Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG
(ISO/IEC JTC1/SC29/WG11 and ITU-T SG16 Q.6)
2nd Meeting: Geneva, CH, Jan. 29 - Feb. 1, 2002

Document: JVT-B001 draft 1
Filename: JVT-B001d1.doc
Generated: 2002-02-01

Title: **JVT (of ISO/IEC MPEG and ITU-T Q.6/16 VCEG) 2nd Meeting Report, Jan 29 – Feb 1 2002**

Status: Output Document Approved by JVT

Purpose: Report

Author(s) or Contact(s):

Gary Sullivan
JVT Rapporteur | Chair
Microsoft Corp.
One Microsoft Way
Redmond, WA 98052
Tel: +1 (425) 703-5308
Email: garysull@microsoft.com

Thomas Wiegand
JVT Associate Rapporteur | Co-Chair
Heinrich Hertz Institute,
Einsteinufer 37,
10587 Berlin
Tel: +49 (0)30 31002 617
Email: wiegand@hhi.de

Ajay Luthra
JVT Associate Rapporteur | Co-Chair
Motorola, BCS
6420 Sequence Drive
San Diego, CA 92121
Tel: +1 (858) 404-3470
Email: Aluthra@gi.com

Source: Rapporteurs | Chairs

---

1. Organization of the Meetings

The Joint Video Team (JVT) or ITU-T VCEG (Q.6/SG16) and ISO/IEC MPEG (ISO/IEC JTC1/SC29/WG11) met at ITU-T Headquarters in Geneva during Jan 29 to Feb 1, 2002. The meeting was chaired by the JVT chair and co-chairs. This document reports the results of that meeting.

2. List of Documents

* Indicates on-time availability on ftp site

& Indicates late availability for non-report documents (alicui tardus opprobrium receptat)

```
JVT-B-TD0*Report     [Sullivan+]   Invitation to the Meeting
JVT-B000* Report     [Sullivan+]   List of Documents
JVT-B001  Report     [Sullivan+]   Report of Geneva Meeting (#2)
JVT-B002  Report     [Sullivan+]   Report of Pattaya Meeting (#1)
JVT-B003  Report     [Sullivan+]   List of Geneva Participants
JVT-B004  Report     [Sullivan+]   List of JVT Experts
JVT-B005* Report     [Sullivan]    AHG Report: JVT Project
JVT-B006  Report     [Wiegand+]    AHG Report: WD & JM Doc & S/W Editing
JVT-B007  Report     [Lindbergh]   AHG Report: Profiles & Applications
JVT-B008* Report     [Kerofsky]    AHG Report: Transform and Quantization
JVT-B009  Report     [Wiegand]     AHG Report: Generalized ERPS
JVT-B010* Report     [Borgwardt]   AHG Report: Interlace
JVT-B011  Report     [List]        AHG Report: Loop Filter
JVT-B012* Report     [Horowitz]    AHG Report: Complexity
JVT-B013* Report     [Viscito]     AHG Report: Buffering
```


INTEL-1105
9,641,841

JVT-B014*	Report		[Benzler]	AHG Report: Transform Size
JVT-B015*	Report		[Marpe]	AHG Report: CABAC
JVT-B016*	Report		[Lim+]	AHG Report: Transport
JVT-B017*	Report		[Suzuki]	AHG Report: GOP Syntax
JVT-B018*	Report		[Johansen+]	AHG Report: Motion Interpolation
JVT-B019*	P2.2*	/3.1	[Sun & Lei]	Global Motion Vector Coding (GMVC)
JVT-B020*	P ?	/ ?	[Winger+]	Interlaced Field Coding Core Experiment A/B
JVT-B021*	P2.2	/3.1	[Hong & Oh]	A modified loop filter of VCEG-O19
JVT-B022*	P2.2	/3.1	[Hong & Oh]	Range decision for motion est. of VCEG-N33
JVT-B023*	Admin		[Koenen]	Remarks on Profiling
JVT-B024*	Info		[Wenger+]	Coding Performance not using MV Prediction
JVT-B025*	P2.0	/3.1	[Wenger+]	On Equivalence of BERs and Packet Loss Rates
JVT-B026*	Info		[Wenger+]	RTP-NAL and RTP packetization
JVT-B027*	P2.2.1/3.1		[Horowitz+]	Scattered Slices Error Concealment
JVT-B028*	Info		[Stockhammer+]	Overview of NAL Concept & VCL/NAL Interface
JVT-B029*	P2.2.1/	?	[Kerofsky+]	Reduced Complexity VLC
JVT-B030*	P2.2.1/3.1		[Zhou]	Simplification of H.26L baseline coding tools
JVT-B031*	P2.2.1/3.1		[Zhou]	16-bit based transform and quantization
JVT-B032*	P2.2.1/3.1		[Zhou]	P-frame coding with interpolative prediction
JVT-B033*	P2.2	/3.1	[v-d Vleuten]	Low-complexity arith. coding implementation
JVT-B034*	P2.0	/3.1	[Halbach]	Enhanced variable-length coding
JVT-B035*	P2.0	/3.1	[Lindbergh]	Updated Profile Framework for JVT codec
JVT-B036*	P2.0	/3.1	[Winger]	Low-Complexity Arithmetic Codec Engine
JVT-B037*	P2.2	/3.1	[Le Maguet+]	PPL, a low Complexity deblocking for JVT
JVT-B038*	P2.2.1/3.1		[Hallapuro+]	Transform and Quantizer - part 1: Basics
JVT-B039*	P2.2.1/3.1		[Hallapuro+]	Transform and Quantizer - part 2: Extensions
JVT-B040*	P2.2.1/3.1		[Wang+]	Results of Core Exp. on Sub-Picture Coding
JVT-B041*	P2.2.1/3.1		[Hannuksela]	Simple Definition of GOP for Random Access
JVT-B042*	P2.2	/3.1	[Hannuksela]	Enhanced Concept of GOP
JVT-B043*	P2.2	/3.1	[Hannuksela]	Coding of Scene Transitions
JVT-B044W	--		[Withdrawn]	[Withdrawn]
JVT-B045*	P2.2.1/3.1		[Bjontegaard]	Low complexity Entropy Coding for Coeff's
JVT-B046*	P2.2	/3.1	[Kimata]	GMVC and GMC switched by MV
JVT-B047*	P2.2.1/3.1		[Zhou]	Removal of EOB Coding Redundancy
JVT-B048*	P2.2	/3.1	[Chen+]	Supporting Film Mode in JVT Codec
JVT-B049*	P2.0	/3.1	[MacInnis+]	Start Codes and Mapping to MPEG-2 Systems
JVT-B050*	P2.2.1/3.1		[Ribas+]	Including Complexity Constraint in HRD
JVT-B051*	P2.2	/ ?	[Yamada+]	Improved transform coding for inter-frame
JVT-B052*	Info		[Sekiguchi+]	Core Exp. Result on Improved MB Pred Modes
JVT-B053*	P2.0	/3.2	[Wien]	ABT Coding for Higher Resolution Video
JVT-B054*	P2.0	/3.1	[Schwarz+]	Core Exp. Result on Improved MB Pred Modes
JVT-B055*	P2.2.1/3.1		[Kurceren]	New Macroblock modes for SP-frames
JVT-B056*	P2.2	/3.1	[Kadono+]	Improved 2D-VLC for High Bit Rate
JVT-B057*	P2.2	/3.1	[Kondo+]	Multiframe Syntax for B-frame Efficiency
JVT-B058*	Info		[Hagai+]	Core Exp. Result on Improved MB Pred Modes
JVT-B059*	P2.2	/3.1	[Hagai+]	Intra Prediction for Improved MB Pred Modes
JVT-B060*	P2.0	/ ?	[Boyce]	Coding Eff. of Various #s of Ref. Frames
JVT-B061*	P2.2	/3.1	[Suzuki]	Adaptive Motion Vector Coding
JVT-B062*	P2.2.1/3.1		[Jeon]	Entropy Coding with Codeword Re-Association
JVT-B063*	P2.2.1/3.1		[Sullivan]	On Random Access and Bitstream Format
JVT-B064&	Info		[Bäse]	Additional Results related to the CABAC AHG
JVT-B065*	P2.2.1/3.1		[Benzler]	ABT Results for Interlaced ITU-R 601 Video
JVT-B066*	Info		[Wedi]	1/8-pel MC for interlaced video
JVT-B067*	P2.2	/3.1	[Suzuki+]	Quantization Tools for High Quality Video
JVT-B068*	P2.2	/3.1	[Sato+]	New Interlaced Coding Tools
JVT-B069*	P2.2	/3.1	[Suzuki+]	Group of Pictures for JVT codec
JVT-B070*	P2.0	/3.1	[Suzuki+]	NAL for MPEG-2 System
JVT-B071*	P2.2	/ ?	[Wang+]	Adaptive frame/field coding for JVT video
JVT-B072*	Info		[Karczewicz]	VLC Coef Coding for High Bitrate
JVT-B073*	Report		[Sullivan]	JVT IPR Status Report
JVT-B074*	P ?	/ ?	[Chujo+]	An improved variable length coding
JVT-B075&	P ?	/ ?	[Kikuchi+]	Multiframe MC using frame interpolation
JVT-B076&	P ?	/ ?	[Kikuchi+]	New picture type for error recovery
JVT-B077*	P ?	/ ?	[Miyamoto+]	Short tap filter for high res sequences
JVT-B078*	Info		[Miyamoto+]	Study on in-loop filtering issues
JVT-B079*	P ?	/ ?	[Lee]	Loop filter using DCT coeff distribution
JVT-B080*	I2.2.1/3.1		[Conklin]	More Results on New Intra Prediction Modes
JVT-B081*	P2.2.1/3.1		[Kim+]	Adaptive use of Double Scan
JVT-B082*	Info		[Kim+]	Brief Result Core Exp. VCEG-O57 (Sub-Pic)

```
JVT-B083W --         [Withdrawn]   [Withdrawn]
JVT-B084* P2.2  /3.1 [Foo]         Downloadable Thresholds for Loop Filter
JVT-B085* P2.2  /3.1 [Foo]         Multiframe Buffer and L-T Buffer [Withdrawn]
JVT-B086* Info       [Kim+]        Results on core Exp. of Sub-Picture Coding
JVT-B087* P2.2.1/  ? [Kim+]        Selection of QP for Sub-Picture Coding
JVT-B088* P2.2  /3.1 [Hannuksela]  MPEG-2 Systems NAL
JVT-B089* P2.2.1/3.2 [Hannuksela]  Complexity-Constrained Generalized HRD
JVT-B090* P2.2  /3.1 [Adachi+]     Core Exp. Results Improved MB Pred Modes
JVT-B091* P2.2  /3.1 [Etoh+]       Proposed Requirements on CABAC
JVT-B092* P2.0  /3.1 [Etoh+]       Proposed Requirements on NAL Specification
JVT-B093* P2.0  /3.1 [Kato+]       Exp. Results of Improved 2D-VLC Proposal
JVT-B094* Info       [Cheong+]     Water Ring Scan for H.26L-based FGS
JVT-B095* P2.0  /3.1 [Zhou+]       Fragile Watermark Detection
JVT-B096W --         [Withdrawn]   [Withdrawn]
JVT-B097& P2.2.1/3.1 [Sun+]        Advanced SP Coding Technique
JVT-B098W --         [Withdrawn]   [Withdrawn]
JVT-B099* Info       [Takagi]      Reversibility of Codes
JVT-B100& P2.0  /3.1 [Marpe+]      Performance of CABAC for Interlaced Video
JVT-B101& P2.0  /3.1 [Marpe+]      New Results on Improved CABAC
JVT-B102* P2.2.1/3.1 [Stockhammer+]Error Robust MB Mode and Ref Pic. Select.
JVT-B103* P    ?/  ? [Liang+]      FastVDO Unified 16-Bit Transform/Quant.
JVT-B104* P2.0  /3.1 [Wenger]      Allow Start Code Emulations
JVT-B105& P2.1  /3.1 [Hamilton+]   Improved CABAC
---Late-Registered Documents Below ---
JVT-B106& P2.2  /  ? [Wang+]       MB level adaptive frame/field coding
JVT-B107& Info       [Andrews]     IPR statement from 8x8
JVT-B108  Report                   Profile & Levels Framework Output Document
JVT-B109  Report                   Random Access and Time
JVT-B110  --                       Withdrawn
JVT-B111  Report     [Wenger]      CE on Scattered Slices
JVT-B112  Report                   CE on SP Pictures
JVT-B113  I2.2.B/3.1 [Sun]         IPR statement from Sharp
JVT-B114  I2.2.B/3.1 [Kerofsky]    IPR statement from Sharp
JVT-B115  Report                   CE on Adaptive MV Coding
JVT-B116  Report                   CE on Interlace Chroma Phase Shift
JVT-B117  Report                   CE on MB Adaptive Frame/Field
JVT-B118  Report/Draft             Joint Working Draft 2 / Joint Encoding Model 2
```

3. Detailed Agenda with Disposition of Contributions

3.1 Administrative Issues

JVT-B000* [Sullivan+] List of Documents

JVT-B001 [Sullivan+] Report of Geneva Meeting (#2)

JVT-B002 [Sullivan+] Report of Pattaya Meeting (#1)

JVT-B003 [Sullivan+] List of Geneva Participants

JVT-B004 [Sullivan+] List of JVT Experts

3.2 Ad Hoc Reports

JVT-B005 [Sullivan] AHG Report: JVT Project

The JVT project ad hoc group report contained a review of the general status of the project, and in particular announced a new ftp site sponsored by IMTC:

ftp://ftp.imtc-files.org/jvt-experts/

The general JVT reflector can be subscribed to at

http://www.imtc.org/scripts/imtc.pl?enter=jvt-experts

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

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

Preparations for this meeting had used the prior ftp site, at which the documents for the meeting were available:

http://standard.pictel.com/ftp/video-site/0201_Gen

JVT-B006 [Wiegand+] AHG Report: WD & JM Doc & S/W Editing
Report of software development status and working draft status.

Recommendation of editing coordination structured into subject-area coordinators overseen by the editor.

JVT-B073* [Sullivan] JVT IPR Status Report
General review of IPR situation with remarks on possible problem areas and open issues.

The group was informed about JVT's patent policy. The IPR status was reviewed. The known patent holders for the JVT project were mentioned. The question regarding the support of a royalty-free baseline was asked. Telenor and 8x8, Inc. indicated an intent to follow sub-clause 2.2.1. (Telenor indicated that the spirit of 2.2.1 was expressed in the Telenor email message of 6 months ago.) Philips indicated to follow sub-clause 2.2.

Thus the Philips situation appears to be the only identified problem in reaching the royalty-free baseline goal.

3.3 UVLC

JVT-B074* [Chujo+] An improved variable length coding
A new variable length code to improve the coding efficiency is proposed. Simulation results show that the proposed VLC outperforms UVLC for high bit rates. The Proposed VLC is better for all sequences when QP is small. The coding gain is up to 6.2% for Inter coding and QP=1. Gains depend on the sequence. For increasing QP, the performance becomes worse than UVLC. The codewords of the proposed VLC are regular and without interleaving. It also provides bit-error detection mechanisms.

JVT-B099* [Takagi] Reversibility of Codes
In this contribution, the usefulness of the reversibility of codes is discussed. A method for systematically producing Reversible VLCs is provided.

Group: results would be needed to study the efficiency of this technique.

JVT-B034* [Halbach] Enhanced variable-length coding
This contribution investigates error resilience properties of two codes, UVLC and the RVLC of H.263, Annex D. It is claimed that the UVLC has been used so far in H.26L standardization, and VLCD is a reversible VLC that offers error resilience/concealment possibilities which are superior to UVLC.

Group: The various features are well understood. However, in order to change the design, the group needs more evidence on how these methods could be applied beneficially in a practical system.

JVT-B029* [Kerofsky+] Reduced Complexity VLC
This document supports the previous proposal VCEG-N36 by repeating remarks about the UVLC performance under bit errors given in VCEG-L23. As mentioned in VCEG-N36, the interleaved structure of the UVLC codewords is complex and can be reduced by using a non-interleaved structure. VCEG-L23 is used to evaluate the sacrifice in removing the resynchronization property of the UVLC.

Group: Provides more flexibility for adaptive codes. Can be used for binarization in CABAC.

JVT-B047* [Zhou] Removal of EOB Coding Redundancy
This document suggests to remove the EOB coding redundancy by transmitting the EOB only for the coded blocks in which the last coefficient is zero. By doing so there is no complexity increase, the error detection feature is still maintained. Up to 1% overall bit-rate saving is measured. Removing the EOB coding redundancy would add one more differentiation to the H.26L coding standard. Defer decision, since it is void if the coeff count techniques are adopted.

Group: Some error detection mechanisms rely on the presence of EOB.

UVLC Topic Area Results:

We want to change the UVLC.

The Exp-Golomb Code provides simplification and flexibility.

Non-synchronizing reversible codes may provide enhanced error resilience.

Propose to adopt the Exp-Golomb codes per JVT-B029 unless the benefit of the enhanced error resilience feature of the other codes is shown.

3.4 Context-Based Adaptive Codes (CAC)

JVT-B045* [Bjontegaard] Low complexity Entropy Coding for Coeff's
An adaptive, low complexity entropy coding method for transform coefficients is proposed. It contains the idea of switching between various VLCs based on contexts. Overall coding gain between 0-9% (QP=10) compared with UVLC coding is reported for Inter Coding and between 0-8% (QP=10) for Intra coding.

Group: impressive gains. Test set used to design tables. Question about complexity was raised.

JVT-B072* [Karczewicz] VLC Coef Coding for High Bitrate
The goal of the proposal is to show that VLC coding can be considerably improved for high rates or/and high resolution material and initiate the discussion is there an interest to use VLC coding in such cases. Proposal uses coefficient counter and parameterizable Exp-Golomb codes. The measured bit-rate reductions are between 0-20 % at low QP values for Intra and between 0-13 % for low QP values for Inter.

CAC Topic Area Results:

Group: Great interest in these methods to achieve bit-rate savings. CAC AHG: Chair Gisle: Mandate is to further study the idea of adaptive codes with regards to the trade-off between coding efficiency and complexity involving the Exp-Golomb and the CABAC entropy coding.

3.5 Other VLCs and Scanning

JVT-B056* [Kadono+] Improved 2D-VLC for High Bit Rate
An improved 2D-VLC encoding scheme using UVLC is proposed. (1) Level scaling 2D-VLC table design, For small QP case, allocation of code number is adaptive for high value. (2) Position adaptive 2D-VLC table design, Long run is removed when encoder know there will be no long run. Simplification: (3) Inverse order zigzag scanning. Combination with 2nd Proposal, Large coefficient should be coded for the last part of scanning, propose inverse order zigzag scanning. The proposal is verified using JVT sequences, and it is shown to reduce 3-7% bits for I-picture and 0-2% bits for P-pictures at small QP. A fourth method is proposed to turn on/off scaling to never degrade performance.

JVT-B093* [Kato+] Exp. Results of Improved 2D-VLC Proposal
Verification results of the improved 2D-VLC encoding scheme, which have been proposed in VCEG-O27, are reported. The results verify the results in VCEG-O27. At the same time, the results show the necessity to avoid drawback on some encoding parameters and/or sequence characteristics.

JVT-B081* [Kim+] Adaptive use of Double Scan
It is proposed to adaptively use double / simple scan. A delta value to adjust the threshold which separates simple and double scan is proposed. Results show that a bit rate reduction at near the QP=24 up to 5% can be achieved. Further consideration of this proposal must be conducted in the light of the results of the CAC AHG.

JVT-B062* [Jeon] Entropy Coding with Codeword Re-Association
In this document a method for improving the current UVLC approach by using fixed re-association table (FRAT) that assigns existing codewords to different symbols according to QP values is proposed. It is proposed to use this approach for high bit rate applications where statistics are different from the cases

with medium and low bit rates. The mapping tables are divided according to quantizer ranges. E.g.: 3 Ranges: QP=1-10, QP=11-20, QP=21-31. Gains up to 8% for QP=1 for Mobile for Intra coding, while normal gains are around 2-3%. Sometimes, results are worse than UVLC. Results are worse for Inter coding than for Intra. Also switches between active and non-active sequences investigated. Encourage further work.

Other VLC Topic Area Result:

Further proposals in this area should show performance on top of CAC result. Coordinate with AHG.

3.6 Context-Based Adaptive Binary Arithmetic Coding (CABAC)

JVT-B015 [Marpe] AHG Report: CABAC
This report summarizes the CABAC related proposals presented at that last meeting, gives an overview of the activities within the CABAC AHG in the interim period between the last and the present meeting, and finally provides a list of CABAC related contributions registered/uploaded so far for the Geneva meeting.

JVT-B101 [Marpe+] New Results on Improved CABAC
New coding elements and new context models for improving coding efficiency of CABAC are proposed, both for inter and intra frame coding. This proposal combines ideas of earlier contributions. Simulation results show performance gains of up to 5.2% BD bit-rate savings relative to the current CABAC specification for intra coding. In the case of inter coding (IPPP) improvements of up to 2.4% BD bit-rate reduction relative to the original CABAC method have been obtained.

Group: Propose to adopt the improved CABAC method of JVT-B101. (revised version expected)

JVT-B100 [Marpe+] Performance of CABAC for Interlaced Video
The improved CABAC entropy coding method is evaluated with regard to interlaced source material. The simulation results show that CABAC significantly improves the rate-distortion performance for the specified set of test sequences. An average BD bit-rate savings of 11%, or equivalently, BD PSNR gains of 0.45 dB in comparison to the UVLC entropy coding method is observed.

JVT-B033* [v-d Vleuten] Low-complexity arith. coding implementation
The proposal significantly reduces the arithmetic coding implementation complexity, while it has a negligible influence on the compression efficiency. Although the method was originally designed and optimized for hardware implementation, it also significantly improves the software execution speed. No degradation against original CABAC.

JVT-B036* [Winger] Low-Complexity Arithmetic Codec Engine
Two arithmetic coding engines are presented: CACM+/ACM98 (improved AC engine from "Arithmetic Coding Revisited", Moffat, Neal, Witten. ACM Transactions on Information Systems, 16(3):256-294, July 1998.) and WAC (Xiaolin Wu's proposed low-complexity AC engine).

Group: Propose to adopt CACM+ of JVT-B036.

JVT-B105 [Hamilton+] Improved CABAC
This contribution proposes a new context-based binary arithmetic coder for use by H.26L for entropy coding of transform coefficients. The proposal considers only the texture coding part. It is demonstrated to produce greater compression performance as compared to H.26L TML9.4 CABAC by up to 8% bit rate reduction for texture data. Overall the gains are up to 5.72 %. The proposal didn't use the common test conditions. Further work is encouraged.

JVT-B091* [Etoh+] Proposed Requirements on CABAC
Requirements for CABAC design are proposed. First requirements are about arithmetic coding core, in terms of feasibility with small mobile devices that require low-power consumption. Second requirements are about context modeling, in terms of coding efficiency degradation sacrificed by the Unified Binarization scheme.

CABAC Topic Area Results:

Propose to adopt the improved CABAC method of JVT-B100. Propose to adopt CACM+ of JVT-B036 and continue to discuss complexity reduction of the arithmetic coding engine in the AHG on CABAC.

JVT-B064 [Bäse] Additional Results related to the CABAC AHG
Notes elsewhere.

3.7 Motion Compensation

JVT-B018* [Johansen+] AHG Report: Motion Interpolation
Thomas Wedi is porting and incorporating his implementation to the JVT software. Although many experts have shown interest in the subject, no new parties have come forward with contributions. This is mainly believed to be due to the big effort required to start this work from scratch and the short time between Pattaya and this meeting. Recommended to continue the AHG with (Chair: Tom-Ivar Johansen, co-chair: Thomas Wedi) with the previous mandate plus finalization of the work for potential inclusion at the next meeting.

JVT-B066* [Wedi] 1/8-pel MC for interlaced video
The contribution analyses the coding efficiency of 1/8-pel motion vector resolution for interlaced sequences. For sequences with active motion (Football, Rugby,...) no coding gain is obtained with 1/8-pel compared to 1/4-pel MCP. The PSNR difference for this kind of sequences is in a range of 0.0 and -0.2 dB. For sequences with moderate motion and high spatial detail (Bus, Flower,...), a gain between 0.5 and 1.0 dB is obtained for higher bitrates of the test set.

JVT-B077* [Miyamoto+] Short tap filter for high res sequences
This contribution reports the experimental results comparing the TML 6-tap filter and the Telenor 4-tap filter. It is reported that the gains are in the range of 0-10% BDBR. It is pointed out that the 4-tap implementation can reduce interpolation calculation up to 50% for that part of the decoder. It was suggested to make filter size a profiling and level issue.

3.8 Macroblock Partition

JVT-B090* [Adachi+] Core Exp. Results Improved MB Pred Modes
This document reports the core experiment results of improved MB prediction modes. B-frame syntax for VCEG-O22, which are newly defined in the CE description, is also described. The results show that up to 8.5% bitrate reduction or 0.41dB PSNR improvement can be seen by the method of VCEG-O22, and up to 5.2% or up to 0.25dB improvement by the method of VCEG-O17.

JVT-B054* [Schwarz+] Core Exp. Result on Improved MB Pred Modes
An improvement of the proposal "Tree-structured macroblock partition" (VECG-O17) is presented. Furthermore, the results of the Core Experiment on Improved Macroblock Prediction Modes defined in VCEG-O61 are reported. The proponents report similar or reduced complexity.

JVT-B052 [Sekiguchi+] Core Exp. Result on Improved MB Pred Modes
This is basically the informative verification report of the technique proposed in VCEG-O22 at Pattaya. Our independent implementation shows comparable coding gain which was reported in Pattaya meeting.

JVT-B058* [Hagai+] Core Exp. Result on Improved MB Pred Modes
VCEG-O61 (VPM2) method is verified by independently developed software. It reduces bit 3.60% for single reference frame and 4.10% for 5 reference frame, in average.

JVT-B059* [Hagai+] Intra Prediction for Improved MB Pred Modes
The method introducing intra-coded segment into VCEG-O22 method is proposed. That segment is coded using intra 4x4 mode coding scheme. It reduces bit-rate by 0.8% comparing to the original VCEG-O22, therefore 4.85% reduction comparing to TML9.0. These additional intra coding modes need to be tested to achieve the performance.

Remarks: VCEG-O22 could require calculation of 4x4 SADs which is computationally complex. How important is encoder complexity as it is not required to check all of the new modes of VCEG-O22? The scheme of VCEG-O22 gives about the same performance regardless of whether the 4x4, 4x8 and 8x4

modes are used. The original aim of VCEG-O22 was to improve coding efficiency at low bit rates – which it achieves, but perhaps not to the extent that may have been expected.

Macroblock Partition Subject Results:

Agreed to include VCEG-O17 into the test model at this meeting. Further information regarding the complexity of VCEG-O22 is requested.

3.9 Multiframe Motion Compensation

JVT-B009 [Wiegand] AHG Report: Generalized ERPS
The AHG report indicates that the goal of integration of ERPS into the design has been achieved.

JVT-B032* [Zhou] P-frame coding with interpolative prediction
This document describes a way to change the baseline codec by introducing an interpolative prediction mode in the existing P-frame coding. In the proposed method, the reference frame number is limited to two, in addition to the forward prediction from the two reference frames, a macroblock can also use interpolative prediction which is similar to the bi-direction prediction mode in the B-frames. Experimental results reveal that this maybe an effective way to cut the reference frame number while still maintaining the coding efficiency.

JVT-B057* [Kondo+] Multiframe Syntax for B-frame Efficiency
The prediction method for referencing B-pictures for coding of the next B-picture are. BDBR between 0-3% are shown overall while for B-pictures, the gains are up to 6%. For further improvement, a new direct prediction mode is introduced which almost doubles the gains when referencing B-pictures for coding of the next B-picture. Various problems to support such a coding method are outlined for the syntax.

Group recommends to form an AHG on B-Picture coding (Chair: Satoshi Kondo) Mandate: finalize B-picture syntax taking into account JVT-B057.

JVT-B043* [Hannuksela] Coding of Scene Transitions
The document proposes composition of scene transitions in the decoder. Component pictures are coded and decoded separately and a transition filter is applied to reconstructed versions of the component picture to obtain a picture to display. The proposal is likely to improve compression efficiency and enhance picture quality in many scene transition cases.

A remark was made that this technique could provide coding gains in case of long fades.

Further consideration and comment on JVT-B043 is requested.

JVT-B075 [Kikuchi+] Multiframe MC using frame interpolation
A new motion compensation method employing frame interpolation with weighted sum of multiple reference frames is proposed. Simulation results show that the proposed method gains SNR improvement of up to 0.5 dB at high bit-rates, even with a simple mean weighting. Significant coding gain of 0.5 to 2 dB is derived for fading sequences. There is no need to calculate a fading factor at the encoder side because the proposed method works as extrapolation of the brightness.

Should include into B-frame AHG work.

JVT-B085* [Foo] Multiframe Buffer and L-T Buffer
Withdrawn.

3.10 Global Motion Compensation and Motion Vector Coding

JVT-B046* [Kimata] GMVC and GMC switched by MV
A modified syntax for GMVC and GMC is proposed. The hook of switching a GMVC or GMC block is integrated into the motion vector syntax. The proposed scheme improves coding efficiency especially in zoomed sequences. Bit-rate savings are reported to be up to 7.8 %.

Request to make experiments with *B-pictures*: it is pointed out that for P-pictures which are temporally further spaced apart, the gain for P-pictures increases. It was pointed out that for B-pictures, the gains are typically smaller.

JVT-B019* [Sun & Lei] Global Motion Vector Coding (GMVC)
The GMVC technique is revised and proposed in this contribution. Compared to the previous GMVC proposal, the "GMVC mode with coefficients" has been rearranged to gain better bit savings. Experiment shows significant bit savings and some visual quality improvements at very low bit rate when global motion is present.

No gains at medium bit-rates.

A comment was made with regard to *skipping at the MV predictor*.

Gains are reduced to 30-50% for B-pictures.

Found an AHG on GMVC/GMC (Chair: Kimata-San, co-Chair: Lainema-San) mandate: study technique for inclusion employing JVT-B046 and JVT-B019. Studying the use of MB skip at predictor.

3.11 De-Blocking Filter

JVT-B011 [List] AHG Report: Loop Filter
The AHG recommends to adopt the improved software into the JM. The reduction in software complexity was estimated to be a factor of 2-4. A remark was made that the reduction in hardware complexity is smaller than the complexity reduction in software. The changes that need to be made to the document were presented to the group. Subjective results were shown and an improvement was seen by the group. It is proposed to adopt this loop-filter. It is also proposed to continue the loop-filter AHG (chair: P. List). Extend Mandate to consider finalization of the loop-filter design. Also consider the case of varying QP on a MB-per-MB basis.

JVT-B021* [Hong & Oh] A modified loop filter of VCEG-O19
This contribution proposes a frame-based loop filter that introduces QP and signal dependent adaptive filtering. The difference to VCEG-O19 is that it does not use a temporal frame buffer to reduce the complexity. Instead it uses different filter coefficients depending on the signal (luma or chroma). An average saving of 15% against the loop filter decoding time is obtained against the existing JM implementation. Subjective comparisons shown so far indicated a potential subjective gain of the proposed method. The proposal will be given further consideration in the loop-filter AHG.

JVT-B037* [Le Maguet+] PPL, a low Complexity deblocking for JVT
A new deblocking algorithm is proposed. The deblocking filter has been tested as a post process function and is to be implemented in the coding loop soon. The algorithm has been compared to the TML9 deblocking in terms of both quality and complexity. The results so far show that it is equivalent to the TML9 deblocking in terms of perceived visual quality (for CIF sequences), and is 3.5 times less complex. Subjective results indicate some improvements of the proposed method while the method gives objective losses up to 0.5 dB BDPSNR. The proposal will be given further consideration in the loop-filter AHG.

JVT-B078* [Miyamoto+] Study on in-loop filtering issues
This contribution concerns two in-loop filter issues. First one is the necessity of in-loop filter. At middle Qp-values, it is argued that the filter should be switched off for medium bit-rates but no results were shown. The other is a comparison between MB-based and Frame-based filtering. It is claimed that MB-based filtering shows the better performance both in coding gain and subjective quality. Group: the differences shown for the two methods are larger than some experts experienced in the past. The adopted design now uses MB-based filtering.

JVT-B079* [Lee] Loop filter using DCT coeff distribution
In this contribution, the reconstructed signal is processed by a 4x4 DCT. The 4x4 DCT coefficients distribution is used to control the deblocking filtering for the 4x4 block boundary in that it provides the horizontal and vertical blocking information. Results: The lastest JM optimized loop filter distributed by Peter List is used for comparison. PSNR results are similar, while the group preferred the subjective quality of the latest JM. Complexity: optimized JM loop filter has lower complexity. The proponent

indicated the potential for further complexity reduction. Bit-rate reduction: optimized JM loop filter has 0.5% more bit-rate reduction. The group remarked that the optimized JM loop filter shows some color bleeding. Encourage further work.

JVT-B084* [Foo] Downloadable Thresholds for Loop Filter
This submission is to propose the use of a flag to indicate whether to use the standard threshold tables or to use user defined threshold tables to improve the subjective quality. The subjective results shown provided some benefit for the proposed technique. Group found the technique to be promising and wants to further investigate it in the loop-filter AHG.

JVT-B061* [Suzuki] Adaptive Motion Vector Coding
This contribution provides a new motion vector coding method for reducing the amount of the motion vector data. The idea of this proposal is to adaptively select the fractional pixel accuracy of differential motion vector components macroblock by macroblock. The proposed scheme provides improvements of 1-4% for similar complexity with JM-1.

Comment on possible problem with overhead bits to signal differential motion vector. Concern was raised about the small gain. Encourage further work.

3.12 SP Frames

JVT-B097 [Sun+] Advanced SP Coding Technique
Instead of quantizing the MC signal, the reconstructed signal is quantized to achieve the synchronization feature. PSNR gains for some experiments indicated the potential of the method. A question was raised with regard to subjective performance. A core experiment to verify and analyse the proposal will be conducted. The CE document will be provided by Nokia and MS. For that, experimental conditions need to be determined: e.g., QP of sequence and QP of switch, frequency of switching. The proposals need to be compared with regards to subjective quality and decoder complexity.

JVT-B055* [Kurceren] New Macroblock modes for SP-frames
In VCEG-O-47 new frame types called, SI-frames have been proposed, to be used in conjunction with SP-frames. The prediction block in SI-frames is formed identically to Intra-frames whereas the rest of the decoding of the SI-frames is similar to SP-frames. The resulting property of the proposed frame is that SI-frame/slice makes use of only spatial prediction and identically reconstructs a corresponding SP-frame/slice, which makes use of motion-compensated prediction. This property provides functionalities in random access, splicing and error resiliency/recovery.

Software is provided that implements SI-frames in TML 9.0 and the corresponding software tools to demonstrate the usage of SI-frames in random access and bitstream switching. The group considers this to be a bug-fix. It is proposed to adopt the feature.

It is requested that for the SP-frame feature to be retained in the JM, a dramatic improvement needs to be made that clarifies it in the document. The proponents are further requested to provide ALL necessary software to use the feature.

3.13 Buffering

JVT-B013 [Viscito] AHG Report: Buffering

JVT-B050* [Ribas+] Including Complexity Constraint in HRD

- (How to measure complexity) It is not clear how complexity can be defined in an implementation independent way. There was no consensus on defining compliance that depends on a limited set of specific implementation designs at some given time (year).
- How to apply it – there is a large mixture of decoder designs. Therefore, even if one can measure it, it is not clear how to apply it in a fashion that supports interoperability

- There was no consensus that VCV should be included in profiles and/or levels.

No consensus in the group on the usefulness of VCV concept. We thank the authors for their work.

JVT-B089* [Hannuksela] Complexity-Constrained Generalized HRD
Proponent Summary: "The document combines the Generalized Hypothetical Reference Decoder design (VCEG-V58) and the Slice-Oriented Hypothetical Reference Decoder design (VCEG-O45)."

VBV parameters should be the normative part of the stream.

Proposal: To allow the VBV parameters (transmission rate, buffer size, initial buffering period) to be sent either in the form of update of the parameter set or the SEI [Supplementary Enhancement Information] packet. (Initial buffering period means VBV delay.) Authors would like to have only one way of sending VBV parameters.

No consensus at this stage. Need to be discussed on Friday (time permitting).

3.14 NAL

JVT-B016 [Lim+] AHG Report: Transport
Review of interim discussions and work on NAL issues, including areas discussed in proposals to this meeting, File Format, and FLC vs. VLC discussion.

Remark: Refer to ToR for scope of work and relationships to other organizations.

JVT-B092* [Etoh+] Proposed Requirements on NAL Specification
Requirements for the NAL Specification are proposed. Requirements to keep functional separation of VCL and NAL, and requirements to have an elementary stream syntax for interoperability testing, are proposed.

Advocates retaining separation of VCL and NAL, with emphasis on efficiency for VCL (only slices and intra support). Remark: Should look at the tools available and then determine how these impact the VCL and NAL. Remark: Note that test conditions can help determine appropriate actions.

Concern of whether RTP design is appropriate for use on IMT-2000. Remark: Wireless can benefit from larger packet sizes. Remark: Compound packet design can help with this.

Very simple elementary stream syntax (one potential NAL) is proposed in the contribution as useful for R&D and interop tests (simpler than use of file format).

Conclusion: Plan is to move toward use of MPEG-4 file format as the defined method (only bitstream switching and below-picture-level addressing (fragments of access units) seems unsupported features in current MPEG-4 file format). Raise as issue with parent bodies.

Principles expressed in this contribution seem to generally be widely supported in the group.

Action items for editing: 1) Add remark to the file format section to clearly indicate its interim nature, 2) avoid/eliminate slight deviations in terminology relative to content of MP4.

JVT-B028* [Stockhammer+] Overview of NAL Concept & VCL/NAL Interface

A summarization of the concept of the H.26L Network Adaptation Layer (NAL) is provided. Purpose:
- reference document on NAL concepts,
- tutorial for those of JVT which are not, or only rudimentarily, familiar with the NAL concept. The document is not a proposal - it describes a concept and the implementation of this concept that were accepted into the H.26L test model."

Group appreciates the tutorial information provided.

JVT-B026* [Wenger+] RTP-NAL and RTP packetization
An information document on the status of the MPEG-4 packetization effort. It includes a) an overview of the IETF standardization process (with references), b) a description of the current MPEG-4 packetization

RFCs and drafts, and c) a short analysis concluding that none of those drafts would be sufficient to implement the functionality of the current H.26L IP/RTP NAL design.

Remark: It is already possible to have a payload format that is optimized for the type of data to be carried (e.g., a format for JVT video).

MultiSL draft format (a working draft in progress in IETF, going to WG last call soon, approx 1yr to get an RFC number "draft standard") is a generic format – not intended to be considered as the best format for every individual type of data. The best format can be designed in a customized fashion.

Remark: A media-specific format should be a subset of the generic MultiSL format. Question: Is this true of AAC format?

Note of last documented design for RTP use of JVT video: JWD or VCEG-N72 (NAL for RTP) and N73 (draft packetization format)

Action: Request information from members/parent-bodies:

- Whether conformance to the draft MPEG-4 "MultiSL" packetization format should be a design constraint on JVT packetization design?
- Regarding definition of "access unit". Is an Access Unit defined as the smallest quantity of data that can be associated with a unique timestamp? Is this definition appropriate for JVT video? (e.g., can a slice be an access unit?)
- Does MultiSL draft support:
    - Support of distinct classes of packets defined such as mode/MV, intra coef, inter coef (e.g., for unequal error protection) within each slice.
    - Compound packets (muxing of multiple slices into one packet)
    - Parts of several pictures in one packet

Where is the best description of the concept of NAL/VCL? Mostly missing from working draft. Request editorial effort to get proper description into the working draft.

JVT-B049* [MacInnis+] Start Codes and Mapping to MPEG-2 Systems

This paper explains some problems that occur when the network or storage layer underlying the JVT video coding layer does not wrap packets precisely around slices, especially the potentially high cost in wasted bit rate. The issue is not unique to MPEG-2 Systems protocols. It explains why unique start codes form the preferred solution in the contributor's opinion, explores how JVT could be made to allow unique start codes in the NAL and the changes to UVLC and CABAC that would be necessary in JVT to make this possible, and proposes a unique start code. A brief overview of how to design an NAL for use with MPEG-2 Systems (Transport and Program Streams) is presented."

Example transport scenarios: Variable-length packets, fixed-length packets (big ones or small ones). Fixed-size transport packets seem to require padding or start codes – and it is potentially very wasteful to use padding. Start codes should be unique. Start codes may also enable leveraging of existing equipment or equipment designs. Propose 32 bit start code prefix.

Could enable UVLC-based start codes with a long codeword devoted to start code and have some way to avoid same-length codewords being used as coefficient data, motion vector data, etc.

CABAC presents its own start code problems (possible emulation prevention per H.263 Annex E).

Potential mapping to MPEG-2 Systems reviewed.

Remark: The specifics of mapping to MPEG-2 systems are out of JVT scope.

Remark: In broadcast, 2^{-24} emulation probabilitymay sound small, but is more often than acceptable during the duration of an entertainment-quality program. Also note that low error rate channels have important high volume applications.

Remark: JPEG-2000 included some use of start codes and study of interaction with arithmetic MQ coder – there could be something to learn from that design.

JVT-B063* [Sullivan] On Random Access and Bitstream Format
Proposal to allow random access point to start with any picture type, not just I pictures. Allow prior references from RA point; providing a means for the decoder to catch up eventually.

Remark: Ideally want I frame to show immediately.

GS: Even with that not a complete decoder reset.

This doesn't *prevent* inclusion of I frame at RA point.

Time tags: Sony has 25-bit IEC method. Supports only PAL and NTSC frame rates.

Need more general scheme, need "true time" (not UTC, but true relative time).

Wants *time based* limit on backward references. A "sync delay" number. If 0, means can immediately decode. If, say 250 ms, means will have to parse/decode 250ms of data stream before able to decode fully. (A 'pre-roll delay' in terms of *picture count* not time.)

This gives the encoder a lot of flexibility to tradeoff random-access responsiveness vs. coding efficiency.

Proposes to permit all 3 of:

- Pre-roll delay (pictures)
- Init delay (time)
- Pre-roll + Init delay (may in some circumstances be faster than either of above)

Q/A: A decoder that starts decoding a given point should init its buffers, start decoding (best effort). Encoder responsibility is then to ensure no further backward references made after promised init point.

Q: Implies some burden on the encoder to keep track of things. Is it practical? Is it back to basics?

Q: Needs more study of effects on loop filtering of intra macroblocks. More study needed.

Q: Item 3: (start with any picture type). How to support fast play/trick play?

A: Encoder can do this if it so chooses. Decoder can choose to shown *only* those RA points attached to I frames (if it wants).

Q: Item 6: (broken link): Want to support "clear" editing process.

Q: Don't like time tag design. (Stephan). Don't think anyone would use it. Maybe supplementary enhancement info (discardable, synced with video).

A: SMTPE time code has problems. Proposal is compatible with SMPTE design, but also references true time.

Q: Init parameters as SEI.

A: Init parameters as SEI. No strong feelings.

Q: Overhead if not SEI?

Remark: Fully support SEI for init parameters. B89 is another proposal re timestamps.

RESULT

(Proposal's numbering):

2. Consider moving RA data down from GOP Layer Will do so as long as it's simple.

3. RA points with any picture type Accepted.

4. Time tag Use simple LSBs of high rate clock. No "fixup" SMPTE stuff, instead simple count, publish conversion to SMTPE in text (accepted pending receipt of working conversion formula).

5. ReInit/ - no for now FFS. REVISIT LOOP FILTER ISSUE IN PLENARY.

6. Broken links FFS.

Proponent (addressing 5): Turn off/on loop filters at boundaries.

General: These people aren't here.

Proponent: Other than that issue, ok.

Discussed the timing indicator part of the contribution.

Similar to MPEG-2 content data description (more compressed form).

Time base should be established at the sequence layer. Parameters described in the document:

Number of basic units in sec in ref clock, base units in picture period, counting type …

At picture level send the frame counter, time stamps (full or partial) …

Open issues: How to represent time is currently an open issue? Use 27 MHz clock?

Recommendations: It is important for JVT to address the issue of how to handle the situation where the capture frame rate is different than the display rate (e.g. film capture and display on TV). Leave this issue open for this meeting. Request further contributions and inputs. Close it in the May meeting.

Set up a separate adhoc group to discuss it (study the issue of timing as it relates to capture time, decoding time and display time, transcoding from MPEG-2) further. (Suggested chairs: Gary S., Sherman Chen).

Think of conversion problem in regard to SMPTE timecode.

JVT-B104* [Wenger] Allow Start Code Emulations
Proposal to explicitly allow start code emulations in such NALs that need start codes, in order not to hurt the coding efficiency for normal symbols (especially in such NALs that do not need start codes). Also advocates disallowing start code emulation prevention bits. It is argued that emulated start codes happen mostly in error prone environments, and there decoders need to be capable to handle incorrect syntax anyway.

JVT-B070* [Suzuki+] NAL for MPEG-2 System
This document describes the contributors view of the requirements of VCL/NAL, and the requirements of NAL for MPEG-2 Systems. It also proposes an NAL for MPEG-2 Systems.

Discusses the role of NAL and relationship to VCL.

Advocating high-level syntax element commonality (not necessarily commonality of exact bit representation). This commonality goal is supported by the group.

Advocating **normative** VCL/NAL boundary specification conformance point to support single decoder for multiple NAL use? No clear disagreement was expressed, however, further study of the issues surrounding conformance point specification is likely to require further study and discussion. Remark: VBV in MPEG-2 video does not include some header data. In H.263, it does. Will need further work to define for JVT. That Issue remains open.

Unique start code advocated for MPEG-2 (not alignment of video packets to PES packets). Advocate start codes in MPEG-2 environment for sequence (NAL specific issue), GOP (open issue), picture (NAL specific issue). Can do by UVLC (not advocated).

Decoder configuration mechanism is necessary (establishment of parameter sets). Agreed.

User data should be supported (Supp Enh Info is in draft, user data is one obvious candidate for use there – Agreed that we need user data). Revisit

Draft SEI syntax has message type ID, then message payload, so it can support user data and other data types. Left to NAL to decide where to carry it. Agreed should be able to synchronize the SEI with appropriate content of VCL stream.

Remark: User data manuf/org. ID codes may be helpful – use of an existing manuf/org ID code registration authority was suggested – noted there is a code used by H.320 and H.323 consisting of 16-bit

country code + 16-bit manuf code. Agreed seems useful to have some reg authority, and perhaps to also allow anonymous user data content.

Add a start code for SEI for bytestream environment. Agreed.

Remark: Consider XML format for SEI..

JVT-B088* [Hannuksela] MPEG-2 Systems NAL
The document proposes an MPEG-2 Transport Packet format where identification of access units is not based on start codes. The proposed format is similar to the planned RTP payload format for JVT. No changes in VCL are required. Moreover, for MPEG-2 Program Streams, the paper proposes a PES packet format that follows the interim JVT file format design.

Advocates considering mapping of VCEG-N72r1 packets (in which first byte is packet type ID) into MPEG-2 Systems transport packets. Assumes start of a new transport packet for each picture if PES header timing information needed. Believes may be possible using proposed techniques to have an MPEG-2 packetization without using start codes. Alternative framing method described. Based on fixed-length packets and ability to synchronize to the transport packet locations.

Remarks: Some of the specifics seem out of the scope of JVT work.

For program stream syntax, advocates using start codes. Proposes considering using file format (or something with similar features) within program stream to carry the data. Assumes use of PES packets containing "boxes" of data. Note: Does not produce unique start codes, so would need an emulation prevention mechanism (e.g., in spirit of JVT-B063) if uniqueness necessary.

Remarks: Some of the specifics seem out of the scope of JVT work.

Action items for JVT: To consider the mapping of the file format features into the bitstream format.

To submit rX version with 2.0 statement.

NAL Subject Area Results:

Adopted Start Code and Bitstream Syntax Method (from JVT-B063):

1) Use MPEG-2/4 start code prefix: 0x00, 0x00, 0x01.
2) Use MPEG-4 visual style of byte alignment stuffing to achieve byte string payload, but flip 0 & 1 values so last byte of payload is never zero (add between 1 and 8 bits of value '1000…').
3) Payload carried after start code prefix, starting with packet type indicator (not emulating MPEG-2 Systems IDs)
4) Types include configuration information, random access point, picture, each slice content type, sequence end?, supplemental enhancement information at each level
5) Encoder emulation prevention method: Search in payload for any string of value 0x00, 0x00, 0x01 or 0x00, 0x00, 0xFF and insert a byte of value 0xFF between the second and third byte. (Average expansion factor is 0.00001 % for random input data)
6) Decoder side: Whenever find 0x00, 0x00, 0xFF, remove the 0xFF. Whenever find 0x00, 0x00, 0x01, declare a next start code detection.
7) At end of payload, remove the last '1' bit and all trailing zeros if any – the remainder is payload prior to alignment padding.

3.15 High-Level Syntax

JVT-B017* [Suzuki] AHG Report: GOP Syntax
This document summaries the discussions of the AHG on GOP syntax.

The report mentions issues such as start codes (addressed elsewhere), FLC versus VLC coding of headers (no specific proposals here, possibly an NAL issue), and general issue of relationship to NAL and VCL. Recommends study of GOP issue in relation to VCEG-O53, VCEG-N52 and VCEG-N72.

JVT-B041* [Hannuksela] Simple Definition of GOP for Random Access
The document proposes signaling of independently decodable GOPs in the slice header structure. It also discusses the relation of independently decodable GOPs and the operation of the multi-picture buffer for

motion compensation. Furthermore, it clarifies how the interim JVT file format (VCEG-O58) supports random access.

Reviews needs of random access: 1) for file storage access, and 2) streaming

Advocates needs of random access as being:

- HRD resync
- Multi-picture buffer reset
- Possible gap in picture numbers (indicating that gap is intentional and not a sign of loss)
- I or SI frame (alternative proposed in JVT-B063)
- Absolute time reference
- File manipulation (cut/paste) capability

Propose to segment the stream into independently decodable GOPs

- First picture number = 0 in each GOB
- Start of GOP flag when picture number is zero

Why not allow start of GOP flag in every picture? Contribution says more bits, potential error resilience impact if loss of first picture of GOP.

Why not allow random access points at P or B pictures? Contribution says consider loop filter impact across MB boder boundaries. Solutions include avoiding border areas for subsequent MC, or a way to turn off the loop filter for some MBs or slices.

Support for random access in file format – file header box identifies random access points, segment box contains absolute time reference, random access positions not aligned with start of file header clump can be identified from slice header structures which are included in the alternative track box.

Intra picture flag in file format is obsolete for random access purposes – not sufficient.

Advocates "complete decoder reset" random access point capability – simple random access capability. No real conflict between this and the JVT-B063 advocacy of other random access capability with eventual perfect recovery. Remark: good to have a sense of the type of reset.

Complete decoder reset enables cut/paste edit capability.

Remark: Should we prohibit wrap-around of picture numbers? Can set modulus in current design. Note that wrap-around is accompanied by a flag indicating wrap.

Remark: Distaste for fields in slice header that are sometimes there and sometimes not – how about taking care of GOP start indication at parameter set level or a different packet/slice type.

On packet type indication: Should try to avoid packet start emulation of MPEG-2 system start code. Perhaps use an emulation prevention byte there or avoid values of MPEG-2 systems start codes for the packet type indicators.

Support for having a complete decoder reset capability.

Requested further detail on exactly what to adopt. (coord with Miska) Doc JVT-B109 was later provided and reviewed.

JVT-B069* [Suzuki+] Group of Pictures for JVT codec
Proponent Summary: "GOP is proposed to support random access. Based on the discussion in the AHG, GOP is proposed incorporating with H.26L design."

Advocates having a random access capability such that: 1) able to quickly and easily identify random access point (at a high level in syntax), 2) absolute timing, 3) identification of whether there is a complete decoder reset at the random access point, 4) random access syntax seems to reside at NAL, with need to synchronize with VCL content.

These four aspects are generally supported.

Syntax proposed: 25-bit IEC 461 timecode (NTSC & PAL), complete decoder reset indicator (does not think all random access points need complete decoder reset), broken link editing indicator.

Proposal expressed that there is a need to address issue of prior references: propose pre-roll count and initialization delay.

Note: HRD/VBV management issue to be addressed at editing/random-access points to enable splicing locations.

Remark: "Parameter set" is currently defined as stream-level information, not synchronous information – we need to be clear on the meaning of terms.

JVT-B063* [Sullivan] On Random Access and Bitstream Format

(Only the random access content part of this proposal is addressed in this section)

notes elsewhere in this report.

JVT-B042* [Hannuksela] Enhanced Concept of GOP
Proponent Summary: "The proposal enhances the concept of independently decodable GOPs presented in JVT-B041 so that disposable chains of pictures (called sub-sequences) can be easily identified and disposed. Such a disposal property may be advantageous in streaming servers, for example."

UEP based on Temporal Scalability. Proposed for "streaming" (IP based distribution) and "storage". Simple addition to the interim file format. MP4 file format may already support it. Needs modification to the decoder multi-picture handling.

Proposal: To modify decoder multi-picture buffer handling so that intentional picture disposal is correctly handled. At the transport level the proposal is to send the bits only as single layer. (Need to study if this impacts non-IP (MPEG-2) based distribution system.).

Note: Verbal report (from Sherman Chen) of Sarnoff IPR.

Possible other IPR (of Teles AG / Wenger) Stephan is requested to formally submit an IPR statement.

3.16 Transform Coding and Quantization

JVT-B008* [Kerofsky] AHG Report: Transform and Quantization
This document summaries the three differing proposal of low complexity transform/quantization design in the view of the AHG chair. No significant coding performance difference has been demonstrated. Differences in transform and quantization implementations are discussed.

No signif. coding perf reported for current quant range.

Luma DC transform changed to Hadamard – Agreed.

2 proposal multiply-free, 1 less efficiently (at least for forward transform).

How was complexity compared? Three suggested primary H/W architectures: gen-purpose dsp, asic, microproc.

16-bit matrix multiply possible in all three proposals.

Shift-add possible with all: harder with TI version, may not be possible for encoder with TI/TML design.

Quantization requirements: TI & TML version does not need normalization on coefficient basis in the absence of a quant matrix (e.g., 3 or more different normalizations), other two do.

Bytes versus shorts for some memory values when looking at amount of memory needed.

Periodic structure to quantization suggested, need QP/M and QP%M for modulus M – used to reduce memory requirement for the multi-norm proposals or to extend quant range (can be used in all three proposals).

Significant interest expressed in finer quantization (Sony Pattaya proposal and others) down e.g., to QP = -12.

Greater bit depth: Extensions in some proposals.

Two main comparison points suggested within current operating parameter range: Multiply free and quantization complexity. Remark: also memory (response: that's a multidimensional issue).

Explorations and discussion areas beyond current operating range:

- Greater bit depth (how to handle dynamic range)
- Finer quantization
- Coarser quantization
- No quantization
- Use with quantization matrices
- Consistency of design with larger block sizes

QP-dependent normalization has an issue with quant matrix use. Asserted that QP-dependence can be eliminated. Should verify that for considering quant matrix use.

JVT-B031* [Zhou] 16-bit based transform and quantization
This document describes the updated TI proposal on 16-bit based transform and quantization. The existing TML transform and quantization is maintained, only change is that the normative scaling factors are introduced to enable 16-bit implementation. The proposal was tested on the H.26L test set and no quality loss w.r.t the 32-bit solution is reported. The extensions to the finer quantization scales, quantization matrices and greater bit-depth are also discussed in this document.

Downscales the forward transform and the inverse transform (with rounding). Scale down B[] reconstruction matrix to 16 bit (if unsigned, otherwise 17). Normative rounding with downshift after each transform dimension. Rounding downshift for inverse quantization.

Remark: Normative intermediate downscaling prohibits nonseparable inverse transform.

Remark: Rounding at each stage of downscaling adds requires extra adds in the matrix multiply implementation.

Remark: Extension to coarser quantization may be a problem. Reply: periodic quantization solution could be adopted to address that if coarser quantization capability is needed.

JVT-B038* [Hallapuro+] Transform and Quantizer - part 1: Basics
Proponent Summary: "Proposal of replacement of the JVT transform with analysis within the current operating range."

Use periodic quantization with 3 normalization factors in transform. Inverse transform coef +1, -1, ½. Forward (example) transform uses +1, -1, 2 factors.

Says intra with very low QP is the critical condition for testing coding efficiency performance, showed test results.

Asserted to be as simple as possible for multiply-free implementation without Hadamard. Remark: Maybe not.

Decoder can be kept fully within 16 bits with 9-bit residual (even for intermediate values).

Only rounding is in final reconstruction.

Memory transform 192 bytes, for quant matrices 288, 144, or 54 bytes.

JVT-B039* [Hallapuro+] Transform and Quantizer - part 2: Extensions
Extension of operating range of JVT-B038 in various ways, Considers 11 and 13 bit residual data, extended quant range, weighting matrices, exact invertibility.

Extension to greater bit depth with no impact on decoder transform or inverse quantization (with 16-bit multplies, but 32-bit memory access). Can add another downshift in encoder for higher bit depth if want 16-bit computation in encoder.

Extended quant range either direction (due to use of QP periodic structure) no change. Results shown for very small QP (down to -12). (Dequant matrices a multiple of 4.) If adopt extended range, then should use same mulitiple of 4 design for both cases.

Doesn't matter how much one extends the quant range due to periodic quantization structure in the proposal.

Extension to exact invertibility – small change to transform, achieves invertibility.

Quantization matrix weighting design shown using Sony method of QP offset.

Cross-verified with independent implementations. Software available pre-Pattaya.

Quant period of 6 used in contribution.

Remark: Example method of encoding shows more than 16-bit downshift. Response: No impact expected if use a smaller shift.

Remark: Example method of encoding shows downscale by 1/16 for 11 & 13 bit extension. Response: No expected impact unless QP extended very small (down to QP=0 no impact).

Perceptual: Demo available – asserted no difference.

Remark: Need to calculate QP%6 and QP/6. Response: One multiply and shift per macroblock when QP changes, or change the period to power of 2, or store the numbers.

JVT-B103* [Liang+] FastVDO Unified 16-Bit Transform/Quant.
Proponent Summary: "This proposal builds on contributions at the Austin and Santa Barbara …"

Described a transform structure, propose adopt the structure, with encoded specification of values of coefficients to use. Several example transforms shown.

Recursive structure shown for larger block-size transform, in which larger block size transform includes the smaller block size transform as an intermediate stage.

Enable three types of transform methods: Direct matrix multiply, multiply-free direct structure, structure with lifting; Remark: Do these get exactly the same results? (e.g., look at need for intermediate rounding for the proposed specific transform). Response there is a need shift in some cases to compensate for scaling.

Advocates separating design of quantization from design of transform, and focuses on the transform. Periodic quantization, for example, can be applied.

Properties available from the family proposed: Exact invertibility, shift-add implementation capability, low bit expansion, equivalent three types of implementation described above).

Asserted that low bit expansion of proposed design provides advantage at small quant values esp. for greater bit depth due to less need for downward scaling to avoid dynamic range expansion.

Remark: Intra results requested, results provided mixed IPPPPPP.

Remark: Results provided use what software? (not provided yet) Show quality loss in some cases? Response: Also shows quality gain in some cases.

Remark: Software presented use R-D quant for his method, not for others. Uses divisions in quantization.

Remark: Lossless case expands the data considerably. Response: Some transforms in the family have been tested with reasonable lossless compression capability.

Remark: What to use for the structure outside of the transform (quantization and inverse quantization). Need definition of what to be used if to adopt something along these lines – needs further definition.

Proposed transform called X5 was used for the results shown. Propose to adopt X5 p=7/16, u=3/8 as starting point and work on downloadable structure capability and quantization/inverse-quant definition.

Remark: How about adopting the Nokia/MS as part of framework to extend upon? Response: exact inversion and tight bit expansion limits are drawbacks of that transform. Remark: Note norm value and need for rounding in X5 matrix multiply implementation.

Remark: Each transform needs an associated quantization method. So would have a finite number of transforms.

Interest in structure? Remark: Not so sensitive in performance to exactly which numbers used, so is there a use for that kind of flexibility? Response: That's probably mostly true in common conditions at 4x4 size, but if extend to higher bit depth, higher resolution (motivating larger block size), may find a need for more flexibility.

Quality comparison to conventional 8x8: Can get as close to DCT performance as desired, depending on particular choice within family.

For quantization and transform as well, perhaps transmit the parameter values or pick a finite set and select among them.

Remark: Mostly an extension beyond current scope proposal rather than a replacement within current scope proposal. Response: Advocate create an activity to study the extension need beyond current range and study transform needs in that extended scope range.

Transform Subject Summary and Results:

Consideration aspects for *current operating range*:

Complexity: Not significant difference, except for rounding requirements (TI has more rounding) and shift-add implementation (TI less friendly), and need for same-result matrix multiply versus shift-add (X5 doesn't appear to satisfy). Some concern over complexity of using a quantization weighting matrix table with Nokia/MS & X5.

Quality: No difference demonstrated. Follow up: someone should confirm the asserted result.

Testing: Less results presented for X5, results for X5 somewhat inconsistent up to 0.5 dB at high rates. Somewhat extra for Nokia/MS.

Verification: Cross-verification of Nokia/MS based on independent implementations.

S/W Availability: TI last week, Nokia/MS pre-Pattaya, X5 not yet (soon).

Consideration aspects for *extensions of operating range*:

- Greater bit depth (how to handle dynamic range): Family approach has less dynamic range problem, Nokia/MS has same inverse transform for 16-bit multiply with 32-bit memory access with greater bit depth – but would need alteration of forward transform, TI don't know
- Finer quantization: Nokia/MS tested to -12, no obvious problem
- Coarser quantization: TI would need some design change
- No quantization: Separate design
- Use with quantization matrices: All compatible with that
- Consistency of design with larger block sizes: Family approach offers extension – and Nokia/MS is part of the family.

Conclusion: Based on the overall considerations described above and the current state of maturity and the desire to make a decision as we are approach CD state, the JVT-B038 proposal has been recommended for adoption. We believe that future proposals of extensions should be able to build upon this design.

Further study of the family approach proposed in JVT-B103 is considered not incompatible with this decision, and further such investigation is recommended.

Group recommends consideration of adjustment of the periodic structure to a period of 8 rather than 6 in the course of ad-hoc interim work (absent any difficulties found in doing so) and extending the range of step size values by one or two factors of two in each direction and adjusting common conditions so that the fidelity range of common conditions remains essentially the same.

JVT-B051 [Yamada+] Improved transform coding for inter-frame

In this contribution, the current 4x4 integer DCT is replaced by 4x4 Hadamard Transform in inter-frame, and shows its advantage considering the balance between complexity reduction and coding efficiency.

Hadamard method tested for inter frames as a complexity reduction.

TML9.0 software, common conditions (one reference picture), little performance penalty (usu. within 0.1 dB, sometimes better).

What about I MBs in P pictures? Used DCT.

Propose a Hadamard versus DCT switching flag in high-level syntax, possibly even at macroblock level.

Open issues noted: How to perform switching (sequence, picture, slice, MB?), criteria, efficiency with other prediction tools (1/8 pel, multi-frame, B-picture) which improve prediction, vlc optimization and scanning order. Not yet results of what switching flag would achieve.

Remarks: For intra there would be a quality penalty, potential for chroma and other visual problems. Noted that this may work well mostly because of good prediction. Subjective quality is a key open question.

Note that this is also a special case of the family proposal, and that selection criteria for when to use different transform.

Note to self: "X2"

How was quantization handled (normalization issue)?

How much complexity reduction relative to JVT-B038?

Recommend further investigation in transform AHG. Do we need multiple transforms, how do we switch, what is the complexity penalty, etc.?

JVT-B067* [Suzuki+] Quantization Tools for High Quality Video

A Weighting Matrix and extension of QP range were proposed in Pattaya. This document proposes how to incorporate with low complexity transform. The preliminary results of VCEG-O52 are reported. It also reports a problem of TML software for high quality video."

Weighting matrix using QP offset approach within existing A[] and B[] matrices with clipping of QP index range. Should not cause overflow problems. DC fidelity emphasized (DC only adjustment, for example). Note that there is interaction between weighting matrix and inverse quantization. (Note: Addressed in JVT-B039) Remark: What about encoder-only techniques in comparison to use of weighting matrix? Suggestion to bring test results especially for high quality video.

Extension of quantizer range to -8 (roughly equivalent to step size of 1 in spatial domain). Use periodic quantization method.

Bug report: TML 9 software crashes for complex scenes with small QP. Response: Probable issue is allocation of buffer in software that assumes no data expansion. Increasing the buffer size fixes the problem. Marta volunteers to fix during integration of JVT-B038.

Prelim weighting matrix results shown for intra image (desire to look at inter later), did not use weighting in second-stage transform. Tried on computer-generated circular zone plate. R-D optimization off. Shows R-D gain around 5% in BDPSNR on that image. Showed one example, some people saw a visual benefit.

Showed computer-generated ramp image at QP=0 with 8 (negative) weighting matrices. All zero, DC using -8, all using -8, etc. Some subjective quality gain asserted as bit precision increased for lower frequency coefficients. Sharpened version of ramp shown emphasizing artifacts resulting from current QP=0. Such artifacts not shown for MPEG-2 (with TM-5 weighting matrix). Different artifacts shown (sharpened) for TML-9 versus VCEG-O25 and different R-D points.

Luma versus chroma balance adjustment not tested yet, but proposed.

Sequence versus still-frame results also would be useful. Would like to see repeat of old H.263+ experiment on encoder thresholding alternative to weighting matrix. Note that weighting seems potentially more useful for larger block sizes than for 4x4.

Group interest expressed in further investigation, reporting results in May, building on the new JVT-B038 design with consideration of the noted issues and interaction with other transform-area investigations.

3.17 Transform Size

JVT-B014* [Benzler] AHG Report: Transform Size

JVT-B053* [Wien] ABT Coding for Higher Resolution Video
Proponent Summary: "In this contribution, the concept on Inter and Intra Adaptive Block Transforms (ABT) is presented. Results are given for interlaced ITU-R 601 source material (the Interlaced CE test set from VCEG-O59), both for Inter-only and Inter&Intra ABT coding. The simulations reveal an improved performance of the proposed scheme, with gains of more than 0.4 dB or 8.5% bitrate savings on average (at high rates, gains of about 1.0dB or 15% bitrate savings can be observed)."

JVT-B065* [Benzler] ABT Results for Interlaced ITU-R 601 Video

Subject Area Results:

Simulations performed according to interlace conditions (CCIR sequences). The sequences are different from the common test set. Coding gains of typically (5-10)% reported. Most gain at high bitrates.

Demonstrations could not be shown due to availability of D1.

Proponent claim visible improvement on CCIR sequences.

Effect of larger transform is mainly on larger picture formats.

For intra coding: added complexity increase due to search of more modes.

Added complexity to decoder: 8 point transforms. More quantization tables. Deblocking is claimed to be less complex.

There is no VLC solution proposed.

Comment: Complexity increase should be quantified.

Proponent consider this a tool for "interlace profile" – or higher profile.

Concern from the group: This is a considerable change

The group find this is a useful tool for high resolution interlace material. Complexity considerations indicate that the tool should not be included in all profiles.

The tool should be considered when defining profiles.

Results have been confirmed by two companies.

Conclusion: Tentatively adopt, provided demonstrations comes out positively.

The AHG work continue.

5-10% gain, esp. at high rate for high-res sequences

The group should consider lower QP range for common test conditions (tests shown here used a low QP).

Need to find a way to do what we now need a D-1 VTR for in terms of demonstration video purposes. A D-1 tape was available for viewing of results at this meeting, but logistical difficulties and scheduling constraints prevented us from being able to view the demo. The proponent indicated that *some* subjective gain would be seen if the demo had been viewed.

Software has been available for some time, cross-verification of results has been performed, including independent implementation with bitstream exchange.

The group recommended considering tentative adoption with follow through on harmonization with the 16-bit architecture 4x4 transform, addition of VLC capability. If adopted, it should be possible to switch off the variable size feature and use only 4x4.

Noted that adoption is tentative, features should be switchable, needs integration with 16-bit transform architecture (e.g., 16 bit 8x8 transform), needs UVLC definition, requiring test results check) for higher than CIF resolution and interlace application – considered non-baseline.

3.18 Robust Transmission

JVT-B024* [Wenger+] Coding Performance not using MV Prediction
The document comments on the efficiency of the in-slice prediction mechanisms of H.26L. At a slice size of one single macroblock, and when deducting the header overhead for the slices, a bit rate overhead of 0.6 to 22 per cent was reported to be observed.

Information document. We thank the authors for the work. Investigation was conducted only for UVLC and not for CABAC. Authors are encouraged to do the tests with CABAC.

JVT-B027* [Horowitz+] Scattered Slices Error Concealment
The document contains a proposal for a new video coding layer based error resilience tool called Scattered Slices. When used over packet lossy links and augmented with an appropriate error concealment technology, it was reported to greatly enhance reproduced picture quality at high packet loss rates, with a small amount of side information. In Scattered Slices, the macroblock ordering in the picture differs from that found in regular raster-scan ordered slices. The penalty for coding macroblocks in an order different from raster-scan order is reported to be less efficient entropy coding, because the in-picture prediction mechanisms (in particular the motion vector and intra-pixel prediction) will, in general, not work as efficiently. However, the total overhead incurred by not taking advantage of in-picture prediction is reported to be normally less than 10% (see JVT-B024 for details).

Demo was shown in H.263 environment. Suggested for the profiles related to the error prone environment with error >= 3% (packet loss rates). Causes <= 10% coding efficiency loss that may be acceptable when there are very few I macroblocks. No hard numbers were presented with regard to performance in JVT environment and based on accepted common condition. Need to compare it with other algorithms under consideration. Group would like to see the results of experiments against JVT common conditions before accepting the results.

Several open issues:

In slice prediction, variable slice size, slice size adaptation to MTU size are open issues.

Use is targeted towards CIF or QCIF resolution. Not recommended for sub-QCIF.

Encourage continuing the work further as part of the Robustness adhoc group. Define core experiment [Stephan W., Thomas S., C.W. Kim, Miska H., Kang, R. Sjoberg]

JVT-B040* [Wang+] Results of Core Exp. on Sub-Picture Coding
The document shows results for the sub-picture coding core experiment (VCEG-O57). The sub-picture coding method allows segmentation of the image to rectangular foreground sub-pictures and to a background sub-picture. The method is reported used to improve error resiliency especially when applied with unequal error protection. The simulation results are reported to show that sub-picture coding outperforms conventional TML in a multicast streaming environment.

Pattaya decision was to do some core experiments [VCEG-O57]. Objective and subjective results (demo of Coastguard, Foreman) were shown. Intra MB update rates will affect the results. Not done yet.

Carphone showed some visible improvement. For Coastguard and Foreman the improvements were not significant. The expected improvement in the sharpness was not perceptible, perhaps due to the quality of the projector. Artifacts at the border of Foreground and background were visible and were somewhat annoying.

Difficult to find Region of Interest. Encoders will be more complex. How to find ROI is not very clear. No consensus that it will be helpful for error resilience.

JVT-B082* [Kim+] Brief Result Core Exp. VCEG-O57 (Sub-Pic)
Report of results of core-experiment of VCEG-O57. Objective results are included and subjective quality was demonstrated.

Demo was shown of Silent sequence with 3% Packet loss. There was a visible improvement in the sharpness of the foreground. Artifacts in the background persisted longer than those in the foreground. Technology needs improvement at the sub-picture boundaries and loss in the background.

JVT-B086* [Kim+] Results on core Exp. of Sub-Picture Coding
Proponent Summary: "The result of core experiment of VCEG-O46, "New image segmentation method" as described in VCEG-O57, "Core experiment description of sub-picture coding"."

Similar comments as for B082. No demo.

JVT-B087* [Kim+] Selection of QP for Sub-Picture Coding
Proponent Summary: "A QP selection method was provided in VCEG-O57 to remove the visible boundary between the foreground and background sub-picture. We propose a more generalized QP selection method to remove the boundary effect while the efficiency of the sub-picture coding technique is maintained. We propose a slice header syntax for sub-picture coding which is the same as proposed in VCEG-O46 except that the background QP is added for the foreground sub-picture slice."

Generalized form. May help in reducing the artifacts in the boundary.

Demo of Carphone for No Error case was shown. Need to do more experiments under packet loss conditions. No conclusions were reached. Authors will provide more results in the next meeting.

[General comment about Sub-Picture Coding Area]

Recommendation: To consider sub-picture technology for Version 2 and not for Version 1 for broader classes of applications and other technologies. [Authors and S.W. Kim objected]

JVT-B102* [Stockhammer+] Error Robust Macroblock Mode and Reference Frame Sel:
Proponent Summary: "This proposal contains two encoder test model extensions to increase the error resilience in combination with multiple reference frames. The first part restricts the selectable reference frames in the rate-distortion optimized reference frame and macroblock selection such that no pixels are used for prediction which have been intra refreshed for error resilience reasons later. In addition the rate-'expected-decoder-distortion' optimized macroblock mode selection presented in VCEG-N50 is extended such that the reference frame is included in the optimization process. Results will be presented based on the Internet test conditions."

Demo was shown for Tempete with significant improvement in error resilience. Does not need any change in the syntax and decoder. Reference frame restriction may also be useful for "dirty" random access.

Proposal: Add the option in the reference encoder to allow the restriction on the selection of reference frames on MB basis and channel optimized combined MB mode & reference frame in the test model encoder. [agreed]

Software implementation is expected to be completed as soon as time slot is made available by the software coordinator.

Continue to do more experiments as part of Robustness ad hoc group.

JVT-B095* [Zhou+] Error Detection Schemes Using Fragile Watermark
Proponent Summary: "A set of error detection schemes using fragile watermark for hybrid codec based video communication were proposed. It improves the error detection rate and error correct detection rate dramatically. To take the advantages of the watermark, a standardization of the schemes should be taken into account."

Loss of PSNR 0.3 to 0.65 dB. Error detection rate from 30 to 60%. Watermark description is missing.

Conclusion: Need to first establish the requirement that we need a new scheme for error detection before we consider it further.

JVT-B076 [Kikuchi+] New picture type for error recovery
Proponent Summary: "We proposed a new frame type called "R-picture" in order for quickly recovery from error (e.g. packet loss). Motion compensated reference frame is intra coded in R-picture and multiplexed into the stream. The server sends R-pictures only when error happens; the codestream sent to the client for error free case stays the same as the normal codestream. Therefore, the size of the sent code-stream does not increase, unlike periodical intra refresh. This scheme is applicable for streaming service

since it does not use a feedback to the encoder. In addition, combination with SP pictures enables perfect drift-free reconstruction."

New Idea. It is applicable only in the system with back channel. Can only be used with SP pictures? More experiments need to be done to clearly show the impact of R-pictures? What is the impact of SI pictures? Authors are requested to compare their approach with other previous investigations done by VCEG (VCEG-M38).

3.19 Interlaced Coding & Progressive/Interlace Interaction

JVT-B010* [Borgwardt] AHG Report: Interlace
Tempete, Mobil favored Frame Coding – In general, pictures with low motion favored frame coding

Bus, Football, Canoa and Rugby favored Field Coding – In general, fast action pictures favored field coding

Conclusions of adhoc group

- Frame or Field coding alone is not a good solution. We need to adapt.
- Need to adapt at least at the picture level

Recommendations of adhoc group

- Adopt option E (picture level adapting) in the standard in this meeting

Issues left open by the adhoc group

Direct mode – Details about block selection need to be made [tentatively settled]

Macroblock level coding – details and CABAC issues

JVT-B020* [Winger+] Interlaced Field Coding Core Experiment A/B

[Meeting Notes]

Experiments done for the core experiment A and B.

Examples:

Rugby, field coding gives up to 2.5 dB (20 to 40%) gain over frame coding.

Tempete, frame coding gives upto about 1 to 3 dB (50 to 75%) gain over field coding.

JVT-B071* [Wang+] Adaptive frame/field coding for JVT video
Proponent Summary: "This document presents the computer simulation results for core experiments A (frame coding), B (field coding) and E (picture level adaptive coding) for interlace coding. Simulations were carried out for various interlace video sequences, including the six common sequences for interlace testing. The simulation results demonstrated the advantages of adaptive coding over frame or field coding. It is recommended that adaptive coding be adopted in JVT."

Core experiment A, B and E were done and reported

Experiment A and B, the results agree with those of JVT-B020

In addition to the sequences described in the core experiments, Coastguard sequence was also used.

Tempete, Coastguard and Mobile & Calendar favoured frame coding. Example: For Tempete frame coding gives the gain of upto 2+ dB over field coding

Rugby, football (other fat motion sequences) favoured field coding. Example: Rugby field coding gains up to 2.5 dB over frame coding

Picture level adaptive coding adapts to the better of the two modes and sometimes little better.

Direct mode – useful mode for field pictures. Fig 3. illustrates a way of doing it when the vectors are taken from the same parity fields, i.e. by scaling up the short vector.

Another approach is to take the vector from the nearest field and scale down a long vector. It is not a complexity issue. Do not know if this will give any performance gain (VCEG-N084, VCEG-O040). No experiments were done or reported.

Group's recommendation about Direct mode - Relatively small issue. Adopt the method described in B071. Request results of comparison the methods in B071 and VCEG-N084/VCEG-O040 and conclude in May meeting whether the decision needs to be modified.

JVT-B068* [Sato+] New Interlaced Coding Tools
Proponent Summary: "New Tools to support interlaced video are proposed in this document. New scanning and MC interpolation filter are proposed to improve coding efficiency."

New Scanning Proposal - Applicable to the frame coded sequences. Gives small gain for frame picture types:

All – 1.7%, IP 1.8%, IBBP 1.3% (Tempete). Small coding gain. IP considerations not known. Need to demonstrate larger gain for acceptance.

MC Interpolation filter – Identified a good issue related to chroma phase shifting, has good potential to reduce visible distortion in chroma. Need to do core experiment to come to a conclusion.

Recommendation – define the core experiment and conduct is as a part of the core experiments in the Interlaced AHG. (Yagasaki-san, Ulrich and K. Y. Yoo will define the core experiments). Authors are encouraged to bring a demonstration of the results, as they may be more visible than what SNR numbers can tell.

JVT-B106 [Wang+] MB level adaptive frame/field coding
Proponent Summary: "This document describes MB-level adaptive frame/field coding for interlaced video materials, and presents the performance comparisons with frame, field and picture-level adaptive frame/field coding. MB-level adaptive frame/field coding is aligned with core experiment C for interlace testing. MB-level adaptive coding provides additional gain over picture-level adaptive coding."

Shows good potential for the sequences where the pictures have mixed motion types (large and small) with-in a picture. Need to get more test sequences with this characteristics – a suggestion was to use Akiyo with Crowd.

Need to further define the interaction with deblocking filter.

Recommendation – Continue to do the core experiments to gather more evidence with regard to the gain, in order to justify the added complexity. Investigate how it interacts with CABAC.

Overall Recommendations of the session

- Adopt frame/field adaptation at the picture level (both text and software)
- Continue to do the core experiments for MB level adaptation for possible inclusion in May meeting if results are positive.

JVT-B048* [Chen+] Supporting Film Mode in JVT Codec
Proponent Summary: "This document describes an approach to deal with coding of film contents vs. 3:2 pull-down operations."

Issue – Film is at 24 frames/sec. Many displays run at ~ 60 fields/sec rate. How to match 24 frames/sec to the display rate that could be different? MPEG-2 inserts two flags: top_field_first and repeat_first_field at the picture layer.

Proposal:

- Define frame_rate and add a frame_rate flag at the sequence layer. Discussion – is it better to define Delta T (picture clock frequency)? Need to have contribution on it with details. (see 063)

- Define display_rate information in Sequence layer.

- Add two flags in "Picture display Extension" film_mode_flag and film_mode_state (similar to MPEG-2).

Whatever solution we come up, we need to also take into consideration the issue of Transcoding from MPEG-2 to JVT

(a general comment - Need to make distinction between the display and decoding process. Display process is out of JVT standard)

GJS comments: 3:2 pattern (one interpretation's "A" is contribution's "D")

Ajay: What about dangling fields?

Arturo: Think about trick modes

Peter: We need to solve the MPEG-2 to JVT conversion problem.

Gary: How about "I don't know" flags? Any unambiguous indication coming out of the encoder is not a solution for the problem of encoders not knowing the timing of the input data.

Note that repeat first field is often used in practice as a compression feature.

JVT-B098 [Tourapis+] Direct Prediction in Interlaced Coding

Withdrawn.

3.20 Profiles and Levels

JVT-B007 [Lindbergh] AHG Report: Profiles & Applications
Proponent Summary: "This document is based on comments on VCEG-014 at the last meeting (described in VCEG-O07), further discussions on the JVT reflector in the Profiles & Apps AdHoc, and private comments. This revised proposal attempts to address all the comments that were made, and proposes solutions for some of the issues that were brought up."

JVT-B025* [Wenger+] On Equivalence of BERs and Packet Loss Rates
Proponent Summary: "This document proposes a (better) definition "Error Limits" section of a profile/level framework document, as made available in the Pattaya ad hoc report on Profile and Level Definitions. It tries to lay common ground for the discussion of error resilience features in the profile/level contents. Theoretical thoughts lead to formula for the conversion of packet loss rates to bit error rates. A definition for the Error Limit column of the table in VCEG-O07 is developed. Finally, higher packet lossy error rates are proposed - however, the bit error rates would have to be reduced according to the formula."

Tentatively accept the general outline at high level for the discussion related to error prone environment. Request feedback from experts about specific numbers / details provided in the document. To be considered for inclusion as informative part of WD-2 related to Profile and Level section's error prone environment related part.

Stephan, will upload a new version with some minor changes with clarification on bullet 2 of section 5.

JVT-B023* [Koenen] Remarks on Profiling

JVT-B035* [Lindbergh] Updated Profile Framework for JVT codec

3.21 Performance Evaluation

JVT-B060* [Boyce] Coding Efficiency of Various #s of Ref. Frames
This contribution provides experimental results about the coding efficiency of various numbers of

reference frames (1, 2, 3, and 5 frames) with quarter and eighth pel motion vector resolution in TML 9.0. This contribution is informational to assist in defining profiles and levels.

Suggestion to include higher resolution sequences. Re-emphasize plan for Profile AHG.

Complexity

JVT-B012* [Horowitz] AHG Report: Complexity
This report covers H.26L complexity reduction work presented in and performed since the Pattaya meeting. It contains a summary of contributions related to H.26L complexity presented at the JVT meeting last December in Pattaya. Further, it contains a summary of complexity related activities occurring since Pattaya including an announcement of complexity related contributions registered for the Geneva meeting.

JVT-B030* [Zhou] Simplification of H.26L baseline coding tools
In this document the coding efficiency of major H.26L baseline coding tools, namely 16x16 intra prediction, multiple reference frame prediction, motion vectors of size below 8x8, hadamard transform in ME and intra prediction decision, RD-optimization, cost-function based zero block decision, and loop filter, is evaluated.

The sequences chosen differ from the JVT test set and it was pointed out that most of them show low activity. A remark was made regarding the use of VQEG sequences for testing. All experiments are done without the loop-filter. When the loop filter is included, for this test set, at QP=20, it is shown that the minimum gain of H.26L over MPEG-4 SP is min 20 %, max 51.5%, and 35% on average.

A remark was made regarding the distinction between encoder and decoder complexity.

3.22 Intra Coding

JVT-B080* [Conklin] More Results on New Intra Prediction Modes
Proponent Summary: "This information document provides additional results in support of the proposal made in VCEG-N54. It shows a compression gain of 10% is achievable with minor changes to the current Intra macroblock coding method by presenting results obtained by coding all frames of the test sequences as keyframes. Visual results are also contained in the document."

Informational relating to VCEG-N54: 3 new directional modes, more predictor pixels left and above. Inverted foreman coding in Santa Barbara. From 1% to 8.3% delta bits improvement on intra.

No penalty from this if simply tack the new methods on to end of the old list (not exactly how it is proposed). Decoder complexity not signif impact.

Software available since approx Santa Barbara.

r1 doc expected.

Adopt.

3.23 Encoding

JVT-B022* [Hong & Oh] Range decision for motion est. of VCEG-N33
This contribution reduces motion estimation complexity using local motion vector statistics. Since it is well-known that motion vector of a block is highly correlated to its neighboring blocks, the motion vectors of the neighboring blocks are used to determine the search range of the block. Experimental results reported that an average of 50% saving of MV encoding time was obtained without loss in visual quality.

Proposed adoption as in JVT-B022.

3.24 Fine-Grain Scalability

JVT-B094* [Cheong+] Water Ring Scan for H.26L-based FGS
Proponent Summary: "In this document, we will notify the limitation of the current FGS if it is directly

adopted into H.26L based scalable coding methodology. Also, we will introduce the water ring scan method as a potential technology that can improve subjective picture quality of a decoded scalable video."

Water ring scan method was presented. It showed some improvements over the current methods being discussed. Authors will further continue the work and bring the results in the next meeting. Continue the discussion in the MPEG adhoc group. We thank the authors for their work. No further action needed at this stage.

4. Meeting Results and Conclusions

4.1 JVT Ad Hoc Committees Established

1. Editing and S/W [Thomas Wiegand, with designated area editing assistance team]
2. Project Management [Gary Sullivan]
3. Deblocking [Peter List] invest. interlace relationship, turning off on slice boundaries
4. VLC [Gisle Bjontegaard]
5. CABAC [Detlev Marpe]
6. Multiframe Motion Prediction [Martin Schlockermann] Mandate: finalize B-picture syntax taking into account JVT-B057, and other multiframe motion prediction work
7. GMVC/GMC [Chair: Kimata-San, co-Chair: Lainema-San] Mandate: study technique for inclusion employing JVT-B046 and JVT-B019. Studying the use of MB skip at predictor.
8. Additional Transforms and Quantization Methods [Matthias Wien] Mandate: ABT, Transform family approach, quant weighting matrices, hadamard
9. Motion Interpolation [Tom-Ivar Johansen]
10. NAL and High-Level Syntax [Young-Kwon Lim, Thomas Stockhammer, Miska Hannuksela]
11. Complexity [Michael Horowitz]
12. Robustness [Michael Horowitz] – Scattered Slices Error Concealment and Error Robust Macroblock Mode and Reference Frame Selection, define and conduct the core experiments.
13. Interlace [Peter Borgwardt, Limin Wang]– Adaptive MB and Chroma Phase distortion; Complete the core experiments on Adaptive MB; Define and complete the core experiments on Chorma Phase correction
14. Film mode / Timing information [Gary Sullivan, Sherman Chen] Discussion on timing (e.g. capture, presentation, etc) and Film mode
15. Profiles and Level [Dave Lindbergh] Discussion on Profiles and Levels
16. Buffering [Eric Viscito] VBV model discussion

4.2 Summary Report to the Parent Bodies (ISO/IEC MPEG and ITU-T VCEG):

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

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

Approx Date	Auspices	Location	Project Milestone
May 6-10, 2002	JTC1	Fairfax, VA, US	JM3, CD
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 on the weekend prior to the next meeting in Fairfax may be necessary to complete the May workplan.

4.2.2 The JVT adopted a second Joint Working Draft (JWD 2) design [JVT-B118], and a Joint encoding test Model (JM 2) non-normative reference encoder description. This includes the adoption of the following changes to the prior JWD 1 and JM 1. The draft document editor (Thomas Wiegand) is asked to produce the new version of the working draft within four weeks of the end of the meeting, and the software coordinator (Karsten Suehring) is asked to produce the corresponding version of the software by the halfway-point between this meeting and the Fairfax meeting in May.

Normative content changes adopted:

- **Deblocking filter JVT-B011 (general usefulness)**
- **Exp-Golomb VLC JVT-B029 (general usefulness without CABAC)**
- **CACM+ CABAC JVT-B036 (general usefulness with CABAC)**
- **Bitstream NAL structure with start code and emulation prevention part of JVT-B063 (use for bitstream environments)**
- **SI Frames JVT-B055 (use for streaming, random access, error recovery)**
- **Intra Prediction JVT-B080 (put in software as configurable feature, use it in common conditions, seek complexity analysis)**
- **Interlace frame/field switch at picture level from JVT-B071 (with field coding as the candidate baseline interlace-handling design)**
- **MB partition alteration VCEG-O17 (general usefulness)**
- **CABAC eff. improve JVT-B101 (general usefulness with CABAC)**
- **Transform JVT-B038 (general usefulness)**
- **Extension of quant range (general usefulness)**
- **Normative picture number update behavior from JVT-B042 (error resilience)**

Tentative:

- **ABT JVT-B053 (prepare completely final and switchable annex description and reference software as candidate for adoption at next meeting – expected to be adopted as a non-baseline feature if this is achieved)**

Non-normative content:

- **Encoding Motion Search Range JVT-B022 (lower complexity method not for common conditions testing)**
- **Robust reference frame selection JVT-B102**
- **Enhanced GOP concept description in interim file format non-normative appendix JVT-B042**

4.2.3 The JVT notes that our working draft *document* is the draft of our future standard, and that it therefore represents the primary and highest-priority representation of our adopted design. The contributors to that draft are reminded that it must contain a complete and full specification of the decoding process for JVT video at a level of detail sufficient to achieve understanding and interoperable implementation of the design (using only the document as the specification of the necessary technical content).

4.2.4 The JVT expressed an interest in the following topics for future consideration:

- **An interesting proposal (Scene Transitions JVT-B043) was presented that needs further consideration and analysis (including other composition-style features such as picture-in-picture), which could be considered for adoption but we could not reach a consensus on its adoption now.**
- **Consideration of making quantization parameter have a period of 8 and adjustment of common conditions a work item for AHG investigation**

- **Consideration of JVT-B109 which was produced and reviewed as an output document draft for consideration and comment, and for potential adoption into draft at next meeting**

4.2.5 The JVT adopted a plan to define and conduct the following core experiments:

- **JVT-B111 Core Experiment on Scattered Slices**
- **JVT-B112 Core Experiment on SP Pictures**
- **JVT-B115 Core Experiment on Adaptive MV Coding**
- **JVT-B116 Core Experiment on Interlace Chroma Phase Shift**
- **JVT-B117 Core Experiment on Macroblock Adaptive Frame/Field Interlace Coding**

4.2.6 The JVT produced JVT-B108 as its draft framework for Profile & Level definition and requests information and comment from its parent organizations and members on its content. In particular, the JVT requests input on what applications are identified for JVT video, which tools appear to be appropriate for those applications, analysis of the complexity of the JVT codec tools, and suggestions on the best mapping of these needs into appropriate profiles and levels.

4.2.7 The JVT requests information and comment from its parent organizations on the consideration of the sub-picture error resilience tool JVT-B040 as a potential candidate for phase 2 work beyond the currently-defined schedule.

4.2.8 The JVT requests information and comments from its parent organizations on the following issues in regard to RTP payload packetization:

- **Remarks on whether conformance to the draft MPEG-4 "MultiSL" packetization format should be a design constraint on JVT packetization design**
- **Remarks regarding the appropriate definition of an "access unit". Is an Access Unit defined as the smallest quantity of data that can be associated with a unique timestamp? Is this definition appropriate for JVT video? (e.g., can a slice be an access unit?)**
- **Remarks on MultiSL draft support in the following areas:**
    - **Support of distinct classes of packets defined such as mode/MV data, intra coefficient data, and inter coef data packets (e.g., for unequal error protection) within each slice.**
    - **Support for "compound packets" (muxing of multiple slices into one packet)**
    - **Support for placing data for parts of several pictures in one packet**

4.2.9 The JVT expressed its intent to move toward use of the MPEG-4 file format as the defined method for JVT video content storage, and directs its working draft editor to clearly indicate the interim nature of the JVT interim file format design in the working draft and to avoid/eliminate any slight deviations of editorial terminology with regard to the drafted interim file format description. The JVT notes that only bitstream switching and addressing of data below the picture level (addressing of fragments of access units) appear to be features in the JVT interim file format design that may not be supported in the current MPEG-4 file format design. The parent bodies are requested to consider these features of the JVT interim file format design for study as we progress toward use of the MPEG-4 file format.

4.2.10 The JVT thanks the ITU for hosting the JVT at its headquarters facilities in Geneva and thanks Ms. Leslie Jones in particular for her assistance. We are especially grateful for the provision of expanded meeting room allocations as the size of our group and its quantity of work grew beyond our initial expectations.

4.2.11 The Genève meeting of the JVT was closed at 21:35 on February 1, 2002.