|  |  |
| --- | --- |
| **ITU – Telecommunications Standardization Sector**STUDY GROUP 21 Question 6**Video Coding Experts Group (VCEG)**76th Meeting: 27 March – 4 April 2025, Virtual | Document VCEG-BX17-v1 |

|  |  |
| --- | --- |
| Question: | 6/21 (VCEG) |
| Source: | **Christof Fersch, Kristofer Kjörling, Janusz Klejsa, Heidi-Maria Lehtonen, Harald Mundt (Dolby Laboratories)** | Email: | christof.fersch@dolby.comheidi-maria.lehtonen@dolby.com  |
| Title: | **Automated expert tuning for H.BWC** |
| Purpose: | Proposal |

|  |  |
| --- | --- |
| **Abstract:** | This document proposes an automated expert tuning for complexity-constrained H.BWC encoders. We provide background and overview of the thinking, provide a starting point for encoder specification to which findings of complexity constrained encoder operation point investigations can be added.  |

##### Background

It is envisioned that the encoder can be configured to operate on heterogeneous devices with varying computational performance, which is facilitated by built-in automation amending the coding toolbox based on signal category, number of channels and sampling frequency subject to a complexity constraint. In a complexity unconstrained scenario, the encoder uses signal-category specific tunings. Once the encoder complexity is constrained, some coding tools can be disabled trading off compression efficiency versus savings in terms of computational cost of encoding. The built-in tuning expert automation enables deployment of an encoder in real-world scenarios and reduces the need for creating case specific tunings. The reasons for such an automation are as follows:

* achieving an optimal trade-off between coding performance and computational complexity requires expert knowledge and it may involve substantial experimental effort;
* the number of possible combinations of signal categories, number of channels, and sampling frequencies, and complexity targets may be excessive;
* the computational complexity of an encoder will be a function of number of channels, and sampling frequency, which can vary substantially even within a specific signal category.

The automatic expert tuning functionality facilitates derivation of encoder tuning from a small set of predefined signal category-based tunings given complexity target and the characteristics of the input signal. This concept is illustrated in Figure 1.



Figure 1. Tuning expert automation of the encoder.

##### Operation of the encoder tuning expert

We illustrate the encoder tuning expert functionality, by providing a hypothetical example for which the complexity target is set to 10% of the encoder complexity associated with the provided complexity unconstrained encoder tunings. The databases used in these examples are the ones listed in the H.BWC Call for Proposals document [1].

For ECG signals, where sampling frequency is greater or equal to 360 Hz, and the number of channels is greater than two, the *default complexity-unconstrained tuning* includes:

* full block matching search (including filtering of the block match candidates);
* full search range for block matching according to default IntraPeriod of 98304;
* enabled use of inter-channel coding;
* enabled DST option (UseDST=1);

while, the *complexity reduced tuning* provided by the tuning expert automation built into the encoder derives the following configuration of tools:

* reduced block matching options by disabling filtering of the block match candidates ( UseBMSigFiltering=0, UseLMSigFiltering=0, UseBMSigMultiHyp=0);
* reduced block matching search by excluding the 2048 block length while keeping block lengths from 1024 down to 128 (LOG2\_MAX\_BLOCK\_SIZE=10, MAX\_SPLIT\_DEPTH=3);
* reduce search range for block matching to 8192 by adjustment of IntraPeriod (IntraPeriod=8192);
* disabled use of inter-channel coding (UseTrafoSignalAdapt=0);
* disabled DST option (UseDST=0).

Figure 2 illustrates the averaged PSNR for the H.BWC baseline and the example fast configuration for the MIT ECG database and Figure 3 shows the corresponding averaged relative encoder runtimes for the baseline and the example fast configuration.

#### A graph with a red line  AI-generated content may be incorrect.

Figure 2. Averaged PSNR for Baseline and example fast configuration for the MIT - ECG data base.

#### A graph of a graph  AI-generated content may be incorrect.

Figure 3. Averaged Relative Encoder Run Times for Baseline and example fast configuration for the MIT - ECG database.

For EEG/EMG signals, where sampling frequency is greater or equal to 256 Hz, and the number of channels is greater than four, *the default complexity-unconstrained* includes:

* full block matching search (including filtering of the block match candidates);
* full search range for block matching according to default IntraPeriod of 98304;
* enabled use of inter-channel coding;
* enabled DST option;
* enabled usage of the Trellis quantizer;

while, the *complexity reduced tuning* provided by the tuning expert automation built into the encoder derives the following configuration of tools:

* disabled block matching (UseBlockMatching=0);
* using only 2048 block length while disabling all the remaining block lengths (MAX\_SPLIT\_DEPTH=0, LOG2\_MAX\_BLOCK\_SIZE=11);
* disabled use of inter-channel coding (UseTrafoSignalAdapt=0);
* disabled DST option (UseDST=0);
* disabled usage of the Trellis quantizer (UseRDOQ=0).

Figure 4 illustrates the averaged PSNR for the H.BWC baseline and the example fast configuration for 32 randomly selected files from the CHBMIT EEG database and Figure 5 shows the corresponding averaged relative encoder runtimes for the baseline and the example fast configuration.

#### A graph with red dots and lines  AI-generated content may be incorrect.

Figure 4. Averaged PSNR for Baseline and example fast configuration for 32 randomly selected files (same for Baseline and fast configuration) from the CHBMIT - EEG database.



Figure 5. Averaged Relative Encoder Run Times for Baseline and example fast configuration for 32 randomly selected files (same for Baseline and fast configuration) from the CHBMIT - EEG database.

Figure 6 illustrates the averaged PSNR for the H.BWC baseline and the example fast configuration for the Ozdemir EMG database and Figure 7 shows the corresponding averaged relative encoder runtimes for the baseline and the example fast configuration.

#### A graph with red lines and black text  AI-generated content may be incorrect.

Figure 6. Averaged PSNR for Baseline and example fast configuration for the Ozdemir – EMG database.



Figure 7. Averaged Relative Encoder Run Times for Baseline and example fast configuration for the Ozdemir – EMG database.

#### Encoder behaviour (specification)

An encoder includes an automated tuning expert facilitating complexity adjustment subject to a provided complexity constraint (see Figure 1).

The encoder tuning expert automation may be implemented by providing an encoder configuration file for a given signal category, where the configuration file enumerates the tools that can be disabled to meet the complexity constraint. The tools are enumerated according to their expected impact on the compression performance for that signal category. During the initialization of the encoder, once the characteristics of the input signal (e.g., sampling frequency, and number of channels) is known, and given the complexity constraint, the run-time complexity of the individual tools may be estimated by evaluating predefined functions of sampling frequency and the number of channels. The coding tools can then be disabled according to their priorities until achieving a configuration of the encoder that satisfies the complexity constraint or resorting to the minimum complexity configuration.



Figure 8. The tuning expert process.

For tools that can be further configured, the list of active tools contains specific configurations that are ordered according to their priorities. For example, for block matching, the list of tools may contain:

BLOCK\_MATCHING: LOG2\_MAX\_BLOCK\_SIZE=11, MAX\_SPLIT\_DEPTH=0;

BLOCK\_MATCHING: LOG2\_MAX\_BLOCK\_SIZE=11, MAX\_SPLIT\_DEPTH=4.

In this case, the first configuration only allows for a block length of 2048, and the second entry contains a configuration including all allowable block lengths. According to this example, the first configuration has a higher priority.

**Conclusion**

We recommend defining the proposed framework in an informative encoder annex of the specification and provide tool prioritization and configurations accordingly as part of the further study of encoder complexity performance trade-off.

**Patent rights declarations(s)**

Dolby Laboratories may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).

##### References

[1] ITU-T, “Call for Proposals on the coding of biomedical waveform data.” ([link](https://www.itu.int/en/ITU-T/studygroups/2022-2024/16/Documents/docs/CfP-H.BWC-TD-PLEN-0286-R1-Clean.pdf))