being prepared as the newest video coding standard of the ITU-T Video Coding Experts Group and the ISO/IEC Moving Picture Experts Group. The main goal of the HEVC standardization effort is to enable significantly improved compression performance relative to existing standards—in the range of 50% bit-rate reduction for equal perceptual video quality. This paper provides an overview of the technical features and characteristics of the HEVC standard.

***Index Terms*—Advanced video coding (AVC), H.264, High Efficiency Video Coding (HEVC), Joint Collaborative Team on Video Coding (JCT-VC), Moving Picture Experts Group (MPEG), MPEG-4, standards, Video Coding Experts Group (VCEG), video compression.**

I. INTRODUCTION

THE High Efficiency Video Coding (HEVC) standard is the most recent joint video project of the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) standardization organizations, working together in a partnership known as the Joint Collaborative Team on Video Coding (JCT-VC) [1]. The first edition of the HEVC standard is expected to be finalized in January 2013, resulting in an aligned text that will be published by both ITU-T and ISO/IEC. Additional work is planned to extend the standard to support several additional application scenarios, including extended-range uses with enhanced precision and color format support, scalable video coding, and 3-D/stereo/multiview video coding. In ISO/IEC, the HEVC standard will become MPEG-H Part 2 (ISO/IEC 23008-2) and in ITU-T it is likely to become ITU-T Recommendation H.265.

development of the well-known ITU-T and ISO/I
The ITU-T produced H.261 [2] and H.263
produced MPEG-1 [4] and MPEG-4 Visual [5]
organizations jointly produced the H.262/MPEG
and H.264/MPEG-4 Advanced Video Coding (A
dards. The two standards that were jointly produ
particularly strong impact and have found their w
variety of products that are increasingly prevale
lives. Throughout this evolution, continued effo
made to maximize compression capability and
characteristics such as data loss robustness, whi
the computational resources that were practical f
ucts at the time of anticipated deployment of e

The major video coding standard directly
HEVC project was H.264/MPEG-4 AVC, whic
developed in the period between 1999 and 2
was extended in several important ways fron
H.264/MPEG-4 AVC has been an enabling techn
ital video in almost every area that was not prev
by H.262/MPEG-2 Video and has substantially
older standard within its existing application
widely used for many applications, including bro
definition (HD) TV signals over satellite, cable,
transmission systems, video content acquisitio
systems, camcorders, security applications, Inte
bile network video, Blu-ray Discs, and real-t
tional applications such as video chat, video con
telepresence systems.

However, an increasing diversity of servic
ing popularity of HD video, and the emergen
HD formats (e.g., 4k×2k or 8k×4k resolution
even stronger needs for coding efficiency supe
MPEG-4 AVC's capabilities. The need is even
higher resolution is accompanied by stereo
capture and display. Moreover, the traffic cau
applications targeting mobile devices and table
as the transmission needs for video-on-demand
imposing severe challenges on today's networks
desire for higher quality and resolutions is a
mobile applications.

HEVC has been designed to address essential
applications of H.264/MPEG-4 AVC and to par
on two key issues: increased video resolution
use of parallel processing architectures. The sy

Manuscript received May 25, 2012; revised August 22, 2012; accepted August 24, 2012. Date of publication October 2, 2012; date of current version January 8, 2013. This paper was recommended by Associate Editor H. Gharavi. (*Corresponding author: W.-J. Han.*)

G. J. Sullivan is with Microsoft Corporation, Redmond, WA 98052 USA (e-mail: garysull@microsoft.com).

J.-R. Ohm is with the Institute of Communication Engineering, RWTH Aachen University, Aachen 52056, Germany (e-mail: ohm@ient.rwth-aachen.de).

W.-J. Han is with the Department of Software Design and Management, Gachon University, Seongnam 461-701, Korea (e-mail: hurumi@gmail.com).

T. Wiegand is with the Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute, Berlin 10587, Germany, and also with the Berlin Institute of Technology, Berlin 10587, Germany (e-mail: twiegand@ieee.org).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2012.2221191

quality, implementation cost, time to market, and other considerations). However, it provides no guarantees of end-to-end reproduction quality, as it allows even crude encoding techniques to be considered conforming.

To assist the industry community in learning how to use the standard, the standardization effort not only includes the development of a text specification document, but also reference software source code as an example of how HEVC video can be encoded and decoded. The draft reference software has been used as a research tool for the internal work of the committee during the design of the standard, and can also be used as a general research tool and as the basis of products. A standard test data suite is also being developed for testing conformance to the standard.

This paper is organized as follows. Section II highlights some key features of the HEVC coding design. Section III explains the high-level syntax and the overall structure of HEVC coded data. The HEVC coding technology is then described in greater detail in Section IV. Section V explains the profile, tier, and level design of HEVC. Since writing an overview of a technology as substantial as HEVC involves a significant amount of summarization, the reader is referred to [1] for any omitted details. The history of the HEVC standardization effort is discussed in Section VI.

II. HEVC CODING DESIGN AND FEATURE HIGHLIGHTS

The HEVC standard is designed to achieve multiple goals, including coding efficiency, ease of transport system integration and data loss resilience, as well as implementability using parallel processing architectures. The following subsections briefly describe the key elements of the design by which these goals are achieved, and the typical encoder operation that would generate a valid bitstream. More details about the associated syntax and the decoding process of the different elements are provided in Sections III and IV.

A. Video Coding Layer

The video coding layer of HEVC employs the same hybrid approach (inter-/intrapicture prediction and 2-D transform coding) used in all video compression standards since H.261. Fig. 1 depicts the block diagram of a hybrid video encoder, which could create a bitstream conforming to the HEVC standard.

reference picture and motion vector (MV) to
predicting the samples of each block. The encod
generate identical interpicture prediction signa
motion compensation (MC) using the MV and
data, which are transmitted as side information

The residual signal of the intra- or interpict
which is the difference between the original blo
diction, is transformed by a linear spatial transfo
form coefficients are then scaled, quantized, e
and transmitted together with the prediction in

The encoder duplicates the decoder proces
gray-shaded boxes in Fig. 1) such that both
identical predictions for subsequent data. There
tized transform coefficients are constructed by
and are then inverse transformed to duplicat
approximation of the residual signal. The re
added to the prediction, and the result of that
then be fed into one or two loop filters to smoo
induced by block-wise processing and quantiza
picture representation (that is a duplicate of th
decoder) is stored in a decoded picture buffer
the prediction of subsequent pictures. In genera
encoding or decoding processing of pictures oft
the order in which they arrive from the source;
distinction between the decoding order (i.e., b
and the output order (i.e., display order) for a

Video material to be encoded by HEVC is
pected to be input as progressive scan imagery
the source video originating in that format or
deinterlacing prior to encoding). No explicit c
are present in the HEVC design to support the u
scanning, as interlaced scanning is no longer us
and is becoming substantially less common fo
However, a metadata syntax has been provide
allow an encoder to indicate that interlace-scan
been sent by coding each field (i.e., the even or
lines of each video frame) of interlaced video
picture or that it has been sent by coding each in
as an HEVC coded picture. This provides an ef
of coding interlaced video without burdening d
need to support a special decoding process for

In the following, the various features invo
video coding using HEVC are highlighted as f

1) *Coding tree units and coding tree block (C*
The core of the coding layer in previous

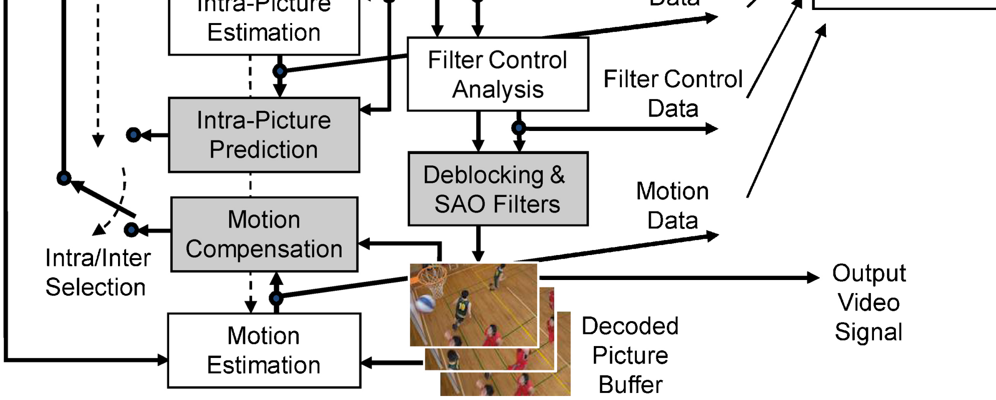

Fig. 1. Typical HEVC video encoder (with decoder modeling elements shaded in light gray).

the macroblock, containing a 16×16 block of luma samples and, in the usual case of 4:2:0 color sampling, two corresponding 8×8 blocks of chroma samples; whereas the analogous structure in HEVC is the coding tree unit (CTU), which has a size selected by the encoder and can be larger than a traditional macroblock. The CTU consists of a luma CTB and the corresponding chroma CTBs and syntax elements. The size $L \times L$ of a luma CTB can be chosen as L = 16, 32, or 64 samples, with the larger sizes typically enabling better compression. HEVC then supports a partitioning of the CTBs into smaller blocks using a tree structure and quadtree-like signaling [8].

2) *Coding units (CUs) and coding blocks (CBs):* The quadtree syntax of the CTU specifies the size and positions of its luma and chroma CBs. The root of the quadtree is associated with the CTU. Hence, the size of the luma CTB is the largest supported size for a luma CB. The splitting of a CTU into luma and chroma CBs is signaled jointly. One luma CB and ordinarily two chroma CBs, together with associated syntax, form a coding unit (CU). A CTB may contain only one CU or may be split to form multiple CUs, and each CU has an associated partitioning into prediction units (PUs) and a tree of transform units (TUs).

3) *Prediction units and prediction blocks (PBs):* The decision whether to code a picture area using interpicture or intrapicture prediction is made at the CU level. A PU partitioning structure has its root at the CU level.

Depending on the basic prediction-type
luma and chroma CBs can then be furth
and predicted from luma and chroma pre
(PBs). HEVC supports variable PB sizes
down to 4×4 samples.

4) *TUs and transform blocks*: The predicti
coded using block transforms. A TU tree
its root at the CU level. The luma CB re
identical to the luma transform block (T
further split into smaller luma TBs. The s
the chroma TBs. Integer basis functions s
of a discrete cosine transform (DCT) are
square TB sizes 4×4, 8×8, 16×16, and 3
4×4 transform of luma intrapicture predic
an integer transform derived from a form
transform (DST) is alternatively specified

5) *Motion vector signaling:* Advanced moti
diction (AMVP) is used, including deriva
most probable candidates based on data
PBs and the reference picture. A merge
coding can also be used, allowing the
MVs from temporally or spatially neig
Moreover, compared to H.264/MPEG-4 A
skipped and direct motion inference are

6) *Motion compensation:* Quarter-sample pre
for the MVs, and 7-tap or 8-tap filters
interpolation of fractional-sample positio
to six-tap filtering of half-sample posit
by linear interpolation for quarter-sampl

DC (flat) prediction modes. The selected intrapicture prediction modes are encoded by deriving most probable modes (e.g., prediction directions) based on those of previously decoded neighboring PBs.

8) *Quantization control:* As in H.264/MPEG-4 AVC, uniform reconstruction quantization (URQ) is used in HEVC, with quantization scaling matrices supported for the various transform block sizes.
9) *Entropy coding:* Context adaptive binary arithmetic coding (CABAC) is used for entropy coding. This is similar to the CABAC scheme in H.264/MPEG-4 AVC, but has undergone several improvements to improve its throughput speed (especially for parallel-processing architectures) and its compression performance, and to reduce its context memory requirements.
10) *In-loop deblocking filtering:* A deblocking filter similar to the one used in H.264/MPEG-4 AVC is operated within the interpicture prediction loop. However, the design is simplified in regard to its decision-making and filtering processes, and is made more friendly to parallel processing.
11) *Sample adaptive offset (SAO):* A nonlinear amplitude mapping is introduced within the interpicture prediction loop after the deblocking filter. Its goal is to better reconstruct the original signal amplitudes by using a look-up table that is described by a few additional parameters that can be determined by histogram analysis at the encoder side.

B. High-Level Syntax Architecture

A number of design aspects new to the HEVC standard improve flexibility for operation over a variety of applications and network environments and improve robustness to data losses. However, the high-level syntax architecture used in the H.264/MPEG-4 AVC standard has generally been retained, including the following features.

1) *Parameter set structure:* Parameter sets contain information that can be shared for the decoding of several regions of the decoded video. The parameter set structure provides a robust mechanism for conveying data that are essential to the decoding process. The concepts of sequence and picture parameter sets from H.264/MPEG-4 AVC are augmented by a new video parameter set (VPS) structure.

typically restricted, and the number of CT
is often varied to minimize the packetiz
while keeping the size of each packet wit

4) *Supplemental enhancement information (*
usability information (VUI) metadata: T
cludes support for various types of meta
SEI and VUI. Such data provide informa
timing of the video pictures, the proper in
the color space used in the video signal, 3-
frame packing information, other display
tion, and so on.

C. Parallel Decoding Syntax and Modified Sli

Finally, four new features are introduced in t
dard to enhance the parallel processing capab
the structuring of slice data for packetization
of them may have benefits in particular applic
and it is generally up to the implementer of
decoder to determine whether and how to tak
these features.

1) *Tiles*: The option to partition a picture i
regions called tiles has been specified.
pose of tiles is to increase the capabili
processing rather than provide error resili
independently decodable regions of a p
encoded with some shared header informa
additionally be used for the purpose of
access to local regions of video pictu
tile configuration of a picture consists
the picture into rectangular regions with
equal numbers of CTUs in each tile.
parallelism at a more coarse level of gr
ture/subpicture), and no sophisticated syn
threads is necessary for their use.
2) *Wavefront parallel processing*: When wa
processing (WPP) is enabled, a slice i
rows of CTUs. The first row is processed
way, the second row can begin to be p
only two CTUs have been processed in
the third row can begin to be proces
two CTUs have been processed in th
and so on. The context models of the
in each row are inferred from those in
row with a two-CTU processing lag. W
form of processing parallelism at a rath

in low-delay encoding, where other parallel tools might penalize compression performance.

In the following two sections, a more detailed description of the key features is given.

III. HIGH-LEVEL SYNTAX

The high-level syntax of HEVC contains numerous elements that have been inherited from the NAL of H.264/MPEG-4 AVC. The NAL provides the ability to map the video coding layer (VCL) data that represent the content of the pictures onto various transport layers, including RTP/IP, ISO MP4, and H.222.0/MPEG-2 Systems, and provides a framework for packet loss resilience. For general concepts of the NAL design such as NAL units, parameter sets, access units, the byte stream format, and packetized formatting, please refer to [9]–[11].

NAL units are classified into VCL and non-VCL NAL units according to whether they contain coded pictures or other associated data, respectively. In the HEVC standard, several VCL NAL unit types identifying categories of pictures for decoder initialization and random-access purposes are included. Table I lists the NAL unit types and their associated meanings and type classes in the HEVC standard.

The following subsections present a description of the new capabilities supported by the high-level syntax.

A. Random Access and Bitstream Splicing Features

The new design supports special features to enable random access and bitstream splicing. In H.264/MPEG-4 AVC, a bitstream must always start with an IDR access unit. An IDR access unit contains an independently coded picture—i.e., a coded picture that can be decoded without decoding any previous pictures in the NAL unit stream. The presence of an IDR access unit indicates that no subsequent picture in the bitstream will require reference to pictures prior to the picture that it contains in order to be decoded. The IDR picture is used within a coding structure known as a closed GOP (in which GOP stands for group of pictures).

The new clean random access (CRA) picture syntax specifies the use of an independently coded picture at the location of a random access point (RAP), i.e., a location in a bitstream at which a decoder can begin successfully decoding pictures without needing to decode any pictures that appeared earlier in the bitstream, which supports an efficient temporal coding

33	Sequence parameter set (SPS)
34	Picture parameter set (PPS)
35	Access unit delimiter
36	End of sequence
37	End of bitstream
38	Filler data
39, 40	SEI messages
41–47	Reserved for future use
48–63	Unspecified (available for system use)

order known as open GOP operation. Good supp
access is critical for enabling channel switchin
tions, and dynamic streaming services. Some pi
low a CRA picture in decoding order and preced
order may contain interpicture prediction referen
that are not available at the decoder. These
pictures must therefore be discarded by a deco
its decoding process at a CRA point. For this
nondecodable pictures are identified as random a
leading (RASL) pictures. The location of splic
different original coded bitstreams can be indica
link access (BLA) pictures. A bitstream splic
can be performed by simply changing the NA
a CRA picture in one bitstream to the value
a BLA picture and concatenating the new bi
position of a RAP picture in the other bitstr
picture may be an IDR, CRA, or BLA pict
CRA and BLA pictures may be followed by R
in the bitstream (depending on the particular
NAL unit type used for a BLA picture). Any R
associated with a BLA picture must always be
the decoder, as they may contain references to
are not actually present in the bitstream due
operation. The other type of picture that can
picture in decoding order and precede it in o
the random access decodable leading (RADL)
cannot contain references to any pictures tha
RAP picture in decoding order. RASL and R
are collectively referred to as leading pictures (
that follow a RAP picture in both decoding or
order, which are known as trailing pictures, c
references to LPs for interpicture prediction.

B. Temporal Sublayering Support

Similar to the temporal scalability feature
MPEG-4 AVC scalable video coding (SVC) e

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**