Multi View Streaming System using 3D Technology

[1]Renuka S Dond, [2]S V. Athawale

[1]T.E. Student, [2]Assistant Professor
Department of Computer Engineering
AISSMS College of Engineering, Kennedy Road, Pune-411001, India

***Abstract*—Two eyes are always better than one and watching is always interesting than reading. Keeping this concept in mind, the visual industry has been keen on developing itself so as to meet the viewers' expectation and show them that effects that they can only imagine in their dreams. From first silent cinema to 4K and 3D cinema that are produced now, the visual industry has evolved a lot. New technologies are introduced and implemented every day. Not just the _lm industry but Phones have also changed from normal cameras into 360-degree experiential recording devices. The technology relies on camera that follows method algorithm that can synthesize new views both with rotational and translational motion of the viewpoint. This enables the ability to perform VR playback of input monoscopic 360-videos _les in full stereo. This seminar focuses on the most upcoming and trending technology that is 3D production technology. By running a complete multi view streaming system we can nourish video experiences without using 3D glasses and VR goggles and take the VR to a next level.**

***Index Terms*- Virtual Reality, Multi view streaming system, DASH.**

1. INTRODUCTION

Current video transmission systems only transmit a single camera view or at most two views for 3D transmissions while video coding standards for the transmission of multiple views have been around for many years. The main reasons why still only single views are transmitted are the high complexity of these video codec, high bandwidth requirement for uncompressed transmission and high bandwidth requirement for insufficiently compressed transmission. The use of multiple views in video transmission however, can lead to a much higher immersion for the user with the help of view interpolation for interactive viewpoint changes or gaze correction to simulate eye contact between conversation partners.

Various coding standards have emerged to optimize the encoding methods in a vital way. Today most widely adopted format for video distribution is H.264, a successor of MPEG-2. H.264 is a widely accepted video codec [1] for Terrestrial, Cable, Satellite and IPTV broadcast which is commonly used across for videoconferencing, mobile video, media players, video chat etc. But 4k technology is still limited and also challenging due to mismatched resolution, new frame consideration, enormous bandwidth and display compatibility [2]. H.265 is competing with next generation video compression formats that are more efficient than H.264 [5]. H.265 named as HEVC (High Efficiency Video Coding) format developed by Video Coding Experts Group (VCEG) and the Moving Picture Experts Group (MPEG).

2. ARCHITECTURE REQUIREMENT

Basic structure of H.265 is similar to H.264. HEVC contains incremental improvement such as flexible partitioning from large to small partition sizes, larger flexibility in transforming block size, better prediction modes, advanced interpolation and de-blocking filters [3]. Shows the architecture of H.265. Larger block improves coding efficiency for higher resolution since group of pixels have same properties. It flexibly adopt smaller transform for impulsive signal and large block size for stationary signal [4]. It has built in concurrent tools to process pixel rate, bit rate and motion estimation. Additional feature includes free-view point video, 360 degree video, augmented reality and 3D movies. Application of combined scalable and multiview HEVC coding includes scalable stereoscopic video, mixed resolution with multiview coding. At architectural level implementation with Dynamic Adaptive Streaming over HTTP DASH, it provides improved level of design to store set of parameters to change dynamically during run time, DASH techniques improves streaming by transferring bitstream with respect to the network traffic, its coupled with HEVC to show better result.

It supports all major instruction sets including AVX/AVX2 andFMA3/FMA4. HEVC comes in two flavors, main profile for conventional video supports 8-bit color data and main 10 profile for 10-bit color data.

3. METHODOLOGY

Our optimization method is based around a system that streams the output of a multiview capturing array over a bandwidth restricted network to a client device using a H.264/MVC codec. Figure 1 depicts the major steps in our system. The incoming multiview video sequence contains the frames of a number of cameras at known positions and their calibration data. The optimization step requires information about the available bandwidth on the transmission channel, information about the capabilities of the video encoder used on the server, most importantly about the relationship between different data rates and the resulting quality. In addition some information about the view interpolation algorithm on the client is needed. We assume it to be in the form of a characteristic curve as described below. Apart from the channel bandwidth, all information remains constant for the duration of the stream and can therefore be exchanged before the transmission starts. The encoder compresses the chosen set of views with a set of encoding parameters that optimizes the quality for the given number of views when using the available data rate.

The result is then transmitted to the client. On the client the sequence is decoded. The views which were dropped during the optimization are reconstructed using the view interpolation algorithm whose details were presented to the server. After this all views from the original video are available on the client ready for display or further processing.

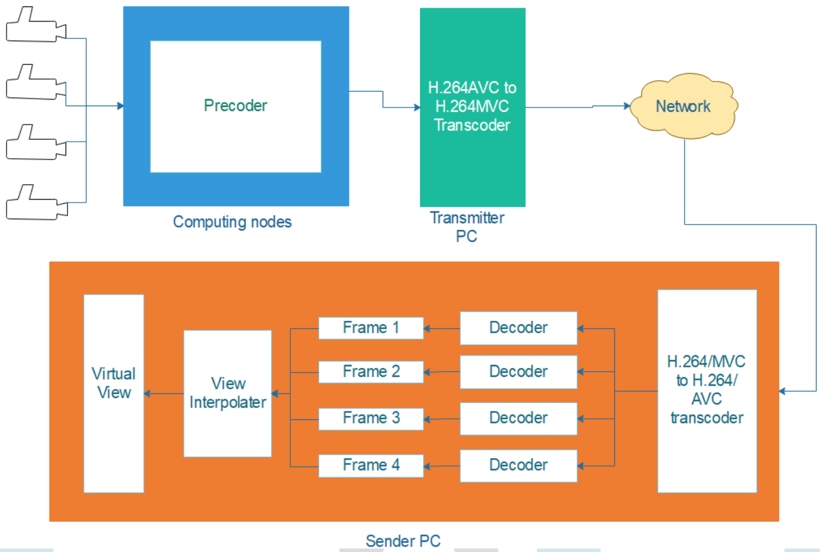

Figure 1: System Architecture

In this paper we studied a complete real-time multi-view video streaming system with view interpolation capabilities running on standard pc hardware. The system consists of the capturing/encoding unit based on a distributed H.264/MVC encoder with five cameras, partnered with five small computing nodes and one desktop PC and the receiving unit with view interpolation running on another desktop computer supported by a dedicated GPU H.264 was chosen due to the wider availability of hardware encoding and decoding support in current CPUs and GPUs. The structural similarities with HEVC offer the possibility to adapt all required algorithms to the newer format, once better hardware support on mainstream hardware is available. In order to achieve the real-time performance the videos are captured by the cameras, pre-encoded into H.264 in the computing nodes and sent to the transmitter PC using raw UDP.

To guarantee the generated video streams are in sync, the internal clocks in the nodes are synchronized via PTP and the frame rate produced by the cameras is equalized in software. By duplicating or dropping out-of-sync frames we limit the absolute maximum deviation of the capture times to half a frame. From there each pre-encoded view is transmitted to the desktop machine using raw UDP streams over a closed network. On the encoding desktop machine the streams are received and the algorithm described in [6], implemented in a custom GStreamer plugin, is used to combine the separate video streams into a single standard-conforming H.264/MVC stream which is then transmitted to the receiver machine via RTSP. This novel encoding approach is very scalable because big parts of the complexity are offloaded to the computing nodes and the remaining complexity of combining the precoded streams into a single multi-view stream is easily handelable on a modern CPU. The role of the computing nodes can be taken by any computer capable of H.264 singleview encoding, a camera connection and a wired network connection.

Today, this could be anything from mobile SoCs over single board computers like the Raspberry Pi to full-sized PCs. This makes the capturing side very scalable at a low cost for the user by simply adding more computing nodes with attached cameras to the system.
On the receiver side the incoming data stream is separated into single H.264 streams again applying the reverse of the algorithm used in the sender. These streams are then decoded using hardware-accelerated H.264 decoders present in the CPU of the receiver machine to get access to all transmitted frames of all cameras. The decoded frames are fed into the view interpolation algorithm [7] running on the GPU and the resulting virtual view, whose position can be interactively changed, is then displayed on screen.

In contrast to other multi-view streaming approaches, our system does not require the additional support of pre-computed depth-maps or any other scene geometry. This makes it independent from costly and often noisy depth capturing cameras. Although this increases the required computing power at the receiver because the depth information has to be estimated by the view interpolation algorithm, it minimizes the perceived delay when the virtual viewpoint is changed.

4. RESULTS

In its current version it is therefore not well suitable as a replacement for existing video communication systems as they require a lower system delay. But by redesigning and optimizing the buffers as well as the interfaces between the stages the delay can be lowered to acceptable levels. The H.264/MVC encoder in this setup is distributed over a total of six independent computers. As multi-view codecs assume that the images from all views have been taken at the same time and the view interpolator relies on this fact to make good estimates about the depth map of the virtual view. To make sure the input to the transcoder stages is in sync we synchronized the internal clocks of all computers in the receiver stage via PTP. This method has been proven to be effective when the system was displayed at CeBIT 2017 and stayed in sync within one frame for over 10 hours repeatedly.

CONCLUSION

HEVC is the next generation technology to provide solution with growing requirement to support mixed format content of video. Today multicore architecture is commonly used to improve performance with low cost and power. HEVC supports parallel processing architecture and tools for multicore. Current trend shows that HEVC will be the market gainer to support higher resolution 4k and 8k in effective way. Leading video providers, content providers and broadcast media support HEVC by adding this decoder in their products. Next is VP9 with greater adoption since it is open source it suits well with commercial applications. When we compare HEVC and VP9, HEVC shows optimized and improved result than vp9 [5].

A real-time multi-view streaming system can be implemented on standard pc hard-ware. Applying new algorithms and making use of the available computing power of modern CPUs and GPUs enables normal consumer machines to capture video streams from multiple cameras, encode them into a single standard-compliant data stream and send it to a receiver machine over a network. The receiver decodes the data stream, decodes all contained frames and feeds them into a view interpolation algorithm which creates a virtual view whose position can be adjusted on the y.

FUTURE-SCOPE

In future we can work on the gap between existing recommendations and solutions that maximize the average user satisfaction. In particular, we can show that an unequal allocation of the storage capacity among different video types as well as camera views is essential to strike for the right balance between storage cost and users satisfaction in interactive multi-view video systems.

ACKNOWLEDGMENT

This paper's success is not just the result of my efforts but of many others too. First of all I would like to thank my guide Prof. S V Athawale whose belief and trust in me made me complete this paper. Secondly I am thankful to all my classmates, friends and all those who have directly or indirectly provided their support during my paper work and the development of this paper. At the end but most importantly I would like to thank my parents and my brother for their constant encouragement and support.

References

[1] Global Media Format Report – 2016 encoding.com

[2] Rakshith.V, "A Low Bandwidth Guide to 4k VideoStreaming", International Journal of Computer Trends and Technology (IJCTT) ,Volume.25,July 2015.

[3] Mahsa T. Pourazad, Colin Doutre, Maryam Azimi, and Panos Nasiopoulos, "HEVC: The New Gold Standard for Video Compression", IEEE Electronics Magazine, 2012

[4] https://sonnati.wordpress.com/2015/01/28/h265-part-iiconsiderations-on-quality-and-state-of-the-art/

[5] Dan Grois, Detlev Marpe, Amit Mulayoff, Benaya Itzhaky, Ofer Hadar, "Performance Comparison ofH.265/MPEG-HEVC, VP9, and H.264/MPEG-AVC Encoders", IEEE 30th Picture Coding Symposium 2013.

[6] T. Lange and T. Herfet, "A distributed real-time multi-view video encoder on consumer hardware," Accepted at BMSB, 2017.

[7] T. Lange, G. Petrovic, and T. Herfet, "Real-time virtual view rendering for video communication systems using opencl," in Broadband Multimedia Systems and Broadcasting (BMSB), 2014. IEEE, 2014, pp. 1–5.