A Real-Time 3D Video Streaming System Using SRTP AND RTSP Protocol

Dr.P.Iyyanar[1], Dr.M.Arunachalam[2], Dr.Anita M Patil[3], Dr. Nilesh Uke[4], Jaya Dipti Lal[5], Dr. V. R Sonawane[6], Dr.R.Rajagopal[7]

[1]Associate Professor, Department of Information Technology, Sona College of Technology, Salem, Tamilnadu

[2]Professor/ Department of Information Technology, Sri Krishna College of Engineering and Technology, Coimbatore

[3]Assistant professor, Department of MCA, Bharati Vidyapeeth, Yashawantrao Mohite Institute of Management, karad

[4]Professor, Department of Computer Science, Trinity Academy of Engineering, Kondwa, Pune

[5]Assistant Professor, Department of E&TC, Shri G S Institute of technology and science, Indore, MP.

[6]Associate Professor, Department of Information Technology, MVPS KBT College of Engineering, Nashik

[7]Associate Professor, Department of Computer Science Engineering, Narasimha Reddy Engineering College, Hyderabad

Abstract

Video streaming has become an attractive research area where effective and network-friendly media coding methods, synchronization, bandwidth, packet loss and delay issues, delivery protocols, and interactive media players and display devices have been examined since the early 1990s. There are many open-source and consumer products available today. It can be used to create an advanced streaming environment that supports multiple type of audio and video. Following the advancement of 2D video streaming technologies, the focus is on the third dimension, mainly because of the promising improvements in 3D displays and multi-view video coding techniques. Multi-view footage includes views from different angles of the same scene that are recorded by several cameras. However, only two of these views can be watched by a specific interface at any given moment. This carries the concept of creating a streaming device that provides stereo content that can be converted to a multi-view by choosing the two viewpoints shown based on the actual user viewpoint. This paper presented the design and implementation of a secure 3D video streaming system using SRTP (Secure Real-Time Transfer Protocol) to transfer the audio and video streams and RTSP (Real-Time Streaming Protocol) protocol used to stream the video contents on the client-side.

Keywords:

3D Video streaming, RTP, SRTP, RTSP, MPEG.

Manuscript received June 5, 2022
Manuscript revised June 20, 2022
https://doi.org/10.22937/IJCSNS.2022.22.6.76

1. Introduction

There are many streaming and video messaging systems that are implemented for different applications. Point-to-point, multicast, and broadcast systems can be applications in real-time. It may be either interactive or non-interactive. A minimum end-to-end delay is required for interactive applications. The time needed to record, encrypt, transmit, decrypt, and view the end-to-end delay. For non-interactive systems, the end-to-end delay is much looser. Thus, an application's interactive state influences its architecture.

Video streamed with these applications may be encoded or pre-encoded in real-time and such choices impact the creation of a streaming program. Three criteria categorize a multimedia application: the number of media concerned, the forms of media assisted and the degree of incorporation. Multi-view video is an extension of the standard two-dimensional video, which incorporates several views from the same scene at any given moment. A perfect multi-view device enables the user to see at every angle the spectator selects at a genuine 3D stereo series. In this research, the system proposes stereo streaming that can be applied to a multi-view streaming device in the future. The application tends to focus on transmitting 3D video streams from a fixed angle.

H.264 Multi-View Extension Codec (MMRG), which is based on Multi-view Video Coding (MVC) technology, is easily compressed, and streamed using standard real-time protocols. Depending upon the available bandwidth or viewer devices, a selective 3D video system is designed to enable the selective transmitting of mono or stereo video [1]. It will use a device based on an independent transmission of both stereo streaming services to accomplish this goal. To have access, a system like this can be created by modifying the original platforms for regular monoscope video streaming. This system created a platform that includes a pre-encoded stereo streaming server and a media player on the Client-side to integrate the display. End-users view the stereo recording by using polarized lenses at the display phase. The receiver can view video content from multiple channels as a stereo with adequate display equipment and enough bandwidth.

1. Scope of the System

Demand for online video content is growing and becoming the primary traffic in wireless networks for streaming video services applications. A substantial share of video streaming services is currently available

Traffic on the Internet. The market of high-definition photos and high-quality smartphone users over the past decade continuously update video sources. According to Cisco's study, 82 per cent of all user Internet traffic is expected to be in the video network by 2021. Due to a significant massive amount of data, different network characteristics and user terminal needs, and the user's context, delivering three-dimensional immersive media to individual users remains a significantly challenging task.

Internet connectivity every day, anytime and anywhere, the amount and behavior of Internet use, particularly media usage, has changed with any device. Web browsing and file sharing were historically the dominant applications. Today's critical portion of the Internet traffic is in real-time streaming and social media applications (e.g., YouTube, Facebook, etc.). Future Internet requires an increased demand for streaming and 3D media through the modification and optimizing the current protocols for streaming over a hybrid network environment for future Internet applications.

Many tools and innovations are emerging in the technology industry today. 3D is considered a technology that is one of the most advanced. It can be presented in many areas such as movie making, medical therapy and engineering to show its significance. 3D technology has matured by improving usable 2D to 3D materials that provide outstanding consistency. 3D technology is available in 3D or 3D video formats. 3D video is spreading and becoming an exciting tool for smartphone and home entertainment applications.

2. Media Streaming System

The video delivery can be defined in one of two ways. One is the process of downloading files, and the other is streaming video. Possibly the simplest form of video distribution is to download. The difference between a video download and a conventional file is the length of your download [2]. Since video files are enormous, the downloading process takes more time. This approach also creates the issue of storage. The client-side needs a large storage area to download a big file while playing before playing. Video streaming is the next step. The compressed video on the sender side is split into video streaming. Then video packets received are decoded and played on the receiver end while the video still arrives in. This allows video streaming to produce and play the video simultaneously. Streaming addresses all the download process drawbacks with a low storage space and low delay.

The family of transportation protocols includes media streaming protocols UDP, TCP, RTP and RTCP [3][4]. UDP and TCP provide basic transport capabilities. It operates with RTP and RTCP. The TCP protocol retains the regulation of congestion, error control, flow control and multiplexing. While TCP supports secure transmission, UDP is suitable for media application streaming. The TCP retransmission characteristic makes delayed streaming applications not appropriate. UDP is, however, used as a video streaming transport protocol [9].

In real-time applications, RTP is an Internet standard protocol. Timestamping is provided for syncing various media streams, sequence numbers for proper ordering, payload types for classifying media package contents and source Id through the Synchronization Source Identifier (SSRC) to differentiate between different media sources.

3. 3D Video Streaming System

Recent development has led to a paradigm shift from the conventional 2D video to 3D technology of Multi-visual Video while also improving video quality and compression through standards such as high-performance video coding [5]. In multi-view, to capture the video from multiple viewpoints, cameras are mounted in predetermined locations. It will be an obstacle to provide such high-quality internet views since Multi Visual Video in various sequences from a different angle requires more bandwidth than a single view than transmitting Multi Visual Video.

3D video has become a visually attractive and expensive affordable technology with the latest stereo camera, compression, and display advancements. In Entertainment and Medical, the 3D video service is expected to be the next big thing. It attracts both academic research and medical imaging and industry innovations. Humans view 3D video with an extra three-dimensional depth resulting in two spatial variations, each showing a one-eyed view. On the other side, the multi-view video allows multiple viewing positions. This allows the viewer to choose or adjust the location of the preferred standpoint to enjoy the video, which is not required in the front center. The typical monoscope perception or the potential stereo perception of the two opposite views are presented in each view location.

Two separate videos will be taken, each one reflecting the one eye perspective to create 3D-perception. The raw data is then coded using a codec to allow video encoding for stereo/multi-view. One of the most common candidates is the new Multi-view Video Codec (MVC) H.264, which significantly enhances the current H.264 Advanced Video Codec (AVC) by using redundancy across different points of view [8]. In the 3D stereoscopic scenario, a reference view is selected from one of the two movies (e.g., the left video) encoded in the standard H.264/AVC codec and examined for temporal redundancy in the view using the motion compensation (MCP) prediction. The second view is inextricably linked to the stream and the reference view, as is the newly created compensation for the second view's discrepancies (DCP).

The encoded frame data and the supplementary information (e.g., view ID and forecast mode) are encapsulated into separately decodable NALs (network abstraction layer), further embedded for real-time streaming in RTP packets. The obtained information is displayed correctly by the visualization device at the user's side. For instance, users can wear shutter glasses for stereoscope perception or install the latest hardware to show stereo/3D videos automatically.

In 3D video streaming, the synchronization and coordination of two view streams for the same video application are crucial. Unlike scalable video (SVC) and multiple description video (MDC), which allow the receiver to play the video with a significant number of layers (SVC) or descriptions (MDC), the closely coupled nature of the two streams in 3D video means that the missing data section of either of the two streams will interrupt normal playback. Communication and synchronization may be considerably more challenging in end-to-end networks since each node serves as both client and server.

3.1 Need for Proposed System

Although Multi-View Video (MVV) and Multi-View plus Depth (MVD) technologies are fascinating, the development of sufficient frameworks to allow end-user distribution over best-in-class network efforts does not advance at the same time. The provision of MVV over-developed networks poses three significant challenges. First, MVV traffic is several times greater than conventional multimedia because it consists of multiple camera-captured video sequences, transmitting more data. The device also knows that the Internet is likely to experience packet loss, delay, and change in bandwidth, which is very important for MVV content. Second, end-user devices' processing capacity, monitor and access connectivity capacities vary, requiring flexible mechanisms to change bandwidth variance when traversing network paths [10].

Second, the video production pattern is to raise the resolution to give the viewer more information. The higher resolution, however, means that the content is even higher. More than two high-resolution views must be transmitted, particularly for the 3D multi-view monitor. A more performance encoder should also be taken into consideration. Third, Codec speed for 3D is too slow, so 3D video output cannot be compressed in real-time. The 3D video can also be delivered by video on demand [12].

The above challenges prompted us to propose a dynamic rate adaptation method and its related multi-view 3D content transmission rate-distortion model to resolve the problem of the differing network bandwidth for Internet video users. The transmission system architectures have been developed using two state-of-the-art techniques: high-performance video encoding (HEVC) and dynamic change streaming for multi-view video transmission using Real-Time Streaming (RTSP).

4.2 Comparison of Various 3D Content Formats

While conventional 3D stereo technology is much easier to supply in the grid because of its low amount of data demand, it poses many problems with the consistency of the experience. The frustration created by the lack of depth and the creation of a 3-D glass wear display would be limited by adopting various users, correcting the 3D display perspective, and bulky glass wear. A multi-view monitor should be used to cope with the 3D stereo 3D screens, but the difficulty arises in the non-dedicated network during different views. Table 1 provides a summary of multiple 3D content types. The multi-visual plus 3D content profile format draws greater interest from the industries and research institutes worldwide [13].

The result is that the multi-view plus depth map is the most effective format for 3D Multi-View Content to be distributed to non-dedicated networks among all 3D content. A limited number of views plus geometry details allow multiple viewpoints within the defined baseline distance to be produced.

Table 1: Comparison of various 3D content formats

3D content Type	Description	Pros	Cons and weakness
Stereoscopic 3D	Linked or compatible views for each eye side by side or up and down	Present 2D transmission networks are very similar	Unable to change depth to various users and glass wear.
View plus Depth	2 stereo views and their respective depth can be made from one view.	In contrast to 3D stereo, more compression quality	Uncomfortable still glass wear, and a single view will contribute to an occlusion area.
Multi-view 3D	Views from various angles are captured in one array	The auto stereoscopic monitor can be assisted without glass wear and more view angles and client depth perception.	Big bandwidth demand because of the large size of the content in the 3D view format.
Multi-view plus depth	Various views of their corresponding depth from various angles	Non-glass wear show help and fewer views are required.	Compared to a 3D Multi View, much less bandwidth is required.

4. Proposed Method For 3D Video Streaming System

Video applications that are three dimensional (3D) have been established in the entertainment business for quite some time, with big blockbusters consistently earning more than their two-dimensional counterparts. With the introduction of 3D sports channels, blue-ray discs, and now the Internet, 3D videos are becoming more popular in the home. Internet distribution of 3D movies is anticipated to be the most commonly used medium of delivery in the future, owing to its inherent flexibility, diverse nature, and growing bandwidth.

It is discussed in this chapter how to transmit 3D movies encoded using the multi-view video coding (MVC) version of the H.264/AVC (Advanced Video Coding) standard utilizing the Real-time Transport Protocol's multi-session transmission (MST) capability (RTP). Depending

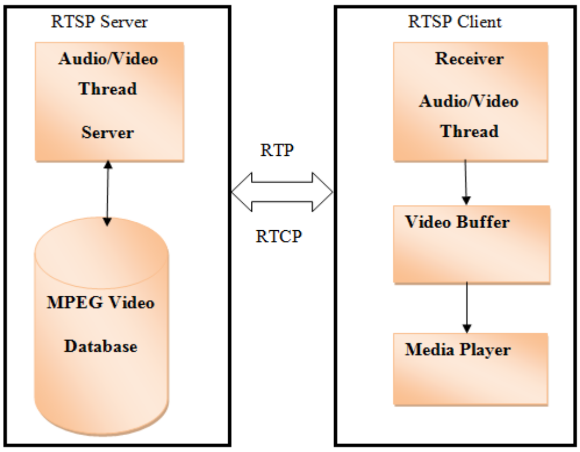

Figure 1: Architecture of 3D video streaming system

on the number of priority levels needed, an RTP session is created for each client. If a frame cannot be duplicated, it is essential to have one. Sessions with a low priority level may be dropped if the network circumstances are not favorable. The number of streams may be changed depending on the network circumstances and available bandwidth, providing the MVC encoded stream the flexibility of simulcast without sacrificing encoder efficiency, which is advantageous in certain situations.

4.1 Video Transmission Over RTP

The suggested system incorporates a client-server application for video streaming that has been developed. RTSP servers receive new connections from RTSP client requests, and they are responsible for sending video to those RTSP clients that have requested it. For packet distribution, it is necessary to schedule the video material. The RTSP server manages the bandwidth required to live video streaming in a dynamic environment [6]. The RTSP client receives video packets from multimedia networks via the RTP transport protocol, and all of the boxes are combined to form video frames. It also decodes the video information and plays the video frames via the player to complete the process. Delivering the real-time packet is accomplished via the use of the RTP channel. As shown in Figure 1, many components make up the video streaming system.

The server software carries out the following functions:

- Acknowledging connection requests from customers.
- Delivering live video streams to clients; scheduling packet delivery.
- Managing bandwidth; and adapting to a changing environment.

According to the regulations outlined in the relevant draught, streaming MPEG videos are carried out in the proposed system. When it comes to real-time video streaming, the RTP protocol provides the ability to packet multiple payload types in appropriate forms for transmission over the Internet. MPEG file structures must be considered when developing this algorithm. The Internet Engineering Task Force defined packet rules for each supported payload type and published them as Internet draughts. These rules are intended to decrease the dependencies between packets while increasing decodable data available at the receiver end [7]. RTP packets come in a variety of sizes. The maximum packet size is limited by the top transmission unit (MTU) on the LAN or WAN. Before the payload data, the application layer places MPEG and RTP headers on the data stream. The addition of UDP and IP headers occurs in the lower levels. Figure 2 depicts the structure of an RTP packet that has been sent out of the network.

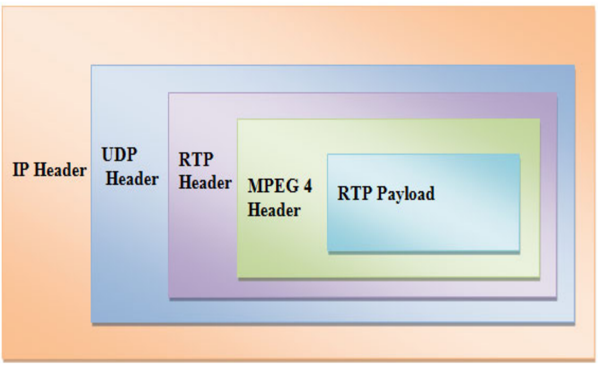

Figure 2: Total packet formats for a video stream

An innovative concept for an effective real-time audio/video streaming system based on the RTP/RTSP protocol. RTSP client and server models are used as audio/video communication endpoints in real-time transmission using the RTP protocol in the proposed system, which is based on true streaming and uses the RTP protocol for information. Cryptographic techniques for secure communication in video streaming systems are based on the DES, triple DES, and AES cryptographic algorithms, which are used to encrypt and decrypt the data streams in both the encoding and decrypting processes. Triple-DES with 192-bit keys is a more secure encryption algorithm than DES with a 56-bit key. AES with a 128-bit key is designed for use in video signaling systems for real-time video processing, and it is also used in encryption. The SRTP protocol is used to transport an audio/video stream securely. This protocol uses the default encryption technique of AES-CBC and AES-CM mode for encrypting and decrypting video streams, as suggested by RFC 3711, for encrypting and decrypting video streams. The secure video communication system is based on the SRTP protocol in this proposed system, which provides a secure video communication system.

4.2 3D Video Transmission OVER SRTP

The combination of inter-view prediction and hierarchical pictures in the Multi-view Video Coding (MVC) addition to the H.264 Advanced Video Coding (AVC) standard leads to increased compression efficiency for multi-view video [11]. In comparison to simulcast coding, the resulting bitstream is much more prone to transmission errors. The purpose of this research is to explain the idea of transmitting MVC-encoded multi-view video across several RTP sessions and to conduct an experimental assessment of its practical effectiveness.

The findings show that by segmenting the transmission of structures according to their importance degree and protecting the elementary streams, transmission faults may be minimized. Using multi-session communication, this system can deliver a more adaptable bitstream in which the number of streams can be adjusted in response to network conditions and, in extreme cases, all but the base-view stream can be dropped to avoid further network congestion while still delivering at least one high-quality view to the end-user, as described in the following section. We have demonstrated in our experiments that confining errors to the RTP session containing the hierarchical B frames mitigates the impact of mistakes on sessions containing the base view and inter-view P frames; error rates of up to 10% have no discernible effect on the received video quality in multi-session transmission.

MVC significantly lowers the amount of data sent by using inter-view redundancies between consecutive video sequences and the regular temporal and spatial redundancies in each view to achieving this reduction. This, on the other hand, makes it more susceptible to transmission mistakes. Keyframe errors have the potential to make the whole bitstream unusable. The inter-view interdependence also implies that when mistakes occur in frames utilized by other views for motion compensation prediction, the faults spread to the different views. The transmission of MVC bitstreams across multiple streams depending on priority and view changes improves the overall quality of 3D video transmissions.

Multi-session transmission (MST) provides end-users and systems with the option to remove views in crowded networks to devote more bandwidth to the base view, resulting in improved quality of experience (QoE). MST may also be utilized to enjoy the benefits of simulcasting without sacrificing the advantages of inter-view similarity exploitation, as seen in the following example [14]. It does, however, cause synchronization difficulties under erroneous channel circumstances when packets arrive outside of the bounds of the channel. This issue has been addressed by using the MVC time-first frame order to resynchronize the frames in the received sequence.

5. Result and Performance Analysis

Multi-view movies include a significant amount of data that is encoded in an interdependent manner. Errors in a frame in one view may spread to other ideas if they are not corrected. When the subsequent frames are bi-predicted from the frames where the mistakes occurred, the magnitude of the impact of these errors may be much more significant. As shown in this article, bi-predicted structures are extensively used in multi-view videos. Consequently, they are more vulnerable to error propagation because of defects in reference frames. Protecting the reference frames may help to enhance the overall quality of multi-view video footage by reducing pixilation. When separating the reference frames from the bi-predicted frames to transmit them across different sessions, the issue of out-of-order arrival of the frames is introduced into the system. MVC, as previously mentioned, uses time-first coding, which implies that it must receive the frames in the correct sequence to decode them. It is almost likely that frames will arrive out of order in a multi-view video during a multi-session transmission of the video. Frames will be sent ahead of those that are required before it.

Because RTP includes features such as timestamps and control mechanisms that may be used to create the synchronization required at the endpoint, it is the best protocol to employ in this situation. This structure ensures that the frames are placed in the proper sequence for decoding. TCP and other conventional transport protocols, such as UDP, are not intended for real-time data delivery. According to how the video is divided, the content of a multi-session broadcast may differ from one session to the next. In this work, the segmentation is based on the significance of data in a multiscreen video about flexible transmission channels, shown by the study results. The 3-view multi-view video is divided into three separate sessions, each with its soundtrack.

Consequently, if the end device is equipped with a 2D video decoder or the network is congested, the first session may be coded independently to give a base view for the second session to code independently. The end device may reject the additional images without degrading the quality of the base view or adding to network congestion by sending data it cannot decode in time for display. Consequently, each multi-view video stream becomes more adaptable. Additionally, since the bitstream extraction is complete, the decoder does not need additional bitstream extraction to display the base view. The second and third streams include all predicted and bi-predicted frames for the other pictures, while the first stream only contains predicted frames. This stream division is used to protect more critical anticipated frames from less critical bi-predicted frames. In bi-predicted frames, errors in the predicted frames may grow exponentially, but this is not true in the other way. Therefore, the stream containing the expected frames may be secured to reduce errors.

The experimental results show that multi-view video decoders can handle bi-predicted frame errors and tolerate a considerable degree of inaccuracy. Except for those frames lost due to network congestion, the above results show that almost all frames are decodable even with a 10% transmission loss rate. On the other hand, errors in bi-predicted frames increase according to the size of the GOP utilized. This requires consideration of many factors, the most critical of which is the decoder buffer. A big GOP increases the decoder's memory needs for prediction since it requires the decoder to keep more frames in memory for a more extended time. When the error rate increases, a buffer overflow occurs, which is hazardous. If a frame in this work has a mistake, the previous reference frame replaces it. On the other hand, the preceding reference frame may have been flushed from the buffer before the error, resulting in the decoder failing.

It should be emphasized that this is not a life-threatening situation and that error control techniques such as Forward Error Correction may be employed to address it (FEC). Additionally, the increased error rates highlight the need of selecting the proper GOP size for a stream. It is critical to consider coding efficiency and transmission errors while choosing the GOP structure throughout the encoding process. Multi-session information of multi-view video may improve the adaptability of the video stream to end devices while also increasing error resistance capabilities by confining errors to a particular spatial region. The experimental results shown above indicate that MST can manage error rates of up to 10% while incurring minimal frame loss when transmission errors are confined to the relatively insignificant bi-predicted frames.

6. Conclusion

Producing high-quality stereoscopic 3D material takes considerably more work than standard video footage preparation. To provide accurate depth perception and visual comfort, 3D videos must be carefully tailored to the viewing circumstances in which they will be presented to viewers. While most stereoscopic 3D material is intended for viewing in movie theatres, where viewing circumstances are consistent, adapting the same content for home television sets, desktop displays, laptops, and mobile devices needs extra modifications. We propose a novel 3D video streaming system that includes automated depth adjustments as a significant feature to solve this issue.

As electrical and computer technology continues to advance at a breakneck pace, multi-view video transmission has gained considerable interest from the academic and manufacturing sectors. 3D video communication systems must address the whole pipeline, from the video source to the display and transmission. Multi-view is the most attractive solution for 3D video streaming with motion parallax since it eliminates the requirement for 3D video perception headgear. It is conceivable that improving multi-view video for distribution and display may result in an increase in consumer demand for true 3D. Numerous challenges remain for 3D systems, from content creation to error resilience and concealment, heterogeneity, achieving true free-viewpoint television, and overall user interfaces quality. A multi-view video service offers a user a variety of perspectives based on the input received from the user. In conjunction with the current trend of moving video streaming service platforms to HTTP adaptive streaming, dynamic adaptive streaming over HTTP (DASH)-based multi-view video services have been extensively researched and developed. However, the view-switching latency in DASH-based multi-view video streaming is much higher than in other methods. The amount of time that elapses between a request for view switching and the rendering of the target view harms the overall quality of the experience. This article discusses some of these problems and offers suggestions for enhancing the overall quality of 3D video streaming throughout the transmission pipeline in multimedia networks such as the Internet that use the RTP, SRTP, and RTSP protocols.

References

1. *K. Birkos, I. Politis, A. Lykourgiotis et al., "Towards 3D video delivery over heterogeneous networks: the ROMEO approach," in Proceedings of the International Conference on Telecommunications and Multimedia (TEMU '12), pp. 60–65, IEEE, Chania, Greece, August 2012.*
2. *Iyyanar, P, Chitra, M 2014, 'Effective Scheme for Audio/Video Streaming using RTSP and RTP'. International Journal of Applied Engineering Research, ISSN 0973-4562, vol.9, no. 23, pp. 19685-19699.*
3. *H. Schulzrinne, S. Casner, R. Frederick, and V. Jacob-son, "RTP: A Transport Protocol for Real-Time Applications",IETF RFC 3550, 2003.*
4. *M. Baugher, D. McGrew, M. Naslund, E. Carrara, and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, 2004*
5. *Google previews Project Starline, a next-gen 3D video chat booth - The Verge*
6. *Iyyanar, P & Chitra, M 2014, 'An Approach for Design and Implementation of Secure Video Streaming using SRTP'. Journal of Theoretical and Applied Information Technology, ISSN: 1992-8645, vol. 62, no.2, pp. 438-444.*
7. *M. Baugher, D. McGrew, M. Naslund, E. Carrara, and K.Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, 2004*
8. *Logitechh 2012, The H.264 Advanced Video Coding (AVC) Standard. Available from: http://www.logitech.com/assets/45120/logitechh.pdf. [23 August 2013].*
9. *Iyyanar, P & J Senthil Kumar, 2016 'An Effective and Secure Scheme for Audio/Vidio Streaming Based on Modified SRTP using AES-CFB Encryption International Journal of Advanced Engineering Technology' vol.7, no.2, pp.449-453.*
10. *ITU-T and ISO/IEC, "ITU-T recommendation H.264 (2010): advanced video coding for generic audiovisual services," Information Technology Coding of Audiovisual Objects Part 10: Advanced Video Coding ISO/IEC 14496-10:2010, 2010.*
11. *A. Vetro, T. Wiegand, and G. J. Sullivan, "Overview of the stereo and multi-view video coding extensions of the H.264/MPEG-4 AVC standard," Proceedings of the IEEE, vol. 99, no. 4, pp. 626–642, 2011.*
12. A. Lykourgiotis, K. Birkos, T. Dagiuklas et *al., "Hybrid broadcast and broadband networks convergence for immersive TV applications," IEEE Wireless Communications, vol. 21, no. 3, pp. 62–69, 2014.*
13. *G.-M. Su, Y.-C. Lai, A. Kwasinski, and H. Wang, "3D video communications: challenges and opportunities," International Journal of Communication Systems, vol. 24, no. 10, pp. 1261–1281, 2011.*
14. *C. Tselios, I. Politis, K. Birkos, T. Dagiuklas, and S. Kotsopoulos, "Cloud for multimedia applications and services over heterogeneous networks ensuring QoE," in Proceedings of the IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD '13), pp. 94–98, Berlin, Germany, September 2013.*