Intra prediction in HEVC

submitted by:
Michael Winkelmann (338024)
(miwinkel@mailbox.tu-berlin.de)

Sommersemester 2012

Abstract

HEVC (High Efficiency Video Coding) is the successive video coding standard of the established H.264/AVC aiming to reduce bitrate of about 50% while preserving same picture quality. To achieve such a good coding efficiency, several techniques are applied of which each includes different coding tools. Intra prediction is a technique where successive parts of an image are tried to be predicted from previously encoded parts. This seminar paper will give an overview over the proposed techniques to realize and improve intra prediction in HEVC.

1 Introduction

With the growing availability of broadband connections and the gaining success of web platforms like YouTube and Vimeo, digital video streaming over the internet is becoming more and more popular and the quality and variety of multimedia and TV services are growing rapidly. Furthermore, as a result of the increasing spread of mobile phones and mobile access to the internet, this process has accelerated further. According to [Cis10], by the end of 2010, Internet video alone represents 40 percent of the consumer internet traffic, where the amount of video shared over P2P is not even included. Until 2014, this percentage is expected to increase to 57 percent.

Moreover, today’s transmission networks have often too less available bandwidth for carrying high quality video contents and the demand for multimedia increases faster than the network infrastructure is able to carry economically [SO11]. Fortunately, new techniques and algorithms for video coding to decrease bitrate while preserving picture quality have been developed.

For these reasons, a new video standard which improves compression efficiency of the existing Advanced Video Coding standard (H.264/AVC) is needed. This standard is called High Efficiency Video Coding and is developed by Joint Collaborative Team on Video Coding (JCT-VC) [HPCC11].

The standardization process started in January 2010 and will be probably finished in early 2013 [HPCC11]. The goal of HEVC is to halve the bitrate by preserving picture quality compared to AVC. Similar to AVC, HEVC is based on an hybrid coding approach combining motion-compensated prediction, intra-picture prediction, 2D transformation coding, loop filtering and adaptive entropy coding. A block diagram of an HEVC decoder given in figure 1.

Intra prediction is going to be considered in this seminar paper. With this techniques, previously coded image parts are considered and used for predicting successive image parts. Bitrate savings in are achieved by encoding only the difference
between predicted and actual image parts. The seminar paper will give an overview over general techniques of intra prediction as well as recent HEVC proposals on this topic. The paper is organized as follows: First, an explanation of the original AVC intra-prediction technique and an overview over the HEVC standardization and codec structure is given. Afterwards, intra prediction techniques proposed during the standardization of HEVC are described.

2 Background

In this section, a short overview about intra prediction in AVC as well as standardization and coding structure of HEVC with focus on intra prediction is given.

2.1 Intra-prediction in AVC

The AVC intra prediction tool provides DC and several directional modes for predicting variable-size blocks. Predecessors of AVC only encoded the differences between DC components of transform blocks to decrease bitrate. With H.264/AVC, advanced techniques for intra prediction have been introduced.

Generally, H.264/AVC distinguishes between macroblocks for with a size of 16x16 pixels containing up to 16 (sub-)blocks with a size of 4x4 pixels (for luminance channel). If such a block or macroblock is encoded in intra mode, a prediction block is linearly generated based on previously encoded and reconstructed (but un-filtered) neighbouring blocks[Ric02] by extrapolation. Afterwards, this prediction block is subtracted from the actual block before encoding. An example which shows a predicted 4x4 block can be seen in figure 4.

Macroblocks usually represent homogeneous regions which can be predicted easily. Otherwise, macroblocks contain complex patterns for which the prediction is much more complicated [Ric02]. For macroblocks, H.264/AVC provides 4 intra prediction modes for luma, which are depicted in figure 3. The first two modes extrapolate pixels from the vertical and horizontal neighbouring pixels, respectively. The third mode estimates the DC coefficient of the block. A DC coefficient is the average of all pixels in a block. The last mode (called Plane mode) performs a planar prediction, by generating a gradient estimated from the neighbouring pixels.

Figure 1: Block diagram of a common HEVC decoder.
Figure 2: Example of intra prediction. Source: [Ric02]

Figure 3: Intra coding modes for macroblocks. Source: [Ric02]
Mode 0 (vertical): extrapolation from upper samples (H), Mode 1 (horizontal): extrapolation from left samples (V), Mode 2 (DC): mean of upper and left-hand samples (H+V), Mode 3 (Plane): a linear ‘plane’ function is fitted to the upper and left-hand samples H.

Source: [Ric02]

Intra-prediction for 4x4 blocks has 9 different mode for luma. These modes are illustrated in figure 4. 8 modes define a certain direction from which the predictor is generated by extrapolation. One mode, called DC mode predicts just the DC coefficient of the performed integer transform.

For chroma blocks, the prediction is similar to the previously described luma prediction for 16x16 macroblocks. Each 8x8 chroma component of a macroblock is predicted from border chroma samples that have previously been encoded and reconstructed. The prediction mode is used for both chroma components.

Besides difference of block and predicted block, the selected intra mode needs to be encoded and signalled to the decoder. As this needs to be done for each 4x4 block, and since 9 mode are possible, this might require a large number of bits. Fortunately, there is a high correlation between intra modes for neighbouring blocks. Thus, a probability for each mode is estimated. For each block, a flag which determines whether the most probable mode is used is encoded. Another flag to signal a mode change is encoded as well. If these flags are unset, the intra mode must be signalling additionally.

Although the intra prediction tools provided in AVC achieve fair bitrate savings, they have several disadvantages. First of all, complex and periodical textures cannot be generated. The coding efficiency decreases rapidly with the distance between reference and predicted pixels. Moreover, artificial edges occur along the directions of intra prediction. With the HEVC standard, those problems are tried to be mitigated.

### 2.2 HEVC standardization process

The HEVC standardization process began in January 2010 with the initial call for papers (CfP) by the ITU/ISO joint committee. Until February 2010, 27 proposals have been submitted.

During the standardization process, several meetings of companies, institutions and people involved in the standardization process where held. In those meetings, the proposals are evaluated and those which have good results are selected to be part of the standard. Papers are selected by increase of coding efficiency and a good trade-off between improvement complexity, coding times and bitrate savings. Furthermore, test models are worked out. This includes the set of test sequences as well as the reference software.
Nearly every meeting, a new version of the reference software, called *HEVC Test Model* (HM), is developed and improved. It includes the latest evaluated coding tools of the submitted proposals. The following version have been released:

- HM1.0, released in October 2010 at the meeting in Guangzhou [TSO10].
- HM2.0, released in March 2011 at the meeting in Guangzhou [WHBO11a].
- HM3.0, released in April 2011 at the Geneva meeting [WHBO11b].
- HM4.0, released in July 2011 at the meeting in Turin [WHBO11c].
- HM5.0, released in November 2011 at San José meeting [MBK11].
- HM6.0, released in February 2012 at the Geneva meeting [MBS12].

Those versions can be considered as milestones in terms of coding efficiency. Unfortunately, there were no documents available yet for the latest version 7.0.

### 2.3 HEVC units

Compared to H.264/AVC, in HEVC, there is no distinction between macroblocks and sub-macroblocks. Instead, there is a hierarchical, quadtree-based structure called *Coding Unit* (CU). CUs have a square shape and can be larger than 16x16 pixels (up to 128x128 pixels) which allows better coding efficiency for large regions of high homogeneity [MTea11].

A CU can be subdivided in a quadtree-based manner adapted to the picture. The root of the quadtree is called *largest coding unit* (LCU) and the leaves are called *smallest coding unit* (SCU) [KM10]. LCU and SCU size must be a power of 2 and
greater or equal to 8. Nearly the whole processing chain is performed in a CU: Intra and inter prediction, transformation, quantization and entropy coding.

Each CU contains one or several variable-block-sized prediction units (PUs) and transform units (TUs). The PU is the unit of intra/inter prediction and multiple PUs can be in a single CU. At PU level, either intra or inter prediction is selected. Furthermore, there are symmetric and asymmetric PU sizes, as shown in figure 5. TU is the transform unit and can contain one or more PUs. For each TU, a spatial block transform and a subsequent quantization of the resulting transform coefficients are performed. Besides the 4x4 and 8x8 integer transform, larger transform sizes are supported in HEVC. Only LCU and the hierarchical depth of CU need to be defined to characterize the various sizes of CU, PU and TU. The structure of HEVC units are depicted in figure 5.

3 Proposed intra-prediction techniques

The initial proposal of Samsung included new tools for intra coding which formed the basis for the standardization process [KM10]. These tools are arbitrary directional intra (ADI), pixel based template matching (PTM), color component correlation based prediction (CCCP). Additionally, multi-parameter intra (MPI) has been proposed; however, it has not been included in the HEVC standard. In the further standardization process, most proposals rely or a very similar to this techniques. The three techniques Angular intra prediction, Planar prediction and Chroma prediction currently implemented in test model HM6.0 [MBS12] will be described in more detail. Moreover, there were some rather eccentric proposals which will be described in section 3.4.

3.1 Directional intra prediction

In directional intra prediction, also called angular intra prediction, predicted pixels are generated by extrapolating neighbouring pixels from a certain direction. The AVC intra prediction uses 8 different prediction directions for 4x4 sub-blocks. Since this number of direction might be insufficient to represent all directional patterns, some proposals in the standardization process suggest to introduce more possible directions to achieve a better prediction. However, more bits are required for signalling those directions [TS10].
The proposed arbitrary directional intra (ADI) by Samsung and BBC [KM10] is an extension of the original AVC intra prediction. It works on larger block sizes and provides up to 33 prediction modes. Similar to H.264/AVC intra prediction, ADI generates pixels by directional extrapolation of neighbouring pixels. Moreover, boundary pixels from the left down can be used for prediction. ADI was the first technique to be included in HM1.0 [TSO10]. When coding the prediction modes, the number of prediction modes is adjusted to the PU size. The majority of prediction modes are defined by the coordinates \((dx, dy)\). Coding efficiency gain is about 2% compared to H.264/AVC high profile. For a PU of size 4, there can be at most 15 directions and for larger block sizes there can be up to 33 directions [MBS12].
However, artifacts occur often along the extrapolated directional texture patterns. To alleviate this problem, pre- and post-filtering is applied for reference samples as well as for the generated directional patterns. The AVC/H.264 standard already
included a pre-filtering process for macroblocks using a low-pass filter on the reference pixels prior to the prediction. In HEVC, this filtering process is extended for all PU sizes larger than 4x4. Moreover, several approaches for post-filtering have been suggested in the HEVC standardization process [IIea10]. In HM 2.0, a linear (2-pel) interpolation method is implemented but it has been shown that DCIM directions perform much better with a more accurate interpolation, such as 32 4-tap interpolation filters [HPCC11].

Other proposals suggest refinements and improvements for HEVC directional intra prediction. The proposal of Nokia, Ericsson and Tandberg additional directions for 8x8 and 16x16 blocks were suggested [UAF10]. France Telecom et al. [Aea10] suggested a rather complex approach to improve coding efficiency. A vector with the largest magnitude of the 2-norm of the gradient field is predicted from neighbouring blocks. If the magnitude exceeds a certain threshold, the prediction is applied also if the DC mode was selected.

Bi-directional intra is another approach [STCY10] where predictors of two different directions are averaged to generate more complex patterns. A bitrate reduction of about 2.1% compared to H.264/AVC main profile could be achieved without increased decoding and encoding times. The proposal [GZGL11a] suggested an improvement of this approach by better exploiting direction-related information. It improves coding efficiency by 0.7% on average compared to HM4.0. Proposal [ZGAG11], called Overlapped Block Intra, can be considered as a variant of bi-directional intra. However, the coding efficiency results were rather moderate and encoding time increased by 1.83x.

Due to high complexity and low coding performance, those coding tools have not been included in the test models.

3.2 Planar prediction

Planar prediction aims for generating smoothly-varying image segments [HPCC11]. In H.264/AVC, planar prediction (signalled as plane mode) tries to approximate a smooth gradient between neighbouring left- and top samples.

In proposal [UAF10], the bottom-right most pixel of a macroblock is signalled explicitly in the bitstream. The samples in between are generated by bilinear interpolation.

The initial proposal [IIea10] assumes that pixels values current block samples and neighbouring samples form a plane. The plane surface is determined by applying a Least Mean Squares method.

3.3 Chroma correlation based prediction

In H.264/AVC, chroma and luma samples are treated completely separated in the prediction process. However, there is a large correlation between chroma and luma samples which can be exploited to increase coding efficiency. Hence, information inferred from reconstructed luma samples can be used for predicting chroma samples.

In proposal [CSH+11], this approach is called color component correlation based prediction (CCCP). An illustration is shown in figure 6. First of all, reconstructed luma samples are downsampled using a bilinear filter to adjust the chroma block size. Afterwards, a segmentation map is generated from the downsampled luma samples by simple thresholding using the mean value. Based on the thresholding results, the

1Because this was an initial proposal, it works on macroblocks instead of CUs or PUs.
content pixels are assigned to the chroma block according to the segmentation map. The prediction is smoothed with an 3x3 averaging filter. The chroma prediction mode is signalled as a special DC mode [HPCC11]. In proposal [CS10], this process was simplified so it only requires integer arithmetic. Compared to other segmentation approaches, the process is not computationally complex since it is uses a simple thresholding approach. Bitrate savings in $Y$, $C_b$, $C_r$ of 1.3%, 6.5% and 5.5% could be achieved and encoding and decoding nearly remained the same. For chroma predictions, the CCCP process replaces the DC prediction. Videos encoded with HM3.0 contained artifacts in chroma channels due to not signalling the CU level. This worsened the subjective quality. This issue was fixed in Panasonic’s proposal [MSN11] and integrated into HM4.0 [WHBO11c]. Chroma prediction was further improved in proposal [MSiS11] by using a better designed downsampling filter. This resulted into bitrate savings of 0.7% on average for chroma without increasing complexity.

Another method based on template matching was proposed in [YTLR10]. Although the overall bitrate savings with an average of 2.6% were convincing, the required encoding and decoding time increasing were too high with 186% and 295%, respectively. Further bitrate savings could be achieved by reusing the CU quadtree from the corresponding luma block for chroma blocks [GZGL11b].

### 3.3.1 Parallel intra coding

Another challenge is the parallelization of intra coding, where several threads work on different partitions on the image plane. In HEVC, these partitions are called tiles and are encoded independent of each other. The more threads are involved in the encoding process, the less is usually the coding efficiency since there are less reference samples available for prediction. Proposal [ZS11] dealt with this issue and introduced a special coding unit, called Parallel Prediction Unit (PPU). The authors tested different partition schemes for 2X parallelism which achieved different encoding and decoding performance:

1. **8x8 prediction with 4x4 residuals**, in which a 8x8 PPU is grouped into 8x4 stripes. The application of this partition scheme increased the bitrate highly by an average of 4% 5%.
2. **Checker board partition**, where 4x4 blocks of a 8x8 PPU are grouped into two sets in a checker board manner. This approach had less impact on coding performance than the previous one.

3. **Stripe partition**, in which a 8x8 PPU is grouped into 8x4 stripes. This approach achieved the best results with a bitrate increases of 0.2% for low delay and 1.3% for all other test conditions.

Moreover, there are other techniques for parallel intra coding in HEVC not directly connected to intra coding, i.e. parallel entropy coding, entropy slices, data levels and more [MCSJ12].

### 3.4 Other approaches

The following approaches are technically interesting but have not been integrated into the test models (yet).

#### 3.4.1 Template matching

Template matching approaches try to improve prediction for periodical and complex textures, which cannot not be generated by directional intra prediction or by calculation of neighbouring pixels. Compared to prediction techniques which are block based, template matching can achieve better prediction for fine structures at the cost of higher encoding and decoding times.

In the Samsung’s proposal [KM10], it is called *Pixel-based Template Matching* (PTM). Three neighbouring pixels from above, left and above-left are considered to predict a sample. A pixel C with minimum template difference determined by is looked up a template area. Afterwards, C is set to the prediction value of the current pixel. In contrast to directional intra prediction, prediction is both based on already encoded neighbouring pixels as well as previously predicted samples. An illustration of this
technique is shown in figure 7. Another proposal which describes pixel-based recursive template matching to generate a prediction block pixel by pixel in raster-scan order [III11].

The template matching average (TMA) tool calculates the average of the first $N$ candidate blocks with the lowest prediction errors to form the predictor. The proposals [LL10] and [WSea10] present a new prediction scheme by considering correlation between lines and pixels instead of blocks. Since the prediction error increases with the distance of reference pixels, these proposals try to attempt to minimize this distance by using pixels and lines for prediction instead of square partitions. Both approaches differ in how to divide a 16x16 block, i.e. in 1x16, 16x1, 2x8 and 8x2 partitions. Template matching is performed subsequently on each target partition based on neighboring pixels as well as previously predicted partitions. The experimental results of [LL10] show bitrate savings of about 4% compared to H.264/AVC High Profile. However, encoding and decoding were more than four times higher.

3.4.2 Multi-parameter intra prediction

Samsung’s proposal [KM10] introduced also a technique called multi parameter intra (MPI). This tool smoothes the luma predicted patterns while maintaining the meaningful details, which results in better coding efficiency from the transform. Moreover, MPI also provides variations to predicted patterns from ADI and Pixel-based Template Matching.

3.4.3 Pyramid and Interleaved Prediction

This kind of prediction takes a downsampled version of the current block and tries to reconstruct the actual block by upsampling it. This approach was suggest in [Slea10]. Proposal [YFLea10] presents a resample-based intra prediction (RIP) technique, which intends to improve coding efficiency by exploiting correlation of pixels instead of blocks. The algorithm has three modes which determine which pixels are signalled first to form the predictor.

3.5 Intra-mode coding

Because in HEVC more intra modes are possible, it is very important to represent them in an efficient way with as less bits as possible. For directional intra, an angle which determines the prediction direction needs to be signalled. As there are up to 30 directions possible, without further optimizations, 5 bits are needed to store this angle. This means that there needs to be a trade-off between the number of intra coding modes and the number of bits. Fortunately, the number of bits can be reduced.

Basically, in codec design, the frequency of coding modes is measured to determine an optimal number of bits for each mode. Similar to entropy coding, most frequent modes usually have special coding symbols whereas less frequent ones are coded with a larger number of bits. The number of bits can be further reduced by differential encoding and by predicting the direction. DCIM [YY10] (Differential Coding of Intra Modes) enables a higher number of intra prediction without increasing overhead for signalling. In this approach, neighborhood pixels are analyzed to estimate the Intra prediction direction and encodes the
selected direction differentially to the previous direction. Moreover, the number of sub-directions increases with PU size. For 4x4 and 8x8 partitions, 14 sub-directions are utilized whereas for larger PU sizes up to 30 directions are available.

One bit per PU is signalled to determine whether DCIM is used. For PUs with enables DCIM, edge detection is performed. With this technique 1.4% and 1.6% gain is achieved compared to HM 2.0 for High Efficiency (HE) and Low Complexity (LC) settings, respectively. Another proposal [MTea11] improved DCIM with a coding efficiency gain of 0.2% and 0.7% for low delay settings. The number of modes varies for different PU sizes. The proposal [PJ11] suggests to code the remaining modes together with the entropy coding.

The chroma intra prediction mode is derived from the corresponding luma intra prediction mode of the same PU and is coded subsequently [NF11]. This implies that chroma intra prediction mode is coded after luma intra prediction modes located in the same PU.

Another important issue is the intra mode coding syntax. Thus, proposals deal with not with coding efficiency of Intra coding but rather simplifying the syntax of intra mode coding. This reduces the decoder complexity and avoids reduces the possibility of bugs in future implementations.

4 Conclusion

As we have seen, intra prediction is a mandatory as well as a highly complex tool for modern video compression to achieve a superior coding efficiency. In HEVC, many new techniques for intra prediction have been integrated. The most effective techniques are directional intra with more directions than in AVC (combined with an efficient mode coding) and chroma prediction based on previously encoded luma samples.

With intra-prediction only, bitrate savings of 29% on average compared to AVC High Profile could be achieved in HM4.0. In HM6.0, bitrate savings are 34% on average compared to JPEG-XR, which is identical to intra-only AVC. Compared to JPEG, which is identical to intra-only MPEG-2 video compression, bitrate savings of about 56% on average can be achieved [LU12]. Moreover, the subjective image quality is even higher [Hor12].

Further improvements are expected within the ongoing development process. The first version of the standard is expected to be finalized by early 2013. Doubling the compression performance compared to AVC then would mean that Ultra HD video could be provided using compressed bit rates similar to those used for HD representations today. Moreover, extensions of HEVC for scalable and multiview coding are expected to be developed.

In the last decade, video coding researchers experimented with different coding strategies, such as compressive sampling and Wavelet-based coding schemes. However, the traditional motion-compensation transform coding is still superior in terms of complexity and coding efficiency [HPCC11].

As all standardized video codecs by MPEG of the past two decades are based on this approach, for a past-HEVC video generation standard it might be possible to make fundamental changes necessary in codec design. Moreover, it might also be possible that coding efficiency of HEVC is already very close to the Shannon limit and thus, another bitrate reduction by 50% might be impossible.
References


