Video Compression Based on Context Adaptive Arithmetic Coding

Abraham. Mathew, and Dr. R. Radhakrishnan

Abstract—In this paper, an efficient multi-view coding tools by using Context Adaptive Arithmetic video Coding (CAAC) is proposed. Using this an efficient single view compression can be obtained. It can also provide a multi view adaptation standard having additional coding gains. Here the main part of the of the encoder is it’s forward path and the video frame to be compressed will be denoted by $F_n$. The frame will be partitioned as two macro blocks of 16 X 16 pixels each and each macro block is encoded as intra or intra mode. On average, for the texture information, the total bit rate can be reduced by 37.2% compared to simulcast High Efficiency Video Coding (HEVC). For depth map compression, gains largely depend on the quality of the captured content. Additionally, a forward compatible solution is proposed offering the possibility for a gradual upgrade from H.264/AVC based stereoscopic systems. With the proposed system, significant rate savings compared to Multi view Video Coding (MVC) are possible. In short, a multi view compression algorithm is entirely based on forward compatible with H.264/AVC and stereo MVC is proposed.

Keywords—Compression, Efficient, Gain, Multi-View.

I. INTRODUCTION

The last years of the twentieth century witnessed an unprecedented and unforeseen progress in computing power, video applications, and network support for video data. Both producers and consumers of video, felt the need of an advanced video codec to replace the existing video compression standards such as H.261, H.262, and H.263. The last of these standard H.263, was developed around 1995 and was outdated by 2001. The two groups responsible for developing video compression standards, ISO-MPEG and ITU-VCEG, felt that a new standard should offer a) an increased compression efficiency b) support for special video applications such as video conferencing, DVD storage, video broadcasting, and streaming over internet c) offer greater reliability. In response to this demand, the ITU started two projects. The first, a short term project that attempt to add extra features to H.263 and this resulted the version 2 of the standard. The second long term project was an effort to develop a new standard for efficient video compression and this project received the code name H.26L also in the same year. The ITU established a new study group, SG16, whose members came from MPEG and VCEG, [2] and approved the new standard. The new standard H.264, is now the part of the huge MPEG-4 project, and is known by several names. The ITU calls it ITU-T recommendation H.264, advanced video coding for generic audio visual services. It’s official title is Advanced Video Coding (AVC). In H.264, slices and macro blocks also have types, an I slice may have only I-type macro blocks, and a B-slice may have B and I macro blocks.

In the market, compression of any of the 3D representation technologies requires forward compatibility with a 2D compression scheme. Because of its superior compression efficiency, H.264/AVC [6] is chosen in a lot of video applications as a basis to be compatible with. On top of this video compression standard, a multi view extension was defined to efficiently compress a multitude of views. This standard was called Multi view Video Coding (MVC) [7]. A subset of MVC, called Stereo Profile, enjoyed large adoption in the market of stereo 3D. Following these activities, there was evidence that an even better 2D compression algorithm was needed to catch up with the ever increasing bandwidth demand of video content. Similar to H.264/AVC standardization, a collaboration of MPEG and VCEG called Joint Collaborative Team on Video Coding (JCT-VC) was created to develop this standard. The process is entering a stabilization phase in its development and the final standard will be called High Efficiency Video Coding (HEVC) [8]. Compared to H.264/AVC, this technology under development already objectively reduces bandwidth with 44% on average [9]. It is estimated that with a subjective comparison, this reduction can even be higher. With an increased performance of single view coding, it becomes more difficult to obtain additional gains from multi view coding of the different views. In this paper, it is shown that multi view coding can still bring significant improvements even when a forward compatible solution with H.264/AVC is facilitated.

II. MULTI VIEW VIDEO CODING

MVC, the multi view extension of H.264/AVC [7], was mainly developed for efficient compression of scenes shot from different viewpoints. For a multi view compression scheme it does not matter if video streams originate from a free viewpoint setup around the scene of interest or from a 3D production camera. The essence is that neighboring views from the same scene contain a lot of correlation. Multi view compression makes use of this aspect by taking into account neighboring views during the compression process of a view. The way MVC uses inter-view correlation is similar to how...
single view compression takes advantage of temporal correlation between successive frames. In a single-view block based codec like H.264/AVC[11], this correlation is reduced with a motion compensation process. With the multi view extension, the same process is applied with a neighboring view which is then called disparity compensation instead of motion compensation. The process of prediction between different views is called inter-view prediction. As an extension of H.264/AVC, MVC provides forward compatibility with its single view variant. One view within MVC is always independently encoded from the other views making this view H.264/AVC compatible [10]. It can be proved that the center view is the H.264/AVC[7]compatible video stream because only temporal referencing is used. This view is also named the base view of a multi view video stream. For the left and the right view, inter-view prediction is applied, making these views only decodable with an MVC compliant decoder. The left and right view of this example configuration can also be called enhancement views. In general, forward compatibility was one of the major advantages facilitating the adoption of MVC in the market. Not only can the H.264/AVC compatible view be decoded from the MVC video stream, it can also easily be extracted from the video stream. This property is called view scalability and can be applied to every view within the multi view video stream. In general, from an MVC video stream, a subset of reference picture level, adaptive inter-view prediction on PU level results. More specifically, each PU[8] indicates the chosen reference frame by means of an index in the reference picture lists. In these lists one or several lower views occur, making it possible to choose for inter-view prediction instead of temporal prediction. On a PU level, the encoder can choose in a Rate Distortion (RD) optimal way if inter-view prediction or inter-frame prediction should be applied. With the proposed multi view compression scheme, forward compatibility with HEVC is guaranteed similarly to the forward compatibility provided by MVC. The same remark can be made for the view scalability aspect of MVC. Additionally, in the proposed compression, the complexity restriction limiting inter-view prediction only for within the same time instance is forced as well. Here the parameter coefficient token is the number of nonzero coefficient and trailing ±1s among the 16 coefficient in the sequence. the number of non zero coefficient can be from 0 to 16. The number of trailing ±1s can be up to the number of nonzero coefficient. Any trailing ±1s beyond the first three are encoded separately as other nonzero coefficients. Parameter trailing-ones –sign-flag is encoded as a single bit for each trailing ±1s signaled by coefficient-token. The trailing-ones are scanned in reverse zigzag order, and the bits that are generated are appended to the binary string generated so far. The next set of bits appended to the binary string encodes the values of the remaining nonzero coefficients. These nonzero coefficients are scanned in the reverse zigzag order, and each is encoded in two parts as prefix and suffix. The prefix is an integer between 0 and 15 and encoded with a variable –size code specified by the standard. The length of the suffix part is determined by parameter suffix length and this parameter can have values between 0 and 6 and adapted while the non zero coefficient are being located and encoded. Parameter total-zero is the number of zeros preceding the last nonzero coefficient, and it will be encoded with a variable –size code.

III. CAAC BASED MULTIVIEW COMPRESSION

In the proposed Multi view CAAC-based Coding, the view is adapted such that it matches the features of MVC as closely as possible. The biggest change in the realization of MVC was enabling a certain view to be inter-view predicted from an earlier decoded view. In MVC, hierarchical coding is enabled.

IV. RESULT AND CONCLUSION

In this paper, a new video compression schemes based on HEVC is presented. For each of the schemes, a relevant test configuration will be described which covers a broad range of application scenarios. Possible applications covered by the test results vary from IPTV to broadcast TV, Personal Video Recording (PVR), and video on optical storage. Outside the scope of these test results are video conferencing and similar scenarios. Because of the technological relevance of compressing two and three views, our proposed multi view compression is evaluated for these cases. The implementation of our proposed multi view compression scheme is based on HEVC Model (HM) 3.0 as developed in JCT-VC during HEVC standardization. The sequences on which the tests are run, correspond to the test sequences used for the Call for Proposals (CfP) Video[5] issued by MPEG. To simulate realistic scenarios as described before, a random access period of 16 frames is configured. This approximately corresponds with a random access of 0.5s up to 0.7s. As a tradeoff between compression efficiency and complexity, four reference frames and a GOB size of eight are used. To challenge the proposed multi view scheme, hierarchical coding is enabled.
Furthermore, an internal bit depth of 10 is used in combination with CABAC and rate distortion optimized quantization (RDOQ). This results in a highly efficient configuration of HEVC.

Finally, to generate results in a 30 to 40 dB PSNR range, Quantization Parameters (QP) between 35 and 47 are selected in steps of 4. Additional to these decisions, it can also be observed that a very compression efficient tool for texture coding did not perform well for depth map coding. This filtering tool, called Adaptive Loop Filter (ALF) [13] only reduced the rate with 0.3% while keeping the quality constant. Therefore, it was decided to disable this tool for depth maps and obtain a significant faster encoding and decoding speed.

To facilitate the performance evaluation, the image epitome size is fixed at 80 × 64 sequences, and the image patch is set to be 8×8. To obtain the image patch set for the coherence term, the image patch sampling period is fixed at 0.5, which means that the patches are sampled at an interval of (4, 0), (0, 4), or (4, 4). The 8 × 8 blocks partitioned by video compression are regarded as the patch sets. We calculate the average PSNR gain at the same bit rate and the equivalent average bit rate savings at the same PSNR relative to H.264/AVC intra coding using the method of [9], with the rate for epitomic priors added to the total rate for each sequence. The Orbi, Interview, Ballet, and Break-dance video test sequences are encoded using the H.264/AVC video coding standard (JM reference software Version 16.0). The selected test sequences cover a range of texture and motion characteristics and the depth maps of these sequences are obtained using a depth-range camera (Orbi and Interview) and a stereo matching algorithm (Ballet and Break dance). Two-second long sequences (i.e., 250 frames from Orbi and Interview sequences and 150 frames from Ballet and Break dance sequences) are encoded, using QP values 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50. An I frame is encoded by every one second. Slices (one row of one slice) are also introduced in order to make the decoding process more robust to errors. The transmission of the encoded bit-stream over an IP core network is simulated by using IP error patterns generated for Internet experiments (7). The individual correlation coefficients (i.e., SSE, R-square) for the Orbi, Interview, Ballet, Break dance color image sequences and for all color image sequences in general are listed in Table 1.

**TABLE I**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>SSE</th>
<th>R-square</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orbi</td>
<td>0.368</td>
<td>0.9161</td>
</tr>
<tr>
<td>Interview</td>
<td>0.261</td>
<td>0.9654</td>
</tr>
<tr>
<td>Ballet</td>
<td>0.084</td>
<td>0.9528</td>
</tr>
<tr>
<td>Breakdance</td>
<td>0.131</td>
<td>0.9607</td>
</tr>
<tr>
<td>All sequence</td>
<td>1.266</td>
<td>0.9273</td>
</tr>
</tbody>
</table>

**REFERENCES**


**Abraham Mathew**, received his B.E in Electronics engineering from Bangalore university, ME in Applied Electronics from Sathyabama University,Chennai. He is currently pursuing PhD from Sathyabama University. His research interest include Video & Image Processing, Signal Processing. He is having more than ten years of experience in teaching. He is a life member of IETE.

**Dr.R.Radhakrishnan**, received his B.E in Electronics and communication engineering from Bharathidasan university, ME in Applied Electronics from Bharathiar University, and PhD from Anna University,Chennai. He is currently working as the Principal of Sri Nandhanam Engineering College, Vellore. He has published more than ten research papers in various international and national journals.