ISO/IEC JTC 1/SC 29/WG 1
(ITU-T SG16)

Coding of Still Pictures

JBIG	**JPEG**
Joint Bi-level Image Experts Group	Joint Photographic Experts Group

TITLE: High Throughput JPEG 2000 (HTJ2K): Call for Proposals (Draft)

SOURCE: WG1

PROJECT: ISO/IEC 15444 (JPEG 2000)

STATUS: Draft

REQUESTED ACTION: Provide feedback and response

DISTRIBUTION: Public

Contact:
ISO/IEC JTC 1/SC 29/WG 1 Convener – Prof. Touradj Ebrahimi
EPFL/STI/IEL/GR-EB, Station 11, CH-1015 Lausanne, Switzerland
Tel: +41 21 693 2606, Fax: +41 21 693 7600, E-mail: Touradj.Ebrahimi@epfl.ch

ISO/IEC JTC 1/SC29/WG1N75031
75th Meeting, Sydney, Australia, Mar. 26-31, 2017

High Throughput JPEG 2000: Call for Proposals (Draft)

The JPEG Committee has launched the High Throughput JPEG 2000 (HTJ2K) activity, which aims to develop an alternate block coding algorithm that can be used in place of the existing block coding algorithm specified in ISO/IEC 15444-1 (JPEG 2000 Part 1). The objective is to increase throughput of JPEG 2000 while otherwise maintaining its unique combination of features, including minimizing the impact of changes on existing codestream syntax and structure, implementations, workflows and content libraries. The output of HTJ2K activity is intended to be published as Part 15 of the JPEG 2000 family of specifications (ISO/IEC 15444).

The JPEG Committee intends to publish a final Call for Proposals (CfP) following its 76th meeting (17-21 July 2017), with the objective of seeking technologies that fulfil the objectives and scope of the HTJ2K activity.

This document is a draft of the CfP, and is offered for a public review period ending 1 July 2017. Comments are welcome and should be submitted to the contacts listed at Section 2.7.2.

1 Background

1.1 Introduction

The success of JPEG 2000 (ISO/IEC 15444-1) across many high-performance applications (D-Cinema, video post-production, film preservation, medical, satellite, etc.) has been due to its unique combination of features.

A recurring issue in these high-performance applications has been the computational complexity of the JPEG 2000 block coding algorithm. The resulting negative impact on throughput can present an obstacle to:

- applications adopting increasing sampling rates (e.g. higher resolutions, or higher frames rates) and sample depths, especially at high-bit rates.
- implementation on power-constrained platforms.

It is therefore desirable to make incremental improvements to JPEG 2000 to increase its throughput, while maintaining its unique combination of features. This includes minimizing the impact of changes on existing codestream syntax and structure, implementations, workflows and content libraries. In particular, the JPEG 2000 block coding algorithm has a substantial negative impact on throughput, and thus a primary candidate for improvement.

1.2 Scope

The High Throughput JPEG 2000 (HTJ2K) activity aims to develop an alternate block coding algorithm that can be used in place of the existing block coding algorithm[1] specified in ISO/IEC 15444-1.
The alternate block coding algorithm is intended to:

- provide substantially higher throughput than the existing block coding algorithm on a given platform (CPU, GPU, FPGA, ASIC…) without substantially reducing compression performance. Software platforms are of primary interest, although others are also important.
- allow transcoding to/from codestreams generated by the existing block coding algorithm, such that the transcoding is mathematically lossless and computationally efficient.
- improve power efficiency such that higher throughput on a given platform should be equivalent to reduced energy consumption per sample.
- be used within all existing parts of the JPEG 2000 standard, with minimal syntactic and structural impact.

The output of HTJ2K activity is intended to be published as a part of the JPEG 2000 family of specifications (ISO/IEC 15444).

[1] The same block coding algorithm is currently used in all parts of the JPEG 2000 work item, except that some optional modifications are introduced by Part 2 (the ultra-fast mode is a very minor extension) and by Part 10 (JP3D allows for the coding of extended blocks).

1.3 Use Cases

1.3.1 General

Unless specified otherwise, HTJ2K is intended to enable the same use cases as those enabled by JPEG 2000 Part 1.

For instance, HTJ2K applies to the JPEG 2000 use cases involving resolution scalability and spatial region-of-interest accessibility. These use cases are particularly relevant to applications that involve some interactive access to a part of the compressed content.

The following summarizes use cases specifically enabled by HTJ2K, while Annex A discusses selected use cases in more detail.

1.3.2 Transcode from J2K to HTJ2K for Efficient Processing

In some applications, an image or image sequence that has already been encoded using JPEG 2000 Part 1 is transcoded to HTJ2K for storage and processing, benefiting from an increase in throughput. This transcoding can be performed on a separate platform or even on the rendering platform itself; the latter is valuable whenever content may need to be decoded or rendered multiple times, as often happens in interactive image communication and rendering applications based on the mechanisms and protocols defined by ISO/IEC 15444-9 (JPIP).

1.3.3 Transcode from HTJ2K to J2K for Legacy Interchange

In some applications, an image or an image sequence is encoded using HTJ2K for local storage and processing, but transcoded to JPEG 2000 Part 1 for interchange with legacy implementations. Such transcoding can be performed incrementally or on-demand.

1.3.4 Transcode from HTJ2K to J2K for Higher Coding Efficiency or Improved Scalability

Some applications may find it useful to transcode from HTJ2K to JPEG 2000 Part 1 for communication to a remote location, taking advantage of quality/SNR scalability and the high coding efficiency of the block coding algorithm specified in JPEG 2000 Part 1.

1.3.5 HTJ2K Native Applications

Some applications will use HTJ2K for both encoding and decoding, without any intermediate transcoding to/from the original JPEG 2000 block bit-stream representation.

1.4 High-level Technical Requirements

1.4.1 Codesteam Syntax

HTJ2K shall reuse the JPEG 2000 Part 1 codestream syntax, as specified in Annex A of JPEG 2000 Part 1, unless necessary to fulfil other requirements therein.

NOTE: The codestream syntax is expected to be modified to accommodate the changes to the decoding process, as permitted in Section 1.4.2.

1.4.2 Minimal Changes to the J2K Decoding Process

Modifications to the JPEG 2000 Part 1 decoding process shall be limited to the Arithmetic Coding and Coefficient Bit Modelling processes specified in Annex C and D of JPEG 2000 Part 1.

1.4.3 Mathematically Lossless Transcoding

The transcoding between codestreams produced by the existing block coding algorithm and those produced by the HTJ2K block coding algorithm shall be mathematically lossless.

1.4.4 Block-based Decoding

Decoding of one HTJ2K block shall not depend on the decoding of any other HTJ2K block.

NOTE: This enables applications in which processing is performed incrementally or on-demand.

1.4.5 Increase in Decoder Throughput

Over a range of bitrates and on a given platform, the throughput of the HTJ2K block decoder should be on average no less than 10 times greater than the JPEG 2000 Part 1 block decoder of the reference specified in Annex D**Error! Reference source not found.**.

1.4.6 Increase in Encoder Throughput

Over a range of bitrates, on a given platform and with identical decoded image, the throughput of the HTJ2K block encoder should be on average no less than 10 times greater than the JPEG 2000 Part 1 block encoder of the reference specified in Annex D.

NOTE: Achieving identical decoded images is possible since only the Arithmetic Coding and Coefficient Bit Modelling processes from JPEG 2000 Part 1 are modified by HTJ2K.

1.4.7 Coding Efficiency Target

On a given platform and with identical decoded image, the HTJ2K codestream should be on average no more than 10-15% larger than the corresponding JPEG 2000 Part 1 codestream.

1.5 Optional Features

1.5.1 Quality scalability

HTJ2K may preserve quality scalability features of JPEG 2000 Part 1, but is not required to do so.

Note: It is expected that the increase in throughput will result in a loss of coding efficiency and quality/SNR scalability, since the existing block coding algorithm obtains a very high degree of quality scalability through the use of multiple coding passes – a primary source of complexity. While preserving quality/SNR scalability is desirable, it is less so than achieving higher throughput and preserving coding efficiency. Applications requiring quality/SNR scalability can continue using JPEG 2000 Part 1 in its present form.

1.6 Royalty-Free Goal

The royalty-free patent licensing commitments made by contributors to JPEG 2000 Part 1 has arguably been instrumental to its success. JPEG expects that similar commitments would be helpful for the adoption of HTJ2K.

2 Call for Proposals

2.1 Introduction

This CfP invites proponents to submit technology contributions that fulfil the scope, objectives, requirements and use cases therein.

Proponents are reminded that they are expected to contribute to the standardisation process, as described in Section 2.6, and attend meeting and present their findings, as specified in Table 1.

2.2 Submission requirements

A submission shall consist of the elements specified in Annex B.

Elements of a submission are provided according to the timeline specified in Section 2.4.

2.3 Evaluation of proposals

The committee plans to select technologies to be included in the standard based on satisfying the requirements and evaluating the results obtained through the evaluation procedure documented in Annex C.

2.4 Workplan

Table 1 specifies the HTJ2K workplan, including submission deadlines. The technical committee intends to prioritize consideration of those submissions that meet these deadlines.

Table 1.Timeline.

27-31 March 2017	Meeting 75. Draft CfP ready for publication.
1 July 2017	Deadline for feedback on Draft CfP.
17-21 July 2017	Meeting 76. Review feedback from Draft CfP and develop Final CfP.
October 1 2017	Registration of intent. Proponents that intend to provide a complete submission shall post element B.1 of their submission as input document to Meeting 77, and indicate their intent to provide a complete submission and contribute to the standardization process.
23-27 October 2017	Meeting 77. Review of registrations of intents. Proponents are required to present element B.1 of their submission. A teleconference bridge is expected to be available.

1 March 2018	No later than this date, proponents shall post elements B.1 through B.3 of their submission as an input document to Meeting 78.
April 2018	Meeting 78. Evaluation of submissions. Selection of one or more submissions as basis for standardization process. If selected, a proponent shall post elements B.4 through B.5 of their submission as an input document to Meeting 78. Proponents are required to attend this meeting in person.
End of Meeting 78 + 15 days	WD
October 2018	Meeting 79. CD
February 2019	Meeting 80. DIS
June 2019	Meeting 81. IS

The above schedule is subject to change, depending on the nature of proposals that are received and the possible need to integrate or merge elements from separate proposals.

2.5 IPR conditions (ISO/IEC Directives)

Proponents are advised that this call is being made in the framework and subject to the common patent policy of ITU-T/ITU-R/ISO/IEC and other established policies of these standardization organizations. The persons named below as contacts can assist potential submitters in identifying the relevant policy information.

2.6 Contribution to Standardization

Proponents are informed that based on the submitted proposals, a standard specification will be created. If they submit a proposal and (part of) the proposed technology is accepted for inclusion in the standard, they will be expected to attend subsequent WG1 meetings and contribute to the creation of the relevant documents. Within this process, evolution and changes are possible as several technologies may be combined to obtain a better performing solution.

2.7 Further information

2.7.1 HTJ2K Ad hoc Group

The HTJ2K Ad Hoc Group was established at the 74th JPEG meeting in October 2016, following the demonstration of a candidate technical solution. The HTJ2K Ad Hoc Group is charged with developing this CfP, collecting use cases and evidence for both requirements and technologies, and facilitating the process leading to the evaluation of proposals, as identified in the above timeline. All interested parties are requested to register to the email reflector of the HTJ2K Ad Hoc Group at:

https://listserv.uni-stuttgart.de/mailman/listinfo/jpeg-htj2k

2.7.2 Contacts

Touradj Ebrahimi (JPEG Convener)
Email: Touradj.Ebrahimi@epfl.ch

David Taubman (HTJ2K Ad hoc Group Chair)
Email: d.taubman@unsw.edu.au

Siegfried Foessel (HTJ2K Ad hoc Group Co-Chair)
Email: siegfried.foessel@iis.fraunhofer.de

ANNEX A – DETAILED USE CASES

A.1. Encoding Use Cases

A.1.1. Lightweight compression on cinematography equipment

The ultimate target in cinematography is often JPEG 2000, but complexity constraints usually prevent JPEG 2000 compression being applied immediately during acquisition. It is desirable for the industry to converge on a common standard output format for cinematography cameras and auxiliary equipment. HTJ2K is a compelling candidate for this format, allowing both high throughput/low power encoding and seamless integration with workflows that involve JPEG 2000, for capture, editing, archival and ultimately distribution.

This use case requires compression of content ranging from full HD at 60fps through to 8K at 120fps, with an important target being 4K at 60fps. Input formats are typically 444 at 12 bits/sample.
Compression ratios of interest for this content are around 10:1. Specifically, bit-rates in the range 2bpp to 4bpp are of interest, but higher bit-rates up to 8bpp may also be required depending on the efficiency of the coding algorithm. Current on-camera encoding strategies involve quantization-based quality control, where quantizer settings are adjusted to ensure a sufficient level of quality, leading to a variable compressed size. For this application, both hardware (FPGA) and software (CPU/GPU) implementations are important.

A.1.2. Virtual Reality Cameras (including 360 degree cameras)

At least one virtual reality capture device already uses JPEG 2000. Low power and low complexity are important for this use case, especially considering that multiple cameras are employed by the acquisition system. At the same time, not all the captured content from a VR camera can generally be viewed at the same time. This implies that the encoded format for VR applications should support region-of-interest access, and perhaps other forms of scalability that are existing strengths of the JPEG 2000 standards.

In this use case, HTJ2K would be used to encode original content produced during image acquisition. Selective communication (e.g., via JPIP) and decoding of content of interest is enabled immediately by this approach, but an interactive browsing experience can be improved further by transcoding to the bit-stream format of JPEG 2000 Part 1. It is expected that this would result in somewhat higher coding efficiency and superior quality scalability, which are both important for highly responsive interactive access to image sources.

This application requires compression of multiple (e.g. 6-8) video streams, typically at full HD resolution. Target compressed bit-rates for this application are 1bpp to 4bpp. Some applications can be latency sensitive, in which case the end-to-end latency may need to be smaller than one frame period.

A.1.3. Compression of earth observation imagery, including hyperspectral imagery

Current commercial satellites are capable of collecting large quantities of imagery. For example, DigitalGlobe's WoldView-4 satellite sensor will be capable of acquiring imagery over 680,000 sqm per day, with permission recently granted to adopt resolutions down to 25cm. This amounts to a data rate of 126 Msamples/s per channel. Hyperspectral sensing at these resolutions will clearly be capable of generating data well into the giga-sample/s range.

At the same time, satellite communication capabilities are limited. At the high end, satellites operating in the X band include those from Satellite Imaging Corporation (800 Mbit/s) and Google's Terra Bella satellites (400Mbit/s). Satellites operating in the S band experience much lower communication bandwidths, in the tens of Mbits/s.

Evidently, compression of satellite imagery is very important. For hyperspectral content, especially, one cannot expect to communicate all content that can be resolved to a base station. As a result, the ability to selectively store and communicate regions of interest and resolutions of interest is expected to be increasingly important. These are capabilities of JPEG 2000 Part 1 as it stands today, and indeed JPEG 2000 is used in some current satellites. However, satellites also have limited power available for performing the compression itself.

HTJ2K addresses the requirements of future satellite imaging applications by providing both a low complexity encoding path and the ability to transcode to the Part 1 JPEG 2000 bit-stream format on demand, on a block-by-block basis if desired. Content transcoded to JPEG 2000 Part 1 is expected to have superior quality scalability and coding efficiency, which can facilitate progressively degradable storage of content and efficient communication of content of interest to a base station.
Similar considerations apply to other platforms for earth observation.

A.1.4. Compression and Distribution of Astronomical Imagery

The planned Square Kilometre Array (SKA) will produce data products, some of which will be multi-dimensional imagery of significant size – multiple TBs up to 1 PB. Current telescopes such as VLA, ASKAP, Arecibo are already producing scientific imagery in similar regimes. The current implementation of JPEG 2000 and its JPIP has enabled the remote visualization to 0.5 and 1.1 TB spectral imaging cubes.

In a current workflow, 4K images with originally 32-bit floating-point intensity data are losslessly encoded and stacked into cubes of 256 components each. Such sub-cubes are generated naturally due to the parallelization of data reduction processing. The sub-cubes are further linked to form a 30,720 component super-cube that can be encapsulated within a single JPX file and remotely accessed via JPIP. The number of components is not limited to these current values and may indeed be required to be much larger. Very often some data has to be reprocessed and new sub-cubes have to be produced. Faster encoding and a lower computational footprint for the encoding will hugely benefit the SKA.

The ability to retrieve and manipulate any arbitrary part of volumetric imagery is crucial to scientific exploration of PB scale data sets. Natural use of multi-component encoding could be a good aid for this, as supported by Part 2 of JPEG 2000.

Another aspect is accurate preservation of the true data. Scientific exploration of astronomy data is done at the lowest signal-to-noise ratio. Some techniques involve aggregation of data to detect the signal from e.g. a radio galaxy. This requirement is often in contrast to the requirement to enable remote exploration. Some optimal solution is desired here.

While the nature of radio telescope data is linear, the visual exploration of the data is usually done in a logarithmic domain. So, the ability to use non-linear transforms is certainly a benefit, as offered by Part 2 of JPEG 2000. Also, because telescopes are taking the images of the sky, non-liner mapping would ideally be defined as part of the compressed representation, as currently it can only be done at the client side with some significant computational cost.

The bandwidth optimization for remote data visualization in astronomy traditionally is achieved using a multi-resolution approach, but the multi-quality approach preserves spatial features in far higher compression regimes and is thus beneficial for studies requiring high spatial resolution.

The aggregated throughput required will be close to 3 Tera Voxels/s; however this won't be done using a single processor.

A.2. Decoding Use Cases

A.2.1. Efficient local rendering from a JPIP client cache

JPIP clients receiving content in JPEG 2000 form typically cache this content and render dynamically from their cache. In interactive applications, where the user pans and zooms across a potentially very large image (or video) surface, the same code-blocks tend to be decoded over and over again. While techniques to cache decoded content are possible, they are clumsy and very memory intensive. As a result it is preferable to render directly from the compressed data, retrieving the relevant code-block bit-streams and reconstructing the viewport of interest on the fly. This can be done in a highly responsive manner, for seemless interactive navigation, even on mobile devices. However, power consumption is non-negligible.

A preferable approach is for the cached code-block bit-streams to be transcoded to HTJ2K on the fly. In this use case, each time a code-block is accessed and decoded, it is transcoded to the HTJ2K format. The next time the code-block is accessed, it can be decoded with much higher throughput, and hence lower resource utilization.

The primary requirement in this case is for decoding with minimal energy consumption, while being able to represent the transcoded content economically. Since compressed content occupies space in a client cache, it is important that the transcoded representation for a code-block be not too much larger than the cached content received originally.

Coded original JPEG 2000 content could have a wide range of bit-rates, extending to lossless.

A.2.2. Medical image compression, including volumetric imagery

Within a hospital environment, networking infrastructure may render the quality scalability feature of JPEG 2000 less important. High throughput / low power compression and decompression have been identified as the most important requirements for this setting.

When the same content is communicated outside the hospital environment, however, coding efficiency and quality scalability both become much more important.

This dichotomy naturally suggests the benefits of a high throughput option to JPEG 2000 that supports reversible transcoding.

In this use case, content is initially compressed using JPEG 2000 with a quality scalable lossy to lossless representation. Content can then be dynamically transcoded to HTJ2K by server equipment so that browsing devices in the hospital environment can be as energy-efficient as possible. At the same time, the original JPEG 2000 content can be served to browsing clients located outside the hospital environment, allowing maximally responsive browsing in lower bandwidth contexts.

This application involves both images and medical volumes, which can be efficiently handled using JPEG 2000 transform extensions found in Part 2 and Part 10 (JP3D).

Lossless compression is likely to be important for this application.

A.2.3. Cost effective visualization of high resolution JPEG 2000 media

An Interoperable Master Format (IMF) conforming to SMPTE standards ST 2067-20 or ST 2067-21 represents a Home Video Master version of a Motion Picture. The SMPTE standards ST 2067-20 and ST 2067-21 require the use of either the Broadcast Profile or IMF Profiles of the JPEG 2000 Part 1 standard using either lossy or lossless compression modes. The content owner has a desire to review the IMF after a Mastering Facility creates it before distributing it to the content licensees worldwide.

On an IMF review workstation, the IMF is transcoded losslessly from JPEG 2000 Part 1 to HTJ2K and is stored on the review workstation’s local storage system. The content owner staff “golden eye” reviews the content carefully when the transcoding operation has completed. The use of HTJ2K on the playback systems reduces the number of CPU devices needed.

A.2.4. Cost effective transcoding of high resolution JPEG 2000 media

An Interoperable Master Format (IMF) conforming to SMPTE standards ST 2067-20 or ST 2067-21 represents a Home Video Master version of a Motion Picture Film. The SMPTE standards ST 2067-20 and ST 2067-21 require the use of either the Broadcast Profile or IMF Profiles of the JPEG 2000 Part 1 standard using either lossy or lossless compression modes. A licensee receives an IMF from a content owner and stores the content using Cloud.

The licensee uses the Cloud Computing to losslessly transcode the IMF to HTJ2K. The licensee then decodes the HTJ2K content and encodes to H.264/AVC using 8 different encoding profiles at different resolutions and bitrates for an OTT video distribution service. One year passes and new devices become available that have VP9 decoding capability, and the content licensee decodes the HTJ2K content again and encodes to VP9 using 10 different encoding profiles at different resolutions and bitrates for the OTT video distribution service. The use of HTJ2K decoding instead of JPEG 2000 Part 1 decoding reduces the content licensee’s Cloud Computing CPU time expense.

A.3. End-to-End Use Cases

A.3.1. Remote viewfinder in videography/cinematography applications

This use case includes various forms of remote monitoring, such as drone piloting and stage-based monitoring. Low latency is especially important. While some existing proprietary methods advertise latencies below 1ms, the optimal trade-off between quality and latency is not precisely known, especially considering that content is typically acquired with a frame interval of 16ms.

While various coding algorithms can be used in this application, there is an interest within the cinematography industry to adopt similar or identical algorithms for both the encoding of the primary video content and the encoding of remote viewfinder content. This makes HTJ2K a particularly interesting option.

For this use case, the typical resolution is full HD at 60Hz. The application is constrained by wireless bandwidths of around 150 Mbit/s, which is equivalent to around 1.25 bits/pixel for full HD at 60fps.

A.3.2. Unified carriage of audio and video streams in live performances

Similar to the above use case, this one involves live video, but the purpose and viewing locations are different. In this use case, video content is exchanged between the main stage and mixing/control locations, carried over the same infrastructure that is currently used for audio. Typically, audio is carried over Gb ethernet and the available bandwidth for the carriage of video is between 200 Mbit/s and 800 Mbit/s. Latencies of less than 10ms have been identified as important for this application. HD and 4K video resolutions are of interest.

This application requires low power encoding and decoding, with the option to use either hardware FPGA platforms or software within modest CPU platforms. The specific combination of features that are particularly relevant to HTJ2K here are real-time software processing and relatively low compressed bit-rates.

A.3.3. Telepresence Systems

This use case can include large displays and image capture configurations, such as display walls. Current display wall resolutions can range into the hundreds of Mega pixels. Some applications for display walls involve local rendering of data sets for visualization. Other applications, however, involve data sharing and interactive communication between sites. For the latter set of applications, low (sub-frame) latency high throughput encoding and decoding tools are required. Moreover, in interactive display applications, it is expected that some parties in the interaction will connect via conventional display systems, which are capable of displaying only a portion of the overall content.

Features of the use case include very high spatial resolutions, the need for low (sub-frame) latency compression, and the need for efficient communication and decoding of windows of interest into the content. These features are all current strengths of JPEG 2000 Part 1, especially when coupled with JPIP (JPEG 2000 Part 9), but it is desirable to reduce the processing load, especially for encoding or decoding of the full content at display-wall resolutions.

For this use case, acceptable bit-rates are expected to lie in the range 0.5bpp to 2bpp, but interactive communication via large display walls is a new area that has not yet been widely explored.

ANNEX B – SUBMISSION REQUIREMENTS

B.1. High-level description and analysis

The submission shall include:

- A high-level technical description of the submission including block diagrams of encoder and decoder.
- Analysis demonstrating that the submission meets all requirements.

B.2. Binary encoder and decoder executables

The submission shall include separate encoder, decoder and transcoder executable programs, configurable via command line or configuration file. The executable programs shall conform to the Evaluation Procedures specified in Annex C.

The encoder executable program shall support the following configuration options:
- Code block size
- Quantization parameters (lossy and lossless)
- JPEG 2000 Part 1 DWT and color transform parameters

The transcoder executable program shall support transcoding from HTJ2K to JPEG 2000 Part 1 codestream, and vice-versa.

The executable programs shall be instrumented to allow accurate throughput measurements of the block coder implementation.

Executable compression or similar tools may be used to prevent reverse engineering or disassembly of the submitted executable files.

The submission shall include the command-line parameters intended for Evaluation Procedures specified in in Annex C.

B.3. Algorithm and design description

The submission shall include a detailed description of the proposed algorithm, including:
- a theoretical discussion on the mechanism used by proposed algorithm to satisfy requirements of Section 1.4; and
- any specific CPU instruction sets from which the implementation would benefit, including vector processing instructions.

Beyond CPU implementations:

- Analysis of GPU implementations may be provided and, if so, focus on available parallelism. More concrete evidence of GPU implementation complexity is also of interest.

- Analysis of hardware implementations may be provided. A minimal analysis should consider memory, use of complex logic units such as multipliers and barrel shifters, as well as identifying the likely critical path dependencies in an encoder and decoder. Additional evidence of hardware implementation complexity is also of interest.

B.4. Technical documentation

A complete technical description of the proposal, sufficient to enable standardization, shall be provided, including:

- Overview of operations
- Codestream syntax
- Lossless transcoding process to and from JPEG 2000 Part 1
- Decoding process

The description shall include all necessary information to create a bit-exact codestream.

B.5. Verification model source code

The submission shall include source code to serve as a verification model, written in a high-level language, such as C or C++.

The source code shall be in the form of a patch against the reference software specified in JPEG 2000 Part 4.

Assembly language or GPU code shall not be included.

Source code shall be documented and understandable.

The source code does not need to correspond to the executable programs of Annex B.2.

All libraries used by the source code shall be available in source code form with ISO/IEC and ITU-T compliant terms.

The codestream output of the source code shall match exactly the ones obtained using the submitted executable files.

ANNEX C– EVALUATION PROCEDURES

C.1. Test Data

C.1.1. General

Proponents are invited to use the points of contact at Section 2.7.2 to retrieve the Test Data. Some of the Test Data is available only to WG1 members.

C.1.2. Video Data

The following test sequences are used:

- APPLE_BasketBallScreen_2560x1440p_60_8b_sRGB_444
- ARRI_AlexaDrums_3840x2160p_24_12b_P3_444
- ARRI_AlexaHelicopterView_3840x2160p_24_12b_P3_444
- ARRI_Lake2_2880x1620p_24_8b_bt709_444
- ARRI_PublicUniversity_2880x1620p_24_8b_bt709_444
- BLENDER_Sintel1_4096x1744p_24_8b_sRGB_444
- BLENDER_Sintel2_4096x1744p_24_10b_sRGB_444
- BLENDER_TearsOfSteel_4096x1714p_24_12b_sRGB_444
- EBU_PendulusWide_3840x2160p_50_10b_bt709_444
- HUAWEI_ScMap_1280x720p_60_8b_sRGB_444
- RICHTER_ScreenContent_4096x2160p_15_8b_sRGB_444
- VQEG_CrowdRun_3840x2160p_50_8b_bt709_444
- VQEG_ParkJoy_3840x2160p_50_8b_bt709_444

For all sequences, 4:2:2 variants either exist or can be derived, and may be added to the test set.

C.1.3. Image Data

The performance of the HTJ2K block coding algorithm is evaluated against the original suite of test images used in the development of JPEG 2000 Part 1, supplemented with the“Waltham1” and “Waltham2” images contributed[2] by Andy Walter from IBM during the development of JPIP.

- AERIAL2_2048x2048_8b_RGB
- BIKE_2048x2560_8b_RGB
- CAFE_2048x2560_8b_RGB
- CATS_3072x2048_8b_Y
- CCITT1_3504x4750_1b_Y
- CCITT2_3072x4352_1b_Y

[2] These images were generated using a high quality 3-pass scanning process, providing exceptional detail while still offering depth of field.

- CHART_1688x2347_8b_RGB
- COMPOUND2_5120x6624_8b_Y
- FINGER_512x512_8b_Y
- GOLD_720x576_8b_Y
- HOTEL_720x576_8b_Y
- MAT_1528x1146_8b_Y
- SEISMIC_512x512_8b_Y
- TOOLS_1524x1200_8b_Y
- TEXTURE1_1024x1024_8b_Y
- TEXTURE2_1024x1024_8b_Y
- ULTRASOUND_512x448_8b_Y
- WATER_1465x1999_8b_Y
- WOMAN_2048x2560_8b_RGB
- XRAY_2048x1680_12b_Y
- WALTHAM1_3600x2600_8b_RGB.tif
- WALTHAM2_3800x2600_8b_RGB.tif

Other images may be used as well, including those from the High Density Camera Array (HDCA) data set.

C.2. Evaluation of Throughput and Bitrate Increase

C.2.1. General

The throughput increase of the HTJ2K block coder compared to the JPEG 2000 Part 1 block coder is measured. To achieve this, the input to the HTJ2K block coder is chosen to be the same as it would be for the JPEG 2000 Part 1 block coder, such that the decoded HTJ2K codestream produces wavelet coefficients that are identical to the wavelet coefficients produced by decoding a corresponding JPEG 2000 Part 1 codestream.

There are two contexts of principal interest for evaluating the HTJ2K block coder:

- The stand-alone context is one in which the HTJ2K block coder is used to compress original content and subsequently decompress the content; and
- The second context involves transcoding between HTJ2K and the existing JPEG 2000 Part 1 representations and vice-versa.

NOTE: Although the evaluation is performed using a subset of all configurations of the JPEG Part 1 block coder, the HTJ2K block coder is expected to support all such configurations .

C.2.2. Stand-Alone Context

The stand-alone context is evaluated by driving the block encoder with quantized sample values (except for lossless tests), and the corresponding block decoder reproducing the identical quantized sample values. This is achieved by configuring the quantization parameters using the derived quantization mode of JPEG 2000 Part 1 and modulating the base quantization step size to sweep through the target test conditions. This JPEG 2000 Part 1 encoding strategy is commonly called "quantizer-driven encoding". This encoding strategy is selected to facilitate a reliable comparison of alternate block coding algorithm. Nonetheless, the HTJ2K block coder is expected to support all other encoding strategies supported by the JPEG 2000 Part 1 block coder.

Stand-alone performance is reported as the bitrate change relative to JPEG 2000 Part 1 and the measured throughput of an implementation. The throughput of encoding and decoding implementations is evaluated on a specific platform detailed in Annex C.4, noting that the throughput is implementation dependent.

Encoding and decoding throughputs are obtained by the evaluation process, and reported in units of Gpel/s (giga pixels per second) along with bitrate change relative to JPEG 2000 Part 1 reported as a percentage, at compressed bit-rates of 0.5bpp, 1.0bpp, 2.0bpp, 4.0bpp and 6.0bpp; lossless configurations will also be tested. Two types of graphs will be produced, as follows:

- bitrate change relative to JPEG 2000 Part 1 (vertical axis) against corresponding JPEG 2000 Part 1 bits/pixel (horizontal axis), allowing coding efficiency to be compared with a JPEG 2000 reference (Annex D); and
- throughput increase (vertical axis) against corresponding J2K bits/pixel (horizontal axis), allowing throughput to be compared with a suitable JPEG 2000 reference (Annex D), as a function of bitrate.

C.2.3. Transcoding Contexts

Applications for HTJ2K involve transcoding in one direction or another. For evaluation purposes, the more challenging task is transcoding from a JPEG 2000 Part 1 codestream to a HTJ2K codestream. Since transcoding is lossless (Section 1.4.3), decompressed imagery is identical regardless of whether transcoding occurred.

Each image/video in the test set is encoded using the JPEG 2000 reference implementation and configurations (Annex D) to target bit-rates of 0.5bpp, 1.0bpp, 2.0bpp, 4.0bpp and 6.0bpp; lossless configurations will also be tested. Since visual weighting is almost always desirable, this initial compression is performed with a defined set of visual weights.

After transcoding to HTJ2K codesteams, the performance of an HTJ2K solution is evaluated in terms of two quantities:

- length of the transcoded codestream; and
- time taken to decode the transcoded content.

C.3. Timing Measurements

The evaluation procedures described above involve the assessment of encoding and decoding times. This is achieved by running the implementation on a test framework on the evaluation platform of Section C.4. To obtain a reliable indication of throughput, implementations should be reasonably optimized, although it is expected that further optimizations are possible.

Throughput is measured for a single-threaded implementation. All implementations should however be capable of utilizing all cores efficiently.

The test framework is designed to minimize or eliminate the following factors:

- Disk I/O presents an overhead that may be considerable for an HTJ2K solution
- Time varying CPU clock rates (i.e., turbo states) may affect the reproducibility of results.
- Caching effects in the system

NOTE: The Final Call for Proposals will detail the test framework, and provide reference timing measurements on the Evaluation Platform.

C.4. Evaluation Platform

The Evaluation Platform uses an Intel i7 processor (Skylake or newer).

NOTE: The Final Call for Proposals will provide a detailed description of the evaluation platform.

C.5. Anchor Evaluation

Anchors are generated by applying this evaluation procedure to the JPEG 2000 Part 1 block coder using selected permutations of block coder modes, e.g. BYPASS.

ANNEX D – JPEG 2000 REFERENCE

D.1. Implementation

The Reference Implementation consists of the Kakadu SDK (http://www.kakadusoftware.com).

D.2. Configurations

The following specify Reference Configurations of the JPEG 2000 Part 1 block coder:

- Code-block size: 32x32, 64x64, 1024x4 (width x height)
- Wavelet Filters: 9-7 (lossy), 5-3 (lossless)
- DWT decomposition levels: 5
- Decorrelating multi-component transform: enabled
- Default code-block style (all flags 0)