parameter settings (to be provided by the JVT)

Discussions: jvt-complexity-list@imec.be
Subscription: Email to owner-jvt-complexity-list@imec.be

16.1.5. Deblocking Filter Analysis

Chairman: Peter List (Peter.List@telekom.de)

Charter: To study the completeness and correctness of deblocking filter specification; to investigate quality and complexity issues for loop filter design in the JVT codec and to assess the potential for improved visual quality, reduced decoder computational complexity, and enhanced design simplicity. It particular this ad hoc group should consider the relationship of the deblocking filter design to interlaced-scan video and the potential usefulness of turning off loop filtering around the boundaries of slices.

Discussions: jvt-loopfilter@berkom.de
Subscription: Contact the chair

16.1.6. High-Level Syntax

Chairmen: Miska Hannuksela (miska.hannuksela@nokia.com), Young-Kwon Lim (young@netntv.co.kr), Thomas Stockhammer (stockhammer@ei.tum.de)

Charter: To Charter: To study the high-level syntax of JVT content (particularly syntax at the slice header level, picture header level, and parameter set level) to determine whether that this syntax is fully and properly specified and supports the carriage of JVT bitstreams for all appropriate transport environments.

Discussions: jvt-car@advent.ee.columbia.edu
Subscription: email majordomo@advent.ee.columbia.edu
saying "subscribe jvt-car"

16.1.7. Timing and Timing Information

Chairmen: Gary Sullivan (garys@microsoft.com) and Sherman Chen (schen@broadcom.com)

Charter: To study issues of timing (e.g. capture, presentation, etc) and Film mode video in relation to the completeness and correctness of the JVT design.

Discussions: Contact the chairs / General JVT reflector

16.1.8. CABAC

Chairman: Detlev Marpe (marpe@hhi.de)

Charter: Charter: To study the CABAC design with regards to completeness and correctness of specification, rate-distortion performance, and complexity.

Nokia Exhibit 2024, p.42
Snap/Hisense v. Nokia
IPR2025-01440

Discussions: jvt-CABAC@hhi.de
Subscription: contact the chair

16.1.9. Joint Model Reference Encoding Method Development

Chairman: Chul-Woo Kim (charlie@mcubeworks.com)

Charter: To further the work on description of example reference encoding methods for the JVT codec design toward eventual adoption as non-normative standard text.

Discussions: Contact the chair / General JVT reflector

16.1.10. Adaptive Block Transforms

Chairman: Mathias Wien (wien@ient.rwth-aachen.de)

Charter: To study the completeness and correctness of the specification of adaptive block transform use in the JVT draft specification; focusing in particular on the degree of harmonization of the ABT design with other changes adopted into the draft.

Discussions: intl-abt-adhoc@ient.rwth-aachen.de
Subscription: email majordomo@ient.rwth-aachen.de
saying "subscribe intl-abt-adhoc"

16.1.11. Intra Coding

Chairperson: Marta Karczewicz (Marta.Karczewicz@nokia.com)

Charter: Consider of possible need to add new prediction modes for chroma prediction and investigate potential for reducing intra coding memory requirements.

Discussions: jvt_intra_adhoc@yahoogroups.com
Subscription: email jvt_intra_adhoc-subscribe@yahoogroups.com

16.2. Summary Information Report and Action Requests from the JVT

The Joint Video Team (JVT) organization reports the following action requests and information to its ISO/IEC MPEG and ITU-T VCEG parent bodies and to its participants:

16.2.1. The JVT forwards the following document to its parent bodies for approval

Title	No.	TBP	Available
Text of Committee Draft of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)	**JVT-C167 (WG11 N4810)**	**Y**	**02/05/12**

16.2.2. The JVT reports the establishment of the following Ad Hoc groups:

Title	Chair	No.	Mtg
JVT Project Management	**Gary Sullivan, Ajay Luthra, and Thomas Wiegand**		**Y**
Text Editing and Reference Software Development	**Thomas Wiegand**		**02/05/11-12**

Nokia Exhibit 2024, p.43
Snap/Hisense v. Nokia
IPR2025-01440

Motion Compensation Interpolation	**Thomas Wedi**		
Study of Complexity	**Michael Horowitz and Jan Bormans**		
Deblocking Filter Analysis	**Peter List**		
High-Level Syntax	**Miska Hannuksela,** Young-Kwon Lim, Thomas Stockhammer		
Timing and Timing Information	**Gary Sullivan and Sherman Chen**		
CABAC	**Detlev Marpe**		
Joint Model Reference Encoding Method Development	**Chul Woo Kim**		
Adaptive Block Transforms	**Mathias Wien**		
Intra Coding	**Marta Karczewicz**		

16.2.3. The JVT future meeting plans proposed by the JVT Chair|Rapporteur are as previously established:

Approx Date	**Auspices**	**Location**	**Project Milestone**
July 22-26, 2002	JTC1	Klagenfurt, AT	JM4, FCD
Oct. 14-18, 2002	ITU-T	Geneva, CH	JM5, AAP Consent
Dec. 9-13, 2002	JTC1	TBD	JM6, FDIS & Rec.
Feb./March, 2003	<Ballot Result>	N/A	IS

It is noted that preparatory ad-hoc meetings within a week of the next meeting in Klagenfurt may be necessary to complete the workplan. Such meetings are to be announced with 30 days notice and approved by SG16 and WG11 management.

16.2.4. The JVT adopted a text of its joint draft for technically-aligned standardization approval by the parent bodies, including in particular approval in ISO/IEC as a Committee Draft [JVT-C167 / N4810]. This includes the adoption of the following changes relative to the editor's draft text JVT-C039 ("JWD2r8"):

Removal of elements not specifying normative video-layer functionality	Removal of draft RTP payload specification (in Annex A of JWD2r8), reference example encoder (in Annex B), interim file format (Annex C), and reference example error concealment description (Annex D)	bug-fix/clean-up
Clarification from discussion	Change "bi-directional" picture → "bi-predictive" picture	bug-fix/clean-up
Conditional JVT-C056 subject to cross-verification and stress testing	Dithered 5-tap Filter for In loop Deblocking	quality enhancement
Conditional JVT-C094 subject to cross-verification	Complexity Reduction and Analysis for Deblocking Filter To adopt modification 1 (removal of recursiveness in "default" filter, and modification 3 (simplified switching for Bs=4, which is I-macroblock boundary). Not modification 2 at this time.	complexity reduction
JVT-C027	Skip Mode Motion Compensation	bug-fix/clean-up
JVT-C028	Context-adaptive VLC (CVLC) coding of coefficients	quality improvement
Conditional	Bounding the complexity of arithmetic decoding	complexity reduction

Nokia Exhibit 2024, p.44
Snap/Hisense v. Nokia
IPR2025-01440

JVT-C038 subject to cross-verification		
JVT-C060 clearer text description for CD will be provided	Improved CABAC	quality improvement
JVT-C061	Fast Arithmetic Coding for CABAC	complexity reduction
JVT-C064	Alteration of byte stream format Adopted with N=1.	definition of system interface
JVT-C066, JVT-C067, JVT-C103	Multiframe Interpolative Prediction Related docs: JVT-C047, JVT-C066, JVT-C067, JVT-C077, JVT-C103 Adopt as in JVT-C067 and JVT-C103 with revised syntax considering JVT-C077 (non-baseline)	very significant quality improvement
JVT-C069	Levels and HRD	clean-up
JVT-C078	Coding of Parameter Sets	definition of system interface
JVT-C083	Signaling of "Clean" Random Access Points	clean-up
Conditional JVT-C089 with limit of maximum of 8 slice groups and subject to further complexity analysis and verification	FMO: Flexible Macroblock Ordering	improved error resilience
JVT-C095	H.320 NAL for JVT Non-system-specific start code structure elements only	Definition of system interface
JVT-C107 and JVT-C140	Adt'l Transforms and Quant. Adoption into non-baseline profile.	New feature
JVT-C114	The improved JVT-B097 SP coding scheme	very significant improvement
Response JVT-C116	Color Space Add equivalent support to MPEG-2 & 4 Video If information not present: undefined or defined by system. “Luma/Chroma” terminology agreed. Sequence level.	bug-fix/clean-up
Response to JVT-C120	Change of MV calculation for direct bi-predictive motion. Send N1, N2 and D at picture level using UVLC. MVf = N1*MV/D MVb = N2*MV/D	bug-fix
JVT-C127	MV Prediction in B Pictures	bug-fix/clean-up
JVT-C135	Pixel Aspect Ratios Add equivalent support to H.263+/MPEG-4 If indication is zero, undefined or defined by system. Add the 1/2, 2/3, and 3/4 variants of the nominal SDTV itemized PAR’s. Sequence level. 8 bit table.	bug-fix/clean-up
JVT-C136	Rounding, QP Origin, Dynamic Range, and |f| item 1: Adopted. item 2: Agree to initialize QP at slice level relative to middle of range. Check current design at MB level and use that with mid-range prediction for slice level. Note: Could change after consideration of JVT-C079. item 3: Adopt the first of the suggested alternative variants. item 4: Adopted. item 5: Adopted. item 6: Adopt subpoint 2, not the rest item 7: Not adopted. item 8: No immediate action requested.	bug-fix/clean-up
JVT-C137	Cropping, Generalized Pan-Scan, Source Formats Cropping and pan-scan concepts adopted. Source format modifications not adopted.	bug-fix/clean-up
JVT-C138	TR's, PN's, Pictures, Frames, and Fields	bug-fix/clean-up

Nokia Exhibit 2024, p.45
Snap/Hisense v. Nokia
IPR2025-01440

	27 MHz support via general time clock support. Dangling field support adopted Separate timing information for each field adopted. SEI for SMPTE-compatible compressed timing adopted Picture numbering aspects for further study	
JVT-C162 subject to cross-verification and some further study	Putting a Reasonable Upper Limit on Binarization Expansion	bug-fix/clean-up
JVT-C166	Definition of HRD for VBR and CBR and low delay mode. Includes method of handling header-level data. CBR can be same as VBR with average rate equal to peak rate. Explicit support of 3:2 pull-down is included. Relationship to other timing-related issues for further study.	Clean-up

16.2.5. The JVT authorizes an editing period for production of the Joint CD text by the editor, to be provided by May 12, 2002.

16.2.6. The JVT reports the following summary information on the draft content of its interface to external systems-level specifications:

The joint video specification defines a common representation of all parameter set data, slice data, and "SEI" data units that are carried in the same way in bitstreams and in packets. Each atomic chunk of such data is referred to as a NAL unit, where NAL is "network abstraction layer". This includes ensuring no emulation of start code prefixes of the bitstream format within a NAL unit.

The joint video specification includes definition of a bitstream format (in a normative annex). This format

- is capable of conveying all data "in band".
- includes sufficient information to identify conformance of the video bitstream and this requires inclusion of sufficient relative timing information.
- is capable of conveying temporal reference (relative timing data) within the video bitstream for use in HRD conformance verification

A packet-oriented interface is also defined within the joint video specification.

To support simple conversion between these two formats, distinct parts of the joint specification describe what data needs to be carried, and

- the bitstream definition section defines how to put all that into a bitstream
- the packet-oriented definition section defines how to put this data into packets and what other information (such as timing information) is necessary for using those packets

Compound packets are not defined within the joint video spec. If used, these are to be defined externally.

In the bitstream format, NAL units are prefixed by a unique start code prefixes and temporal reference data is included at the picture or slice level. In the packet format, timing information support is defined externally.

Further information on these aspects is available in document JVT-C168.

16.2.7. The JVT reports the following summary information regarding Requirements and the definition of normative Profiles and Levels

Nokia Exhibit 2024, p.46
Snap/Hisense v. Nokia
IPR2025-01440

Regarding Profiles

- The Baseline Profile includes
  - I and P picture types
  - In-loop deblocking filter
  - Interlace (level-dependent, levels 2.1 and above)
  - 1/4-sample motion compensation
  - Tree-structured motion segmentation down to 4x4 block size
  - VLC-based entropy coding
  - Flexible macroblock ordering (maximum 8 groups)
- This includes all normative decoding features except:
  - B pictures
  - CABAC
  - Adaptive block-size transforms
  - 1/8-sample motion compensation
  - Mixing intra and inter coding types within a macroblock
  - Data partitioning
  - SP & SI "switching" pictures
- One additional profile is defined, the Main Profile. This adds the following differences relative to the Baseline Profile:
  - B pictures
  - CABAC
  - Adaptive block-size transforms
- Notes on Profiles and Levels
  - MV range will be limited
  - A limit is imposed on extreme aspect ratios
  - Number of reference pictures at highest supported picture size is level-dependent
  - Number of reference pictures increases when picture size is smaller than the highest supported size for the level, never exceeding 15
- TBD's
  - Exact values of MV range limits
  - Whether to allow smaller than 8x8 bi-predictive motion in B pictures for Main profile
  - Whether to include adaptive B picture interpolation coefficients for Main profile

Regarding Levels (in both profiles) are summarized with example typical format support as follows

- Level 1=QCIF@15 (Intermediate Levels 1.1=CIF@7.5 and 1.2=CIF@15)
- Level 2=CIF@30 (Intermediate Levels 2.1=HHR and 2.2)
- Level 3=SDTV (Intermediate Levels 3.1 and 3.2)
- Level 4=HDTV
- Level 5=SHDTV (1920×1088×60p)

Regarding Requirements

  - No 4:2:2 source format support is currently defined (pending further study and sufficiently mature contribution to next meeting)

16.2.8. Regarding WG11 National Body comments on technical content of the JVT draft, the JVT thanks the USNB to WG11 for its comments and reports that the actions taken by the JVT are consistent with all of the USNB requests.

16.2.9. Regarding the IPR status of the JVT draft text, the JVT reports that the specification of the Baseline Profile Tools as defined in the Joint CD is complete and is fully consistent with the Terms-of-Reference goal of achieving a royalty-

Nokia Exhibit 2024, p.47
Snap/Hisense v. Nokia
IPR2025-01440

free baseline profile status for technology essential for implementation. Further reports to the JVT regarding the IPR status of the draft text are requested.

16.2.10. The JVT thanks the US National Institute of Standards and Technology, and Charles Fenimore in particular, for aid in providing equipment and support for demonstrations of video technology designs at the meeting.

16.2.11. The JVT thanks the USNB to WG11, and the sponsors of the social event, including Contentguard, INCITS, MPAA, Microsoft Corporation, NIST, RIAA, and Rightscom for hosting the JVT and providing excellent arrangements for our work.

The 3rd (Fairfax) JVT meeting was closed at 1:30pm Friday, May 10, 2002

Nokia Exhibit 2024, p.48
Snap/Hisense v. Nokia
IPR2025-01440