Development of mixed screen content coding technology: SCC-HEVC extensions framework

DRAGORAD MILOVANOVIC, ZORAN BOJKOVIC
University of Belgrade
Studentski Trg 1, 11000 Belgrade
SERBIA
dragoam@gmail.com   http://www.zoranbojkovic.com

Abstract: The status of high efficiency video coding development is reviewed based on public available documents in HEVC version 4 standardization process. Here, the requirements and results of call for proposals on screen content coding (SCC) tools for HEVC are presented. The most significant additional coding tools are highlighted. To achieve high coding efficiency in SCC-HEVC, new coding tools that explore the correlation in screen content are developed: IBC block vector search, inter block search, palette mode, and adaptive colour transform. The performances of the coding tools are approved in verifications tests through cooperation phase of standardization process.

Key-Words: image and video technologies, mixed content coding, standardization process

1 Introduction

HEVC (High Efficiency Video Coding) is the joint video coding standardization project of the ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG). The technical development has been performed by the Joint Collaborative Team on Video Coding (JCT-VC) since 2010. The scope of this group was extended to continue working on Format Range Extensions (RExt) and Screen Content Coding (SCC) as extensions of HEVC, expected to be finalized in early 2016.

The design of HEVC incorporates the latest state-of-the-art technologies and algorithmic advances to address the persistent demand for broader usage of video content, video migration to broadband networks, diversification of mobile devices, ever-higher resolutions for cameras and displays, and increasingly high video quality [1]. HEVC achieved first edition 1E of the specification in January 2013.

Another milestone has been achieved for HEVC by finalizing the Range Extensions FReXt amendment, with technology allowing efficient compression of video content for colour sampling formats beyond 4:2:0 and up to 16 bits of processing precision. In particular, the lossless and near lossless range of visual quality is more efficiently compressed than is possible with the first version 1E technology. The amendment is integrated into a new edition of HEVC to be published as the second 2E version in October 2014.

Further Screen Content Coding Extensions remain in development for completion in early 2016 for video containing rendered graphics, text, or animation as well as (or instead of) camera-captured video scenes. Rec. ITU-T H.265 | ISO/IEC 23008-2 fourth version 4E (High efficiency video coding with extensions for screen content coding 06/2016) refers to the integrated text additionally containing screen content coding extensions, additional supplement enhancement information, and corrections to various minor defects in the prior content of the specification. The third edition published by ISO/IEC as ISO/IEC ISO/IEC 23008-2:2016 corresponds to the fourth version (4E).

Screen content coding is defined as the coding of video containing a significant proportion of rendered (moving or static) graphics, text, or animation rather than, or in addition to, camera-captured video scenes. The different characteristics between screen content and camera-captured content make the encoding of screen content different from camera-captured content. To achieve high coding efficiency for screen content, new coding tools should be considered to explore the correlation in screen content.

In the first part of the paper, the SSC-HEVC standardization framework is reviewed based on the JCT-VC working documents. In the final part, the most significant developed coding tools are highlighted.
2 HEVC FRExt extensions

High Efficiency Video Coding (HEVC) is the current joint video coding standardization project of the ITU-T Video Coding Experts Group (ITU-T Q.6/SG 16) and ISO/IEC Moving Picture Experts Group (ISO/IEC JTC 1/SC 29/WG 11). The Joint Collaborative Team on Video Coding (JCT-VC) was established to work on this project. The scope of this group was extended to continue working on Format Range Extensions (RExt), Scalable HEVC (SHVC) and Screen Content Coding (SCC) as extensions of HEVC.

The main steps of HEVC technical developments are organized in four phases (Table 1):

1E. The HEVC first base specification finalized in 2013
2E. Fidelity range extensions (FRExt), Scalable video coding (SHVC) and Multi-view video coding (MV-HEVC) extensions finalized in 2014
3E. 3D video coding (3D-HEVC) extension finalized in 2015
4E. Screen Content Coding (SCC) extensions will be included in the fourth version of HEVC, which is expected to be finalized in 2016.

The Fidelity Range Extensions (FRExt) expands the range of bit depths and color sampling formats supported by the standard, and include an increased emphasis on high-quality coding, lossless coding, and screen-content coding. The drafted range extensions for HEVC include support for the 4:2:2 and 4:4:4 enhanced chroma sampling structures and sample bit depths beyond 10 bits per sample. Additional areas of work include coding of screen content (graphics and other noncamera-view or mixed content), very high bit-rate and lossless coding, coding of auxiliary pictures (e.g., alpha transparency planes), and direct coding of RGB source content [2,3]. With HM 16.0 reference software the RExt branch was merged into the mainline HM software.

The several relatively small changes to the decoding process have been developed for the 4:4:4 chroma sampling, near-lossless or lossless encoding, and screen content, that improve coding efficiency [2].

Intra-picture block prediction (IBC). Prediction operate by copying blocks of previously decoded regions within the same picture, in a similar manner to how motion compensation operates when referencing other decoded picture.

Smoothing disabling for intra-picture prediction. Encoder disable a smoothing pre-filtering that is otherwise applied to intra-picture spatial prediction signal.

Transform skip mode modification. Applies both to lossy and lossless mode, and include enabling horizontal and vertical DPCM coding modes for residual signals, supports of transform skipping for any block size, and rotation of 4x4 residual signals for more efficient entropy coding.

Cross-component decorrelation methods. Correlation between different color components are typically larger in RGB color representation, compared to YCbCr (where the chroma components are already substantially decorrelated differences relative to the luma). Also, the penalty (in terms of bit rate increase) of not exploiting such correlations is naturally larger in color formats without subsampling of components. Therefore, methods for inter-component prediction are being investigated as possible additional elements to be applied within the encoding/decoding processes.

Improved compression in lossless and near lossless coding. The HEVC first edition already enables lossless compression by skipping the transform, quantization and loop filtering, whereas prediction (motion-compensated or intra-picture) and entropy coding are used mostly as is. Substantially different techniques have been proposed that may lead to additional improvements when coding at very high fidelities.

Whereas the HEVC first edition and its FRExt extensions enable improved compression of screen content by the simple options of the block copying and transform bypass mode, other, more sophisticated methods particularly suitable for SCC coding synthetic image structures (which have characteristics such as sharp edges and repetitive patterns) are under development.

2.1 Improvement of SCC efficiency

Screen content has several new features not previously available in camera-captured content, including [4]:

Sharp content. Compared with camera-captured content, screen content usually has with sharp edges. The sharp edges make the transform, which works well to reduce redundancy for camera-captured content, not suitable for screen content. To help the encode sharp content, transform skip has been designed for screen content.
Large motion. Screen content may contain large motions. For example, when browsing a web page, a large motion exists when scrolling the page. Conventional motion estimation algorithm may not be able to handle large motion well. Thus, new motion estimation algorithms to handle the large motions for screen content may be required.

Unnatural motion. Screen content may contain unnatural motions, such as fading in and fading out. The conventional motion model, which usually only takes translational motions into consideration, may not be able to handle screen content well.

Repeating patterns. Repeating patterns are also very common in screen content. For example, the screen content may contain the same letter many times. The same letters belong to a repeating pattern and using previous letters to predict the current letter may be quite efficient. To utilize the correlation among repeating patterns, Intra Block Copy (IBC) has been developed. With the help of IBC, similar blocks can be predicted and encoded efficiently.

The different characteristics between screen content and camera-captured content make the encoding of screen content different from camera-captured content. To achieve high coding efficiency for screen content, new coding tools should be considered to explore the correlation in screen content.

2.2 SCC competition phase

The Screen Content Coding (SCC) extensions improve compression capability for video containing a significant portion of rendered (moving or static) graphics, text, or animation rather than (or in addition to) camera-captured video scenes. Example applications include wireless displays, remote computer desktop access, and real-time screen sharing for videoconferencing. SCC will be included in the fourth version of HEVC, which is expected to be finalized in February 2016.

The intent of document Requirements for an extension of HEVC for coding of screen content is to provide necessary information that can be used to help select new coding tools and develop new extensions of HEVC targeted at these specific video applications. When compressing these new types of content, e.g. screen content, any new tools in future HEVC extensions should significantly outperform the current HEVC standard or they should provide capabilities not available in those standards. Furthermore, the document includes quality requirements for the given applications so that tests performed in MPEG can be designed to determine whether coding tools developed in MPEG meet those quality requirements.

Document [5] summarizes requests in the following three groups:

Mixed content. In general, this type of video signal contains a mix of camera-captured natural videos, images, and computer-generated graphics and text. The camera-captured natural videos and images constitute a portion or portions of a source video frame and may have been compressed with visible compression artefacts. The camera-captured natural videos and images usually have been chroma-subsampled to YUV 4:2:0 and then chroma-upsampling to YUV 4:4:4, and thus may have significant chroma redundancy. Non-camera-captured medical content with overlaid and/or supplemental computer-generated graphics and text also falls into this category. Requirements for mixed content are as follow.

- Quality: up to visually lossless for all or part of the picture;
- Bitrate: comparable to that supported by HEVC Main 10 4:4:4;
- Configurations: including a “low-delay” configuration;
- Codec complexity: based on the framework of HEVC Main 10 4:4:4, with low extra complexity.

Text and graphics with motion. In general, this type of video signal represents the computer-generated graphics and text commonly seen on computer displays. Typical motion observed in this type of content includes window switching and moving, text scrolling, etc. This type of content usually exhibits clear textures and sharp edges with distinct colours. In addition, the text and graphics material in this group is generally not subject to prior lossy compression or processing. Requirements for text and graphics with motion are as follow.

- Quality: up to visually lossless, mathematically lossless;
- Bitrate: comparable to that supported by HEVC Main 10 4:4:4;
- Configurations: including a low-delay configuration;
- Codec complexity: based on the framework of HEVC Main 10 4:4:4, with low extra complexity.

Animation content. In general, this type of video signal represents the computer-generated animation content commonly seen in gaming and data visualization. Typical characteristics of this type content include fast motion, rotation, smooth shade, 3D effect, highly saturated colours with full resolution, clear textures and sharp edges with distinct colours. In addition, the animation content material in this group is generally not subject to prior lossy compression or processing. Requirements for animation content are as follow.

- Colour sampling format: RGB or YUV 4:4:4;
- Resolution: same as those supported by HEVC Main 10 4:4:4;
- Frame rate: same as those supported by HEVC Main 10 4:4:4;
- Bit-depth: 8 or 10 bits per colour component.
The Joint Call for Proposals for coding of screen content represented the start of standardization process [6]. Because there is evidence that significant improvements in coding efficiency can be obtained by exploiting the characteristics of screen content with novel dedicated coding tools, a CfP is being issued with the target of possibly developing future extensions of the HEVC standard including specific tools for screen content coding.

The proposed technologies are evaluated based upon objective metrics and through subjective testing. Results of these tests are made public, taking into account that no direct identification of any of the proponents will be made (unless it is specifically requested or authorized by a proponent to be explicitly identified).

**Test material.** Below is a list of the 4:4:4 screen content sequences to be used. Both the provided RGB and YCbCr formats of each sequence shall be processed. Subjective testing will be performed on only the 1920x1080 and 1280x720 resolutions and using only the RGB format.

**Parameters and conditions.** The two test conditions are **lossy** (the decoded compressed content is not necessarily numerically identical to the uncompressed content) and **mathematically lossless** (the decoded compressed content is numerically identical to the uncompressed content).

**Coding constraint conditions.** **AI** (All Intra - all pictures are coded as Intra pictures), **LD** (Low delay - the first picture is an Intra picture, and there are no backward references for inter prediction /bi-prediction may be applied, but only without picture reordering/) and **RA** (Random Access - Intra picture every 16, 32, and 64 pictures for 20 fps, 30 fps, and 60 fps sequences, respectively).

**Subjective test procedure.** Subjective testing will be performed only on the bitstreams generated using the AI and LD coding constraints, and only for the 1280x720 and 1920x1080 RGB sequences. Four QP values are associated with each anchor sequence. For subjective testing, submitted bitstreams associated with only the three larger of these four QP values will be used. Thus, subjective testing will be performed on the three sequences comprising the lower-rate subset. **Anchors** have been generated by encoding the source sequences using the software and configuration files described below from the HM-12.1 RExt-5.1 software.

The tests were performed in four Laboratories located in Europe. The report Results of CfP on Screen Content Coding Tools for HEVC summarizes the result of the Joint Call for Proposals for Coding of Screen Content, for possibly extending HEVC by associated tools [7]. Seven responses (Qualcom, NCTU&ITRI, MediaTek, Huawei, Microsoft, Mitsubishi, InterDigital) to the Joint Call for Proposals were received during the April 2014 meetings of ITU-T SG 16/Q.6 (VCEG) and ISO/IEC JTC 1/SC 29/WG 11 (MPEG). As a result, it was concluded that the responses provided justification for launching a new standardization project with the goal of extending HEVC, tentatively to be finalized in late 2015. The technical development are performed by JCT-VC.

The study of tools included in the proposals showed many commonalities in terms of the basic concepts, however with various differences in detailed implementation as well as encoder and decoder complexity. The following categories of tools were identified from the proposals:

- **Block based intra block copy** with various aspects of partitioning, search range, mode switching and displacement vector coding;
- **Line based intra copy**;
- **Palette mode** with various aspects of coding the colour palette itself as well as the indices;
- **String matching based on dictionary coding**;
- **Cross component prediction and adaptive colour transform**;
- **Miscellaneous small modifications of existing tools** (such as de-blocking, RDPCM, etc.);
- **Encoder-only modifications**.

From the results of objective (PSNR vs. bit rate) testing of the submitted proposals, substantial bit rate savings were observed to indicate a potential benefit of dedicated tools for screen content coding. The result of the formal subjective assessment also showed a clear improvement for several proposals in comparison to the anchor. Furthermore, various cases showed a visual quality very close to full transparency – even for lossy coding.

Therefore, after the submissions to the CfP provided evidence that significant improvements in coding efficiency can be obtained by exploiting the characteristics of screen content with novel dedicated coding tools, it was planned to develop an extension of the HEVC standard that provides the following:

- Compression tools with a benefit in compression performance for video with screen content or mixed content characteristics;
- A minimum set of profiles for maximum interoperability between devices in the foreseen application domains is targeted.

This extension are developed by JCT-VC as a twin text specification, extending ISO/IEC 23008-2, with the following tentative timeline: First Test Model 14/04, PDAM 10/2014, DAM 02/2015, FDAM 10/2015.
2.3 SCC collaboration phase

The development of SCC extensions for HEVC are based on set of core experiments (CE) specifies tools under investigation and time line of simulation and cross-check reports. Common test conditions (CTC) for screen content experimentation specifies test scenarios under consideration, test sequences, basic encoder configuration, and objective / subjective evaluation of visual quality (Table 2).

The document Common test conditions for screen content coding defines common test conditions and software reference configurations to be used in the development of HEVC screen content coding (SCC) extension [8]. Common test conditions (CTC) are defined for conducting experiments in an environment where outcome of the experiments can be compared. This document defines two coding modes, three configurations, two encoding formats and four quantization parameter (QP) values.

Table 2. Description of core experiments in SCC-HEVC.

<table>
<thead>
<tr>
<th>CE</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CE1</td>
<td>Color gamut scalability</td>
</tr>
<tr>
<td>CE2</td>
<td>Line-based Intra copy</td>
</tr>
<tr>
<td>CE3</td>
<td>Palette mode</td>
</tr>
<tr>
<td>CE4</td>
<td>String matching for sample coding</td>
</tr>
<tr>
<td>CE5</td>
<td>Inter-component prediction, adaptive color transforms</td>
</tr>
</tbody>
</table>

HEVC Screen content coding test model (SCM) provides an encoder-side description, which serves as a tutorial for the encoding model implemented in the HM reference software [9-14]. The purpose of the document is to share a common understanding of the reference encoding methods supported in the software, in order to facilitate the assessment of the technical impact of the proposed new technologies during the standardization process. The common test conditions and software reference configurations that should be used for experimental work are described in CTC.

- The only changes in SCM1 with respect to the HEVC Range extensions test model TM7 are in block vector search for intra block copy (IBC) and hash-based search method for the inter mode [9-10].
- The main changes from SCM1 to SCM2 are the addition of palette mode, residual adaptive colour transform, modifications to the block vector predictor, and block vector difference coding for IBC [10,11].
- The main new coding tool from SCM2 to SCM3 is adaptive motion vector resolution. Other aspects of the design, such as the hash-based inter search and IBC search, palette coding, and adaptive colour transform, have also been improved [11,12].
- The main changes from SCM3 to SCM4 are unification of intra block copy (IBC) and inter signalling, grouping of bypass bins for palette indices and escape values, and enabling the palette mode for non-4:4:4 chroma formats [12,13].

3 SCC coding tools

The coding tools described in SCM4 test model are implemented in the four point release of the 16th HEVC reference software. For an understanding of the encoder upon which SCM4 is built, it is necessary refer to the HEVC test model 16 update 2 of encoder description and the reference software HEVC-16.4+SCM-4.0 [14]. Reference software is being made available to provide a reference implementation of the HEVC standard being developed by the JCT-VC. One of the main goals of the reference software is to provide a basis upon which to conduct experiments in order to determine which coding tools provide desired coding performance. It is not meant to be a particularly efficient implementation of anything, and one may notice its apparent unsuitability for a particular use. It should not be construed to be a reflection of how complex a production-quality implementation of a future HEVC standard would be.

An outline of developed SCC coding tools is follow.

**IBC block vector search**. In order to evaluate the rate-distortion (RD) cost of using the Intra BC mode, for each CU, block matching (BM) is performed at the encoder to find the optimal block vector. In SCM, first a local area search is performed. This is followed by a search over the entire picture for certain CU sizes.

![Fig. 1](image1.png)

**Fig. 1** Spatial candidates for IBC block vector prediction [13].

**Inter block search**. Compared to the HEVC Range extensions test model 7, SCM modifies the inter block search in two ways. The inter search is modified to adapt to the characteristics commonly found in screen content sequences. Furthermore, the inter block search is extended to the whole picture using hash-based techniques. Hash-based search is applied only to 2N×2N blocks. For one block, two hash values are calculated in a similar way but with different CRC truncated polynomials. The first hash value is used for retrieval and the second hash value is used to exclude some of the hash conflicts.

![Fig. 2](image2.png)

**Fig. 2** Search candidates for the modified method [13].
Palette mode. The basic idea behind a palette mode is that the samples in the CU are represented by a small set of representative colour values. This set is referred to as the palette. It is also possible to indicate a sample that is outside the palette by signalling an escape symbol followed by (possibly quantized) component values. This is illustrated in Figure 3. In this example, the palette size is 4. The first 3 samples use palette entries 2, 0, and 3, respectively, for reconstruction. The blue sample represents an escape symbol. For decoding a palette-coded block, the decoder needs to have the following information: the palette entries and palette indices or escape symbol for each sample (in case of escape symbol, additional component values - possibly quantized). In addition, on the encoder side the appropriate palette to be used with that CU needs to be derived.

Adaptive Colour Transform (ACT). ACT performs in-loop colour space conversion in the prediction residual domain using colour transform matrices based on the YCoCg and YCoCg-R colour spaces. ACT is turned on or off adaptively at the CU level using the flag cu_residual_act_flag. ACT can be combined with Cross Component Prediction (CCP), which is another inter component de-correlation method already supported in HEVC RExt. When both are enabled, ACT is performed after CCP at the decoder, as shown in Figure 4. Care is taken on the encoder side when performing ACT, in order to avoid doubling the encoder complexity by searching all the possible modes twice - in both the original colour space and the converted colour space.

Adaptive motion vector resolution. Adaptive MV resolution allows the MVs of an entire picture to be signalled in either quarter-pel precision (same as HEVC version 1) or integer-pel precision. Hash based motion statistics are kept and checked in order to properly decide the appropriate MV resolution for the current picture without relying on multi-pass encoding.

4 Conclusion
Mixed screen content has several new features not previously available in camera-captured content, including sharp content, large motion, unnatural motion and repeating patterns. The different characteristics between screen content and camera-captured content make the encoding of screen content different from camera-captured content.

To achieve high coding efficiency for SCC-HEVC, new coding tools are considered to explore the correlation in screen content: block based intra block copy (with various aspects of partitioning, search range, mode switching and displacement vector coding), line based intra copy, palette mode (with various aspects of coding the colour palette itself as well as the indices), string matching based on dictionary coding, cross component prediction and adaptive colour transform, miscellaneous small modifications of existing tools (such as de-blocking, RDPCM), encoder-only modifications.

Here, the SCC-HEVC extensions profiles under consideration in fourth edition (4E) are Screen-Extended Main, Screen-Extended Main 10, Screen-Extended Main 4:4:4, Screen-Extended Main 4:4:4 10 [15, 16].

The most significant SCC-HEVC developed additional coding tools are IBC block vector search, inter block search, palette mode, and adaptive colour transform. The performances of the new tools are approved in verifications tests through cooperation phase of standardization process (Table 3, Table 4).

References:

Table 3. FRext-HEVC standardization track.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Draft text</td>
<td>TM1</td>
<td>TM2</td>
<td>TM3</td>
<td>TM4</td>
<td>TM5</td>
<td>TM6</td>
<td>TM7</td>
<td></td>
</tr>
<tr>
<td>Reference software</td>
<td></td>
<td>WD1</td>
<td>WD2</td>
<td>WD3</td>
<td>WD4</td>
<td>WD5</td>
<td>WD6</td>
<td>WD7</td>
</tr>
<tr>
<td>Conformance test</td>
<td></td>
<td></td>
<td></td>
<td>HM12</td>
<td>HM-12.1</td>
<td>HM-13.0</td>
<td>HM-14.0</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. SCC-HEVC standardization track.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Draft text</td>
<td>SCMC1</td>
<td>SCMC2</td>
<td>SCMC3</td>
<td>SCMC4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference software</td>
<td></td>
<td>WD1</td>
<td>WD2</td>
<td>WD3</td>
<td>WD4</td>
<td>HM-14.0</td>
<td>HM-15.0</td>
<td>HM-16.2</td>
<td>HM-16.4</td>
</tr>
</tbody>
</table>


