Color Space Compatible Coding Framework for YUV422 Video Coding

Lujun Yuan[1], Guobin Shen[2], Feng Wu[2], Shipeng Li[2], Wen Gao[1]

[1]Institute of Computing Technology, Chinese Academy of Science, Beijing
[2]Microsoft Research Asia, Beijing

{ljyuan, wgao}@jdl.ac.cn, {jackyshen, fengwu, spli}@microsoft.com

ABSTRACT

By keeping more chrominance components, YUV422 format provides much better fidelity and visual quality than YUV420 format and is used extensively in high end video applications such as studio and video archiving. A new video coding framework for YUV422 video sources is proposed in this paper. The proposed framework features color space compatibility to the more popular YUV420 syntax. Specifically, the chrominance components are separately into two parts and coded differently. The first part, together with luminance component, conforms to the YUV420 layout and is coded the same way as a normal YUV420 video to produce a YUV420-compatible base bit stream. The second part, i.e., the remaining chrominance components, is coded to generate an enhancement chrominance bitstream for improving the chrominance quality. This is in sharp contrast to the YUV422 coding method of MPEG-2/4 standards where all the chrominance are coded together and in the same way. Consequently, the resulting YUV422 bitstream can be easily converted to a YUV420 bitstream by simple truncation instead of undergoing an expensive transcoding process. New coding modes are also introduced for more efficient coding of the enhancing chrominance components. Performance-wise, the new framework also outperforms existing methods thanks to the new coding modes introduced.

1. INTRODUCTION

With the development of digital video processing and network communications, trends in consumer electronics market reveal that people are becoming more concerning about video quality. YUV422 video meets this trend. YUV422 format provides much better fidelity and visual quality than YUV420 since it keeps more chrominance information. In fact, YUV422 format is widely used in studios and in archiving applications. Obviously, a prerequisite condition for adopting YUV422 format is that users must have YUV422 capable video equipments. While the new devices are emerging, most commercial digital TV Sets or set-top boxes for home use can only accept YUV420 input. It is impractical for TV stations to broadcast both YUV422 and YUV420 bitstream simultaneously for the same content since the channel is rare yet expensive resource. As a result, the YUV422 bitstream must be transcoded to a YUV420 bitstream either at the server side or at the client side. The latter is more practical in the future since there may be customers that subscribed to higher video quality. Unfortunately, if the YUV422 video sources are coded by traditional MPEG-2 or MPEG-4 standard [1, 2], the transcoding process is very expensive and generally leads to quality degradation.

Scalable video coding, ranging from traditional layered coding to more advanced FGS, PFGS [3~5], has received a lot of attention since last decade because of their excellent capability in handling different (or dynamic) scenarios. Motivated by the philosophy behind scalable video coding, a new color space compatible video coding framework for YUV422 format video coding is proposed in this paper. By color space compatibility we mean that the resulting YUV422 bitstream can be easily adapted to the color space that a specific decoder may accept.

Both the MPEG-2 and MPEG-4 standards support YUV422 format. In these standards, all the chrominance components are coded together and in the same way. On the contrary, in the proposed video coding framework, the chrominance components are coded in a different way. Specifically, the chrominance components are separated into two parts: one part, together with luminance component, conforms to the YUV420 layout. The other part is the remaining chrominance components. The first part is then treated as a normal YUV420 video to generate a base bitstream whose syntax is compatible with corresponding YUV420 syntax. The second part is coded separately to produce an enhancement chrominance bitstream which is used to enhance the chrominance quality. Clearly, a YUV422 bitstream can be easily truncated to yield a YUV420 bitstream. No transcoding process is needed. We also introduce new coding modes in this framework for efficient coding of the enhancing chrominance components.

The rest of paper is organized as follows. Section 2 first analyzes the traditional YUV422 format video coding and then proposes a new video coding framework for YUV422 video. Experimental results are given in Section 3. Finally, Section 4 concludes this paper.

2. PROPOSED VIDEO CODING FRAMEWORK

Both YUV420 and YUV422 formats are supported in MPEG-2 and MPEG-4 standards. The positioning of luminance and chrominance samples of YUV420 and YUV422 and their corresponding macroblock structures are illustrated in Figure 1. Clearly, the vertical resolution of YUV422 chrominance component is twice that of the YUV420. The internal organization of chrominance blocks of YUV422 in MPEG-2/4 is illustrated in Figure 2-A. Each chrominance block is composed of continuous lines from a chrominance picture. It can be observed that for the YUV422 bitstream generated by MPEG-2/4 to be decodable by a YUV420 decoder, a transcoding process (including decoding the chrominance components, down sampling, and re-encoding the chrominance components) is indispensable. One can not simply skip or discard the bitstream of every other chrominance block (say, U2 and V2) because of the misalignment between the pixels in U1 and V1 and those in luminance blocks Y3 and Y4. In fact, pixels in U1 and V1 are only associated to luminance blocks Y1 and Y2.

To achieve color space compability, the chrominance blocks are composed from alternate lines of a chrominance picture, as illustrated in Figure 2-B. This subtle modification to the chrominance block formation method drastically changes the association between luminance blocks and the resulting chrominance blocks. Clearly, U1 and V1 are associated to all the four luminance blocks and their physical positioning confirms to YUV420 layout[1]. As a result, a YUV420 decoder can simply discard every other chrominance blocks of a YUV422 bitstream and the resulting video still can be displayed correctly. Note that, U2 and V2 are also associated with all the four luminance blocks. If U2 and V2 are available, the chrominance signal of the resulting video will be enhanced uniformly. Therefore, they are called enhancement UV (EUV) hereafter.

2.1. Coding framework

According to above analysis, YUV422 video sources can be separate two layers: a base layer consisting of the luminance components and the chrominance components (e.g., U1 and V1) that conform to YUV420 layout, and an enhancement layer consisting of the remaining chrominance components (e.g., U2 and V2). For the base layer, we adopt traditional video coding method such as MPEG-2 or MPEG-4 to generate a YUV420-compatible base layer bitstream. For the enhancement layer, there is no restriction on the coding method for the enhancement chrominance layer. We choose similar coding technology as that of the base layer with additional new coding modes. In this way, we can reuse most of the available modules which is very important for both software and hard decoders. The block diagrams of the proposed color space compatible YUV422 codec is shown in Figure 4, 5.

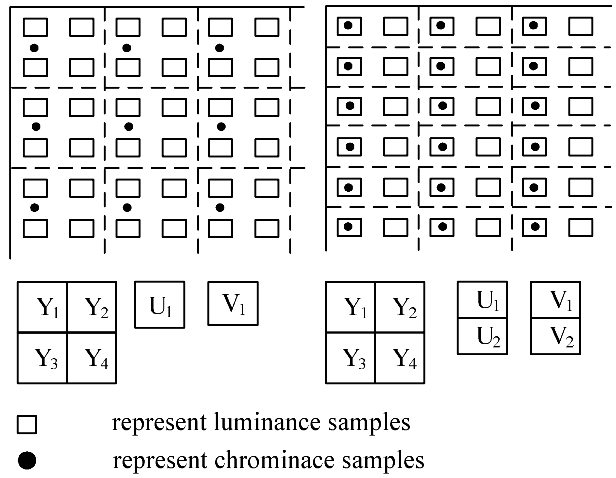

Figure 1. The position of luminance and chrominance samples of YUV420 and YUV422 format and their macroblock structures

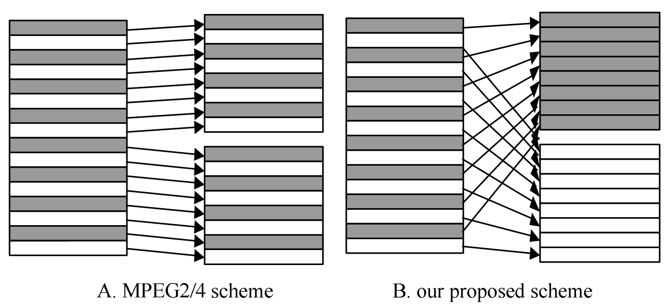

Figure 2. The internal organization of the chrominance block for YUV422 format

2.1.1 Proposed YUV422 encoder

The proposed YUV422 encoder is shown in Figure 3. The format of the input video is YUV422. It first goes through a "chroma separation" process which manipulates the layout of the chrominance components and form the YUV420 input and the enhancement chrominance (EUV) input. The flow chart for encoding the YUV420 (within dotted box) is exactly the same as common YUV420 encoder. The resulting base bitstream (B-BS) is fully syntactically compatible with common YUV420 decoder. The flow chart for the EUV branch is more or less the same as that of the YUV420 branch except two places: firstly, no ME (motion estimation) is performed on EUV (or UV in general). All chrominance components use MVs

[1] In this paper, U1 and V1 are obtained through simple 2:1 down sampling from the chrominance of YUV422. No filtering process is applied and it will be our future work.

that are derived from Y component. Secondly, there is a mode selection for EUV (see switch S1 in Figure 3). For EUV, a block can either be temporally predicted using MCP (motion compensation prediction) from *previous* reference EUV data or be "spatially" predicted from the base UV component in the *current* YUV420 frame. The resulting enhancement bitstream (E-BS) is then combined with the B-BS to form the final output bitstream.

Basically, there are two ways to organize the resulting base bitstream and enhancement bitstream. One way is to embed the enhancement bitstream into the base bitstream and the other is to concatenate the enhancement bitstream to the end of the base bitstream. For the first method, a YUV422 decoder can sequentially decode all the frames while a YUV420 decoder needs to skip the enhancement bitstream for each frame, it is suitable for streaming or broadcasting applications. For the second method, a YUV422 decoder has to seek and decode the base bitstream and the enhancement bitstream before proceeds to the next frame while a YUV420 decoder can sequentially decode all the frames without touching the enhancement bitstream, it is suitable for download-and-play applications.

2.1.2 Proposed YUV422 decoder

The proposed YUV422 decoder is shown in Figure 4. The decoding process of the YUV422 decoder basically reverses that of the encoder. That is, the input bit stream is first parsed to form B-BS and E-BS. The flow chart of B-BS branch is exactly the same as a common YUV420 decoder. For E-BS, the flow chart is the same as that of the B-BS except that the difference data is added to either its MCP or its SP, as determined by the mode information, to form the reconstructed EUV frame. After the YUV420 frame and EUV frame are ready, a chrominance composition process is applied to form the final YUV422 output which can be used to render. An additional color space conversion module (to convert YUV422 to RGB) is necessary if the display accepts only RGB input.

2.2. New coding mode

Due to the new EUV block formation method, two possible predictions are available when coding EUV blocks. Similar to normal YUV420 coding (and YUV422 coding in MPEG-2/4 as well), EUV blocks can use the EUV blocks in the previous reconstruct frame as reference. We call it temporal prediction (TP). At the same time, EUV may also use the base UV (chrominance components in the YUV420 layout) from the *same* frame as reference. We call it spatial prediction (SP). SP is very efficient since EUV is intrinsically highly correlated to the corresponding base UV, thanks to the new chrominance block organization method. Therefore, EUV may have three coding modes: Intra, TP and SP. It is the new SP coding mode that enables the proposed framework to outperform other existing methods in terms of coding efficiency for the YUV422 video.

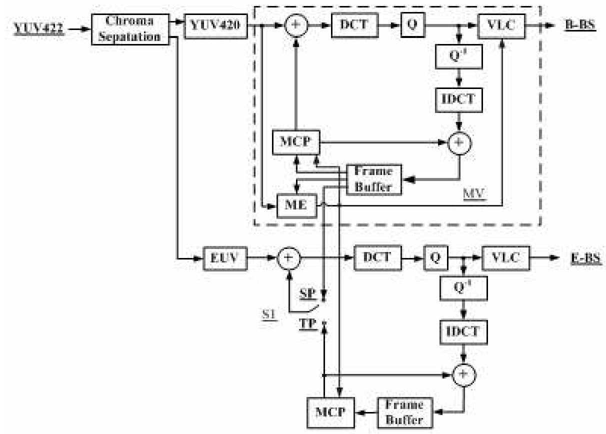

Figure 3. The proposed YUV422 encoder

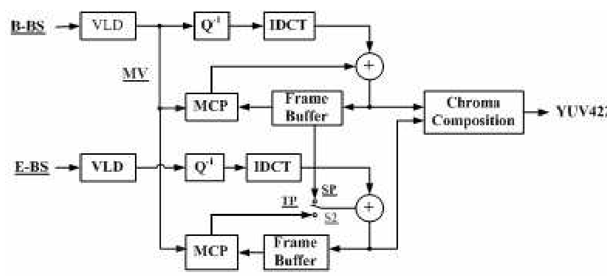

Figure 4. The proposed YUV422 decoder

3. EXPERIMENTAL RESULTS

Besides the new functionality (i.e., the color space compability), the proposed coding framework actually results in better coding efficiency as compared to normal YUV422 coding because of the new coding modes. In our experiments, two high definition JVT test video sequences, Night and Preakness, were used. The resolution of the sequence is 1280×720p, and the frame rate is 30. Firstly, the video sequences are coded with MPEG-2 encoder [6] with the following fixed quantizer scales, 4, 6, 8, 10, 12, 14, 16, 18, 20, and 22. Secondly, the video sequences are coded with the proposed YUV422 encoder, which is modified on the same MPEG-2 encoder, with the same fixed quantizer scales. The GOP length is 15 and no B picture is used. The Rate-PSNR curves for these two encoding methods are depicted in Figure 5 and 6.

For Night sequence, our proposed encoder has similar coding efficiency as the standard MPEG-2 encoder. For the Preakness sequence, our encoder consistently outperforms the standard MPEG-2 encoder in the whole bit rate range. The reason for these results is two-fold. First of all, the proposed chrominance formation method generally leads to slight loss of DCT transform gain because the intra-block correlation of the chrominance blocks is reduced. On the other hand, the prediction gain resulting from the new SP coding mode often outweighs the loss in transform gain. It is observed from experiments that a significant percentage of blocks are coded with SP mode for Preakness sequence. This accounts for the large performance gain for this sequence. However, for Night sequences, the percentage of SP mode is much less and the result is just comparable to the normal MPEG-2 encoder.

4. CONCLUSIONS

In this paper, a new video coding framework for YUV422 format video is proposed. The resulting bitstream be easily adapted to YUV420 color space by simple bitstream truncation instead of an expensive transcoding process which would otherwise be indispensable if traditional method of YUV422 coding methods such as MPEG-2/4 were used. To achieve the color space compatibility, the chrominance components of YUV422 video are organized in a different way to ensure the correct association between the chrominance pixels and the luminance pixels. Two of so-organized chrominance blocks, together with the four luminance blocks, are treated as a YUV420 video to generate a YUV420-compatible bitstream and the rest two chrominance blocks is coded to yield an enhancement chrominance bitstream. New coding modes of the enhancement chrominance components are also presented. Besides the new functionality, the overall coding efficiency of the new framework is also improved as compared with corresponding traditional method. Our future work includes applying filtering when obtaining YUV420-compatible chrominance out of YUV422 chrominance and enabling B-frames to get more experimental results.
The similar idea can be extended to color space compatible coding of YUV444 video straightforwardly.

5. REFERENCES

[1] ISO/IEC 13818-2: "Information technology–Generic coding for moving pictures and associated audio information: Video," ISO/IEC JTC1/SC29/WG11, Nov. 1994

[2] ISO/IEC 14496-2: "Coding of audio visual objects: Visual," ISO/IEC JTC1/SC29/WG11, Nov 2000

[3] M. Ghanbari, "Two-layer coding of video signals for VBR networks," IEEE J. SAC, Vol. 7, No. 5, June 1989, pp 771-781

[4] W. Li, "Fine granularity scalability in MPEG-4 for streaming video," ISCAS 2000, Geneva, Switzerland, May 28-31, 2000

[5] F. Wu, S. Li, Y.-Q. Zhang, "A framework for efficient progressive fine granularity scalable video coding", IEEE trans. Circuits and Systems for Video Technology, vol. 11, no. 3, 332-344, 2001.

[6] MPEG Test Model 5, ISO/IEC JTC/SC29/WG11 Document. April, 1993

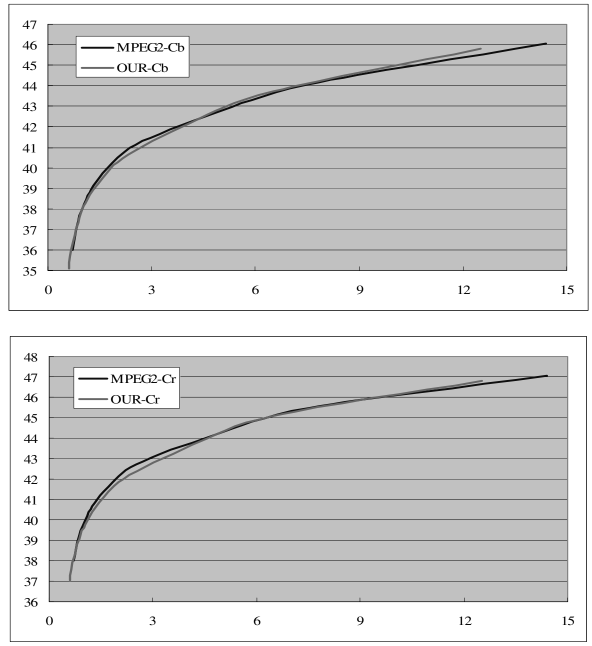

Figure 5. Rate-PSNR curves for chrominance components for Night sequence

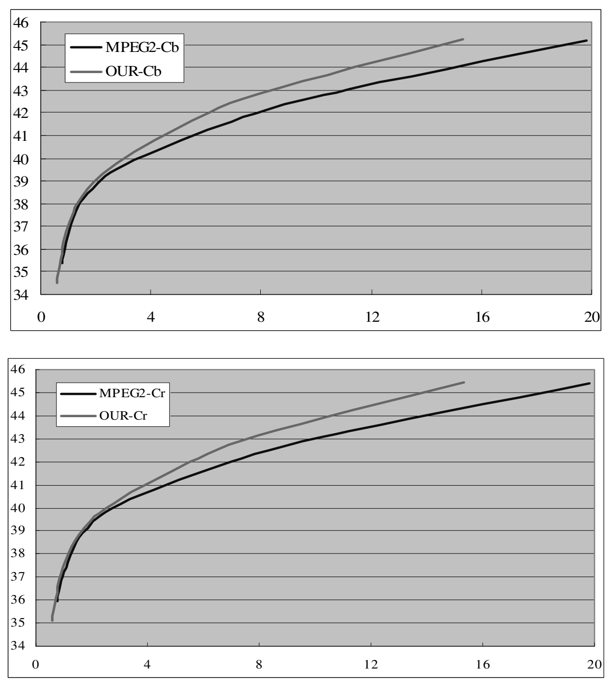

Figure 6. Rate-PSNR curves for chrominance components for Preakness sequence