INTERNATIONAL TELECOMMUNICATION UNION

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

H.264
Implementors' Guide

(30 July 2010)

SERIES H: AUDIOVISUAL AND MULTIMEDIA SYSTEMS

Coding of moving video

Implementors Guide for H.264: "Advanced video coding for generic audiovisual services"

Summary

This document is a compilation of reported defects identified in ITU-T Recommendation H.264 (2010-03). It must be read in conjunction with the Recommendations to serve as an additional authoritative source of information for implementers.

This Implementors' Guide contains all updates submitted up to and including those at Study Group 16 meeting, July 2010, in Geneva, and was approved on 30 July 2010.

Document History

Version	Date	Description
1	30 July 2010	Initial version - completed at the ITU-T Study Group 16 meeting, Geneva, 19 July – 30 July 2010

Contact Information

ITU-T Study Group 16 / Rapporteur Question 6/16	Gary J. Sullivan Microsoft USA	Tel: Fax: Email:	+1 425 703-5308 +1 425 706-7329 garysull@microsoft.com
Editor H.264	Heiko Schwarz Fraunhofer HHI Germany	Tel: Fax: Email:	+49 30 31002 226 +49 30 392 7200 heiko.schwarz@hhi.fraunhofer.de

Table of Contents

IMPLEMENTORS' GUIDE FOR ITU-T H.264:
"ADVANCED VIDEO CODING FOR GENERIC AUDIOVISUAL SERVICES"

1   Introduction

This document is a compilation of reported defects identified in ITU-T Recommendation H.264 approved 2010-03. It must be read in conjunction with the Recommendations to serve as an additional authoritative source of information for implementors. The changes, clarifications and corrections defined herein are expected to be included in future versions of H.264.

Upon discovering technical defects with H.264, please provide a written description directly to the editor of the Recommendation with a copy to the Q6/16 Rapporteur. The template for a defect report is located at the end of the Guide. Contact information for these parties is included at the front of the document. Return contact information should also be supplied so a dialogue can be established to resolve the matter and an appropriate reply to the defect report can be conveyed. This defect resolution process is open to any interested party. Formal membership in the ITU is not required to participate in this process.

2   Scope

This guide resolves defects in the following categories:

–   editorial errors

–   technical errors, such as omissions and inconsistencies

–   ambiguities

In addition, the Implementors' Guide may include explanatory text found necessary as a result of interpretation difficulties apparent from the defect reports.

This Guide will not address proposed additions, deletions, or modifications to the Recommendations that are not strictly related to implementation difficulties in the above categories. Proposals for new features should be made in through contributions to the ITU-T.

3   References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Recommendation. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Recommendation are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Recommendation does not give it, as a stand-alone document, the status of a Recommendation

–   ITU-T Recommendation H.264 (2010-03), Advanced video coding for generic audiovisual services

4   Nomenclature

In addition to traditional revision marks, the following marks and symbols are used to indicate to the reader how changes to the text of a Recommendation should be applied:

Symbol	Description
[Begin Correction]	Identifies the start of revision marked text based on extractions from the published Recommendations affected by the correction being described.
[End Correction]	Identifies the end of revision marked text based on extractions from the published Recommendations affected by the correction being described.
...	Indicates that the portion of the Recommendation between the text appearing before and after this symbol has remained unaffected by the correction being described and has been omitted for brevity.
--- *SPECIAL INSTRUCTIONS* --- *{instructions}*	Indicates a set of special editing instructions to be followed.

5 Technical and Editorial Corrections to H.264

5.1 SVC deblocking filter

Description:	The SVC specification in ITU-T Rec. H.264 (2009-03) had a problem affecting conformance of the SVC decoding process to the non-SVC decoding process for the base layer when using (and only when using) the Scalable High profile. The issue (in the SVC deblocking filter specification) could arise if there was a coded transform coefficient level that was non-zero, but was associated with a quantization scaling matrix weight that was so small that the final reconstructed transform coefficient value would end up actually becoming equal to zero. A correction of the text was required, because the problem affected the need for the output of the SVC decoding process to match the output of the non-scalable decoding process for the base layer of an SVC bitstream. The correction that was proposed and agreed during the Study Group 16 meeting, October 2009, in Geneva to address the issue actually affected some cases in which that zero-coefficient problem would not actually arise. In particular, it affected the result of the decoding process for the Scalable Baseline Profile, although the Scalable Baseline profile did not actually suffer from the original problem (because of its lack of support for quantization scaling matrices).

A large-volume implementer of the Scalable Baseline profile then discovered that this modification – which was to repair a problem in the Scalable High profile – had now affected their products that implemented the Scalable Baseline profile; and they submitted VCEG-AN10 and MPEG M17633 to request reconsideration of the change in H.264 (2010-03). They proposed an alternative correction that would be compatible with the large volume of existing Scalable Baseline decoders.

But there was a negative reaction to that proposal from another company working toward a new deployment of products for the Scalable High profile – pleading for the adopted correction to stay as-is because it has become part of new product designs under development since its adoption. |

	The agreed compromise is to modify the correction so that it is profile-specific: Implementers of Scalable High profile can use the adopted correction of H.264 (2010-03) and the prior known large-scale deployment of Scalable Baseline profile can remain unaffected.
Reference:	http://ftp3.itu.int/av-arch/video-site/1004_Dre/VCEG-AN10_r3.doc
	http://www.itu.int/md/T09-SG16-100719-TD-WP3-0132/en

[Begin Correction]

5.1.1 G.8.7.4.3 SVC derivation process for the luma content dependent boundary filtering strength

...

– Otherwise, if any of the following conditions is true, bS is set equal to 2:

– cTrafo[mbAddrP] is equal to T_8x8 and either the array rS_L contains non-zero samples for the 8x8 luma block containing sample p_0 or ((slicedeIdc[mbAddrP] & 127) is equal to 0 and the 8x8 luma transform block coded in sliceP and associated with the 8x8 luma block containing sample p_0 contains non-zero transform coefficient levels),

– cTrafo[mbAddrP] is equal to T_4x4 and either the array rS_L contains non-zero samples for the 4x4 luma block containing sample p_0 or ((slicedeIdc[mbAddrP] & 127) is equal to 0 and the 4x4 luma transform block coded in sliceP and associated with the 4x4 luma block containing sample p_0 contains non-zero transform coefficient levels),

– cTrafo[mbAddrQ] is equal to T_8x8 and either the array rS_L contains non-zero samples for the 8x8 luma block containing sample q_0 or ((slicedeIdc[mbAddrQ] & 127) is equal to 0 and the 8x8 luma transform block coded in sliceQ and associated with the 8x8 luma block containing sample q_0 contains non-zero transform coefficient levels),

– cTrafo[mbAddrQ] is equal to T_4x4 and either the array rS_L contains non-zero samples for the 4x4 luma block containing sample q_0 or ((slicedeIdc[mbAddrQ] & 127) is equal to 0 and the 4x4 luma transform block coded in sliceQ and associated with the 4x4 luma block containing sample q_0 contains non-zero transform coefficient levels).

– Otherwise, if profile_idc is equal to 83 and any of the following conditions is true, bS is set equal to 2:

– cTrafo[mbAddrP] is equal to T_8x8 and the array sTCoeff[mbAddrP] contains non-zero scaled transform coefficient values for the 8x8 luma transform block associated with the 8x8 luma block containing sample p_0,

– cTrafo[mbAddrP] is equal to T_4x4 and the array sTCoeff[mbAddrP] contains non-zero scaled transform coefficient values for the 4x4 luma transform block associated with the 4x4 luma block containing sample p_0,

– cTrafo[mbAddrQ] is equal to T_8x8 and the array sTCoeff[mbAddrQ] contains non-zero scaled transform coefficient values for the 8x8 luma transform block associated with the 8x8 luma block containing sample q_0,

– cTrafo[mbAddrQ] is equal to T_4x4 and the array sTCoeff[mbAddrQ] contains non-zero scaled transform coefficient values for the 4x4 luma transform block associated with the 4x4 luma block containing sample q_0.

…

– Otherwise, bS is set equal to 0.

The variable interProfileConformanceFlag is derived as follows.

– If DQId is greater than 0, interLayerDeblockingFlag is equal to 0, and any of the following conditions is true, interProfileConformanceFlag is set equal to 1:

  – profile_idc is equal to 83 and constraint_set1_flag is equal to 1,

  – profile_idc is equal to 86 and constraint_set0_flag is equal to 1.

– Otherwise, interProfileConformanceFlag is set equal to 0.

When interProfileConformanceFlag is equal to 1 and both mbType[mbAddrP] and mbType[mbAddrQ] specify an Inter macroblock prediction mode, it is a requirement of bitstream conformance that the following constraints are obeyed:

– When cTrafo[mbAddrP] is equal to T_8x8 and the array sTCoeff[mbAddrP] contains at least one non-zero scaled transform coefficient value for the 8x8 luma transform block associated with the 8x8 luma block containing sample p_0, the bitstream shall not contain data that result in an array rS_L for which all sample values are equal to 0 for the 8x8 luma block containing sample p_0.

– When cTrafo[mbAddrP] is equal to T_4x4 and the array sTCoeff[mbAddrP] contains at least one non-zero scaled transform coefficient value for the 4x4 luma transform block associated with the 4x4 luma block containing sample p_0, the bitstream shall not contain data that result in an array rS_L for which all sample values are equal to 0 for the 4x4 luma block containing sample p_0.

– When cTrafo[mbAddrQ] is equal to T_8x8 and the array sTCoeff[mbAddrQ] contains at least one non-zero scaled transform coefficient value for the 8x8 luma transform block associated with the 8x8 luma block containing sample q_0, the bitstream shall not contain data that result in an array rS_L for which all sample values are equal to 0 for the 8x8 luma block containing sample q_0.

– When cTrafo[mbAddrQ] is equal to T_4x4 and the array sTCoeff[mbAddrQ] contains at least one non-zero scaled transform coefficient value for the 4x4 luma transform block associated with the 4x4 luma block containing sample q_0, the bitstream shall not contain data that result in an array rS_L for which all sample values are equal to 0 for the 4x4 luma block containing sample q_0.

[End Correction]

G.7.4.2.1.1 Sequence parameter set data semantics

The semantics specified in subclause 7.4.2.1.1 apply with substituting SVC sequence parameter set for sequence parameter set. Additionally, the following applies.

profile_idc and **level_idc** indicate the profile and level to which the coded video sequence conforms when the SVC sequence parameter set is the active SVC sequence parameter set.

constraint_set0_flag is specified as follows.

– If the sequence parameter set data syntax structure is included in a sequence parameter set RBSP, the semantics specified in subclause 7.4.2.1.1 apply.

– Otherwise (the sequence parameter set data syntax structure is included in a subset sequence parameter set RBSP), constraint_set0_flag equal to 1 specifies that all of the following conditions are true:

   – the coded video sequence obeys all constraints specified in subclause G.10.1.1.

   – the output of the decoding process as specified in subclause G.8 is identical to the output of the decoding process that is obtained when profile_idc would be set equal to 83.

constraint_set0_flag equal to 0 specifies that the coded video sequence may or may not obey all constraints specified in subclause G.10.1.1 and that the output of the decoding process as specified in subclause G.8 may or may not be identical to the output of the decoding process that is obtained when profile_idc would be set equal to 83.

> NOTE 1 – The output of the decoding process may be different, if the array sTCoeff contains non-zero scaled luma transform coefficient values for a transform block of a macroblock that is coded in an Inter macroblock prediction mode, but all reconstructed luma residual samples of the array rS_L that are associated with the transform blocks are equal to 0. In this case, the boundary filter strength that is derived as specified in subclause G.8.7.4.3 can depend on the value of profile_idc.

constraint_set1_flag is specified as follows.

– If the sequence parameter set data syntax structure is included in a sequence parameter set RBSP, the semantics specified in subclause 7.4.2.1.1 apply.

– Otherwise (the sequence parameter set data syntax structure is included in a subset sequence parameter set RBSP), constraint_set1_flag equal to 1 specifies that all of the following conditions are true:

   – the coded video sequence obeys all constraints specified in subclause G.10.1.2.

   – the output of the decoding process as specified in subclause G.8 is identical to the output of the decoding process that is obtained when profile_idc would be set equal to 86.

constraint_set1_flag equal to 0 specifies that the coded video sequence may or may not obey all constraints specified in subclause G.10.1.2 and that the output of the decoding process as specified in subclause G.8 may or may not be identical to the output of the decoding process that is obtained when profile_idc would be set equal to 86.

> NOTE 2 – The output of the decoding process may be different, if the array sTCoeff contains non-zero scaled luma transform coefficient values for a transform block of a macroblock that is coded in an Inter macroblock prediction mode, but all reconstructed luma residual samples of the array rS_L that are associated with the transform blocks are equal to 0. In this case, the boundary filter strength that is derived as specified in subclause G.8.7.4.3 can depend on the value of profile_idc.

[End Correction]

5.2   SPS information in scalability information SEI message

Description:	The scalability information SEI message provides the possibility to signal the sequence parameter sets and subset sequence parameter sets that are referenced by a scalable layer representation. According to the semantics, the corresponding syntax elements num_seq_parameter_set_minus1[ ] and num_subset_seq_parameter_set_minus1[ ] indicate the number of sequence parameter sets minus 1 and subset sequence parameter sets minus 1, respectively. It is however possible that a scalable layer representation does not reference any sequence parameter set (only subset sequence parameter sets) and that a scalable layer representation does not reference any subset sequence parameter set (only sequence parameter sets). The second case always occurs for the base layer. The same issue is present in the specification of the view scalability information SEI. In order to correct this issue, the naming, syntax, and the semantics of the syntax elements num_seq_parameter_set_minus1[ ] and num_subset_seq_parameter_set_minus1[ ] are modified.
Reference:	http://www.itu.int/md/T09-SG16-100719-TD-WP3-0132/en

[Begin Correction]

G.13.1.1   Scalability information SEI message syntax

...

if(parameter_sets_info_present_flag[i]) {		
num_seq_parameter_set_~~minus1~~[i]	5	ue(v)
for(j = 0; j <= num_seq_parameter_set_~~minus1~~[i]; j++)		
seq_parameter_set_id_delta[i][j]	5	ue(v)
num_subset_seq_parameter_set_~~minus1~~[i]	5	ue(v)
for(j = 0; j <= num_subset_seq_parameter_set_~~minus1~~[i]; j++)		
subset_seq_parameter_set_id_delta[i][j]	5	ue(v)
num_pic_parameter_set_minus1[i]	5	ue(v)
for(j = 0; j <= num_pic_parameter_set_minus1[i]; j++)		
pic_parameter_set_id_delta[i][j]	5	ue(v)
} else		
parameter_sets_info_src_layer_id_delta[i]	5	ue(v)

[End Correction]

[Begin Correction]

G.13.2.1   Scalability information SEI message semantics

...

num_seq_parameter_set_~~minus1~~[i **]** ~~plus 1~~ indicates the number of different sequence parameter sets that are referred to by the primary coded VCL NAL units of the current scalable layer

representation. The value of num_seq_parameter_set_minus1[i] shall be in the range of 0 to 32~~1~~, inclusive.

seq_parameter_set_id_delta[i][j] indicates the smallest value of the seq_parameter_set_id of any sequence parameter set required for decoding the representation of the current scalable layer, if j is equal to 0. Otherwise (j is greater than 0), seq_parameter_set_id_delta[i][j] indicates the difference between the value of the seq_parameter_set_id of the j-th required sequence parameter set and the value of the seq_parameter_set_id of the (j − 1)-th required sequence parameter set for decoding the representation of the current scalable layer. The value of seq_parameter_set_id_delta[i][j] shall not be greater than 31. When j is greater than 0, the value of seq_parameter_set_id_delta[i][j] shall not be equal to 0. When parameter_sets_info_present_flag[i] is equal to 1, the primary coded VCL NAL units of the current scalable layer representation shall not refer to any sequence parameter set for which the value of seq_parameter_set_id is not indicated by the syntax elements seq_parameter_set_id_delta[i][j] for the current scalable layer and the syntax elements seq_parameter_set_id_delta[i][j] for the current scalable layer shall not indicate any sequence parameter set that is not referenced in any primary coded VCL NAL unit of the current scalable layer representation.

num_subset_seq_parameter_set_minus1[i] ~~plus 1~~ indicates the number of different subset sequence parameter sets that are referred to by the primary coded VCL NAL units of the current scalable layer representation. The value of num_subset_seq_parameter_set_minus1[i] shall be in the range of 0 to 32~~1~~, inclusive.

seq_parameter_set_id_delta[i][j] indicates the smallest value of the seq_parameter_set_id of any sequence parameter set required for decoding the representation of the current scalable layer, if j is equal to 0. Otherwise (j is greater than 0), seq_parameter_set_id_delta[i][j] indicates the difference between the value of the seq_parameter_set_id of the j-th required sequence parameter set and the value of the seq_parameter_set_id of the (j − 1)-th required sequence parameter set for decoding the representation of the current scalable layer. The value of seq_parameter_set_id_delta[i][j] shall not be greater than 31. When j is greater than 0, the value of seq_parameter_set_id_delta[i][j] shall not be equal to 0. When parameter_sets_info_present_flag[i] is equal to 1, the primary coded VCL NAL units of the current scalable layer representation shall not refer to any sequence parameter set for which the value of seq_parameter_set_id is not indicated by the syntax elements seq_parameter_set_id_delta[i][j] for the current scalable layer and the syntax elements seq_parameter_set_id_delta[i][j] for the current scalable layer shall not indicate any sequence parameter set that is not referenced in any primary coded VCL NAL unit of the current scalable layer representation.

num_pic_parameter_set_minus1[i] plus 1 indicates the number of different picture parameter sets that are referred to by the primary coded VCL NAL units of the current scalable layer representation. The value of num_pic_parameter_set_minus1[i] shall be in the range of 0 to 255, inclusive.

pic_parameter_set_id_delta[i][j] indicates the smallest value of the pic_parameter_set_id of any picture parameter set required for decoding the representation of the current scalable layer, if j is equal to 0. Otherwise (j is greater than 0), pic_parameter_set_id_delta[i][j] indicates the difference between the value of the pic_parameter_set_id of the j-th required picture parameter set and the value of the pic_parameter_set_id of the (j − 1)-the required picture parameter set for decoding the representation of the current scalable layer. The value of pic_parameter_set_id_delta[i][j] shall not be greater than 255. When j is greater than 0, the value of pic_parameter_set_id_delta[i][j] shall not be equal to 0. When

parameter_sets_info_present_flag[i] is equal to 1, the primary coded VCL NAL units of the current scalable layer representation shall not refer to any picture parameter set for which the value of pic_parameter_set_id is not indicated by the syntax elements pic_parameter_set_id_delta[i][j] for the current scalable layer and the syntax elements pic_parameter_set_id_delta[i][j] for the current scalable layer shall not indicate any picture parameter set that is not referenced in any primary coded VCL NAL unit of the current scalable layer representation.

[End Correction]

[Begin Correction]

H.13.1.3　View scalability information SEI message syntax

...

if(parameter_sets_info_present_flag[i]) {		
num_seq_parameter_set~~_minus1~~[i]	5	ue(v)
for(j = 0; j <= num_seq_parameter_set~~_minus1~~[i]; j++)		
seq_parameter_set_id_delta[i][j]	5	ue(v)
num_subset_seq_parameter_set~~_minus1~~[i]	5	ue(v)
for(j = 0; j <= num_subset_seq_parameter_set~~_minus1~~[i]; j++)		
subset_seq_parameter_set_id_delta[i][j]	5	ue(v)
num_pic_parameter_set_minus1[i]	5	ue(v)
for(j = 0; j <= num_pic_parameter_set_minus1[i]; j++)		
pic_parameter_set_id_delta[i][j]	5	ue(v)
} else		
parameter_sets_info_src_op_id[i]	5	ue(v)

[End Correction]

[Begin Correction]

H.13.2.3　View scalability information SEI message semantics

...

num_seq_parameter_set~~_minus1~~[i **]** ~~plus 1~~ specifies the number of different sequence parameter sets that are referred to by the VCL NAL units of the representation of the current operation point. The value of num_seq_parameter_set~~_minus1~~[i] shall be in the range of 0 to 32~~1~~, inclusive.

...

num_subset_seq_parameter_set~~_minus1~~[i **]** ~~plus 1~~ specifies the number of different subset sequence parameter sets that are referred to by the VCL NAL units of the representation of the current operation point. The value of num_subset_seq_parameter_set~~_minus1~~[i] shall be in the range of 0 to 32~~1~~, inclusive.

[End Correction]

5.3   SVC derivation process for prediction weights

Description:	The variables baseLumaLogWD and baseChromaLogWD is equations G-112 and G-115 are not defined. The correct variable names are refLayerLumaLogWD and refLayerChromaLogWD.
Reference:	http://www.itu.int/md/T09-SG16-100719-TD-WP3-0132/en

[Begin Correction]

G.8.4.2.1   SVC derivation process for prediction weights

...

6. The variables $logWD_L$, w_{XL}, o_{XL} are derived by:

$$logWD_L = \underline{refLayer}\text{base}LumaLogWD \qquad\qquad (G\text{-}112)$$
$$w_{XL} = aRefLayerLumaWeightLX[\ refIdxLXWP\ ] \qquad\qquad (G\text{-}113)$$
$$o_{XL} = aRefLayerLumaOffsetLX[\ refIdxLXWP\ ] * (\ 1 << (\ BitDepth_Y - 8\ )\ ) \qquad (G\text{-}114)$$

7. When ChromaArrayType is not equal to 0, the variables $logWD_C$, w_{XC}, o_{XC} (with C being replaced by Cb and Cr and iCbCr = 0 for Cb and iCbCr = 1 for Cr) are derived by:

$$logWD_C = \underline{refLayer}\text{base}ChromaLogWD \qquad\qquad (G\text{-}115)$$
$$w_{XC} = aRefLayerChromaWeightLX[\ refIdxLXWP\ ][\ iCbCr\ ] \qquad\qquad (G\text{-}116)$$
$$o_{XC} = aRefLayerChromaOffsetLX[\ refIdxLXWP\ ][\ iCbCr\ ] * (\ 1 << (\ BitDepth_C - 8\ )\ )\ (G\text{-}117)$$

[End Correction]

5.4   Profile specific level limits for MVC

Description:	Two of the constraints specified in H.10.2.2 (Profile specific level limits) do not represent level-specific constraints – they are actually profile-specific constraints that apply to all levels. These constraints are moved from their current location to subclauses H.10.1.1 and H.10.1.2.
Reference:	http://www.itu.int/md/T09-SG16-100719-TD-WP3-0132/en

[Begin Correction]

H.10.1.2   Stereo High profile

...

–   MVC sequence parameter sets shall have num_views_minus1 less than 2.

– For each access unit, the value of level_idc for all active view MVC sequence parameter set RBSPs shall be the same as the value of level_idc for the active MVC sequence parameter set RBSP.

– The level constraints specified for the Stereo High profile in subclause H.10.2 shall be fulfilled.

[End Correction]

[Begin Correction]

H.10.2.2   Profile specific level limits

a) ~~In bitstreams conforming to the Multiview High profile, MVC sequence parameter sets shall have frame_mbs_only_flag equal to 1 for all levels.~~

a~~b~~) In bitstreams conforming to the Stereo High profile, MVC sequence parameter sets shall have frame_mbs_only_flag equal to 1 for the levels specified in Table A-4.

c) ~~In bitstreams conforming to the Stereo High profile, MVC sequence parameter sets shall have num_views_minus1 less than 2 for all levels.~~

[End Correction]

Annex: ITU-T Rec. H.264 Defect Report Form	

DATE:	
CONTACT INFORMATION	
NAME:	
COMPANY:	
ADDRESS:	
TEL:	
FAX:	
EMAIL:	
AFFECTED RECOMMENDATIONS:	
DESCRIPTION OF PROBLEM:	
SUGGESTIONS FOR RESOLUTION:	

NOTE - Attach additional pages if more space is required than is provided above.

_____