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

Document: JVT-C099r1
Filename: JVT-C099r1.doc

Title:	**Main Profile for MPEG-4 AVC and H.264**
Status:	Input Document to JVT
Purpose:	Proposal
Author(s) or Contact(s):	Sherman (Xuemin) Chen, Alexander (Sandy) MacInnis (Broadcom) (with comments from Haoping Yu (Thomson Multimedia), Minhua Zhou (Texas Instruments), Ping Wu (TandbergTV), Dzung Hoang and Eric Viscito (Globespan))
	Tel: Email: schen@broadcom.com
Source:	Broadcom

1. Introduction

As we are making good progress on the technical work of H.26L (also called MPEG-4 AVC, or JVT codec), we would like to start discussions on definition of the main profile (MP).

As stated in the JVT document "Profile and levels framework"[1],

- "Profiles should be defined ONLY when their requirements are well-understood".
- "Profiles should ONLY be defined when there is solid support for the Profile, expressed as deployment commitments from multiple companies".
- "Profiles shall only by defined if adequate Level definitions are available".

Definition of a main profile is well justified because the Digital Broadcast Television industry has

- Well-understood requirements for its enhanced TV services,
- High-volume deployment commitments.

It is expected that the main profile will be very useful in enhanced digital TV services over cable, satellite and DSL, etc. This profile might also be considered for entertainment video streaming applications.

2. Requirements of the Main Profile

The main profile should serve the needs of the Digital Broadcast Television industry. The attributes of a useful profile include:

- Entertainment quality service
  - Video/Audio quality designed for large display viewing in the home
  - Support for high-quality delivery of film content to televisions
  - Accurate and continuous synchronization of video/audio services
  - Support for access points for low delay channel changes
  - Random access to support trick modes and personal video recording
  - Reliability
- High coding performance for entertainment quality video at resolutions higher than HHR (Half Horizontal Resolution) video.
- Balance of coding efficiency and complexity of implementation geared towards inexpensive consumer-level appliances using low-cost decoders with limited CPU power and limited cache/memory.
- High leverage of existing MPEG-2 network delivery infrastructure. It must be suitable for simple mapping to MPEG-2 Transport and Program Streams.
- Deployment to vast numbers of consumer-level receivers (interlaced and progressive displays)
- Deployment under well-defined network conditions, i.e., low error rates and low jitter conditions (cable and satellite networks).

It is expected that this profile operates at a bitrate over 300kbps, a frame rate range up to 60fps, and frame resolutions of up to 8100 MBs/frame, i.e. high definition. A definition of Levels within this profile constrains the bitrate range, frame rate range, and frame resolutions into sub-ranges.

3. Definition of video coding tools for the main profile

The video coding tools to be included in this profile are suggested in the following table [2]:

Video Tools	Main Profiles	Restrictions
I-picture	X	
P-picture	X	
B-picture	X	Disposable B pictures only, Motion vector block sizes are limited. (see below).
Motion vector block sizes of 16x16, 16x8, 8x16, 8x8, 4x8, 8x4 and 4x4	X	Bi-directional motion compensation is not allowed for MC blocks smaller than 8x8. The direct mode can only be applied to the MBs with their co-located MBs having motion vector block size larger than or equal to 8x8.
Reference frames	X	Limited by the amount of memory in a hypothetical decoder, i.e. 1 or 2 reference frames at max resolution for a level ; More reference frames allowed when resolution drops as long as the total frame buffer size does not exceed the amount required for max resolution.
16 bit transform and quantization	X	Extended quantization range QP within (-12, 43)
Loop filter	X	
1/4 pel motion vector resolution	X	
Interlaced coding tools	X	Frame/field-based only
Intra prediction modes	X	
Exp-Golomb VLC	X	
CABAC	X	Use of CABAC is limited to bitrates less than or equal to one-half of the maximum allowed for a given level
Startcodes (in NAL)	X	Consistent with MPEG-2 video and MPEG-2 Systems
Slice types	X	Include MPEG-2 style slices (type 2) without requiring independence from other slices and resulting loss of coding efficiency.
HRD	X	Consistent with MPEG-2 Systems STD

Table 1. Tools in the TV profile

Excluded tools:

- 1/8 pel motion vector resolution
    - Does not provide substantial benefit
    - refer to experimental results in VCEG-O42, JVT-B060,JVT-C036

- Motion vector blocks at size 4x4 (not all cases allowed)
    - Do not allow bi-directional MVs with 4x4, 4x8 or 8x4. Prefer to limit the number of 4x4 MVs in a picture. One suggestion is that the limit be proportional to the max number of MBs per picture. For example, one could allow up to 25% of MBs to be 4x4 while the rest are 16x16, or any other pattern with the equivalent total number of MVs. In this example, for a level with N MB/s picture max, up to (16*.25 + 1*0.75)*N = 4.75 * N MVs /picture would be allowed. One could choose different parameters, such as allowing 6N or 8N MVs/ picture.
    - Unlimited use of 4x4 MVs does not add significant value vs. reasonable limitations, but it causes significant added complexity.
    - Refer to VCEG-O42 experimental results (overall average 0.869 dB for types 1-7 vs 1 vs 0.699 dB for types 1-4, 0.17 dB delta) (overall average 8.34 % bits for types 1-4 vs. 9.12 % for types 1-7, delta of 0.78%)
    - Refer to JVT-C036 simulation results.

- SP and SI frames
    - Designed for other applications.
    - The added complexity can not be justified by the feature.

- 32 bit transform – ABT
    - Not mature at this time
    - Increase in coding gain vs. complexity of implementation unclear at this time

- JVT-B043 Coding of Scene Transitions
    - Increases decoder complexity (e.g. must decode two frames in one frame time).
    - There are other approaches currently used in the industry to handle scene transitions

- Error-resilience tools (data partitioning, etc.)
    - Existing network resiliency approaches are adequate
    - TV Profile assumes more reliable network delivery than, for instance, wireless

- GMC / GMVC
    - Increased complexity and arguable applicability for all broadcast material

- B frame as a reference for prediction (JVT-B057)

Guide on usage:

- I-pictures and GOP structure. This would help meet the stated requirement to make random access and trick modes work well. The suggested GOP size is at least once per 30 frames, or once per 15.

4. Definition of Levels

Within MP, four levels are defined as specified in Table 2. The level definitions are based on the intended application areas. Further studies may be needed to finalize the numbers in Table 2. Note that the interlace tools don't apply to level L1.

Profile	Level	Maximum Bitrate	Maximum Total Pixels per Second	Maximum MBs per Frame	Maximum HRD Buffer Size (units of 16384 bits)
Main	L1	1Mbps	3,041,280	396 (i.e. 352x288)	62
Main	L2	3Mbps	6,082,560	792 (i.e. 352x576)	184

Main	L3	5Mbps	12,441,600	1620 (i.e. 720x576)	306
Main	L4	20Mbps	62,208,000	8160 (i.e. 1920x1088)	1221

Table 2. Definition of Levels in the Main Profile

All Levels of Main Profile allow a maximum of 60 frame/sec subject to the other constraints of the level, e.g., pixels per second.

The maximum number of bits per macroblock needs to be restricted in order to ensure reliable decoding of worst case streams. The suggested limit is 384 bytes maximum per any one macroblock with the only exception being if the standard includes an escape to encode pixel values directly as PCM with no motion compensation, then the limit is 384 bytes plus the marcoblock header. There should also be a limit of bits per macroblock averaged over a row of macroblocks. The suggested limit is a maximum of 192 bytes per macroblock averaged over any one row of macroblocks.

Reference:

[1] Dave Lindbergh and Ajay Luthra, “Profile and levels framework”, JVT-B108, Feb. 2002.
[2] Thomas Wiegand, Joint Working Draft 2 / Joint Encoding Model 2, JVT-B-118, March 2002.
[3]Le Maguet, TV Profile and Levels, JVT-C036, May 2002.

(Append for Proposal Documents)

JVT Patent Disclosure Form

International Telecommunication Union
Telecommunication Standardization Sector

International Organization for Standardization

International Electrotechnical Commission

Joint Video Coding Experts Group - *Patent Disclosure Form*

(Typically one per contribution and one per Standard | Recommendation)

Please send to:

JVT Rapporteur Gary Sullivan, Microsoft Corp., One Microsoft Way, Bldg. 9, Redmond WA 98052-6399, USA
Email (preferred): Gary.Sullivan@itu.int Fax: +1 425 706 7329 (+1 425 70MSFAX)

This form provides the ITU-T | ISO/IEC Joint Video Coding Experts Group (JVT) with information about the patent status of techniques used in or proposed for incorporation in a Recommendation | Standard. JVT requires that all technical contributions be accompanied with this form. *Anyone* with knowledge of any patent affecting the use of JVT work, of their own or of any other entity ("third parties"), is strongly encouraged to submit this form as well.

This information will be maintained in a "living list" by JVT during the progress of their work, on a best effort basis. If a given technical proposal is not incorporated in a Recommendation | Standard, the relevant patent information will be removed from the "living list". The intent is that the JVT experts should know in advance of any patent issues with particular proposals or techniques, so that these may be addressed well before final approval.

This is not a binding legal document; it is provided to JVT for information only, on a best effort, good faith basis. Please submit corrected or updated forms if your knowledge or situation changes.

This form is *not* a substitute for the *ITU ISO IEC Patent Statement and Licensing Declaration*, which should be submitted by Patent Holders to the ITU TSB Director and ISO Secretary General before final approval.

Submitting Organization or Person:	
Organization name	**Broadcom**
	15435 Innovation Drive, San Diego, CA 92128
Mailing address	
Country	**USA**
Contact person	**Sherman Chen**
Telephone	1-949-585-6185
Fax	
Email	schen@broadcom.com
Place and date of submission	5/1/2002
Relevant Recommendation | Standard and, if applicable, Contribution:	
Name (ex: "JVT")	**H.26L**
Title	**Main Profile for MPEG-4 AVC and H.264**
Contribution number	**JVT-C099r1**

(Form continues on next page)

Disclosure information – Submitting Organization/Person (choose one box)

☒ 2.0 The submitter is not aware of having any granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution.

or,

The submitter (Patent Holder) has granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution. In which case,

☐ 2.1 The Patent Holder is prepared to grant – on the basis of reciprocity for the above Recommendation | Standard – a free license to an unrestricted number of applicants on a worldwide, non-discriminatory basis to manufacture, use and/or sell implementations of the above Recommendation | Standard.

☐ 2.2 The Patent Holder is prepared to grant – on the basis of reciprocity for the above Recommendation | Standard – a license to an unrestricted number of applicants on a worldwide, non-discriminatory basis and on reasonable terms and conditions to manufacture, use and/ or sell implementations of the above Recommendation | Standard.

Such negotiations are left to the parties concerned and are performed outside the ITU | ISO/IEC.

☐ 2.2.1 The same as box 2.2 above, but in addition the Patent Holder is prepared to grant a "royalty-free" license to anyone on condition that all other patent holders do the same.

☐ 2.3 The Patent Holder is unwilling to grant licenses according to the provisions of either 2.1, 2.2, or 2.2.1 above. In this case, the following information must be provided as part of this declaration:

- patent registration/application number;
- an indication of which portions of the Recommendation | Standard are affected.
- a description of the patent claims covering the Recommendation | Standard;

*In the case of any box **other than 2.0** above, please provide the following:*

Patent number(s)/status _____

Inventor(s)/Assignee(s) _____

Relevance to JVT _____

Any other remarks: _____

(please provide attachments if more space is needed)

(form continues on next page)

Third party patent information – fill in based on your best knowledge of relevant patents granted, pending, or planned by other people or by organizations other than your own.

Disclosure information – Third Party Patents (choose one box)

☒ 3.1 The submitter is not aware of any granted, pending, or planned patents *held by third parties* associated with the technical content of the Recommendation | Standard or Contribution.

☐ 3.2 The submitter believes third parties may have granted, pending, or planned patents associated with the technical content of the Recommendation | Standard or Contribution.

For box 3.2, please provide as much information as is known (provide attachments if more space needed) - JVT will attempt to contact third parties to obtain more information:

3rd party name(s) _____

Mailing address _____
Country _____
Contact person _____
Telephone _____
Fax _____
Email _____
Patent number/status _____
Inventor/Assignee _____
Relevance to JVT _____

Any other comments or remarks: