Reduction of jitter in 3D video by transmitting over multiple network paths

S. Vishwa Kiran*

Department of CSE,
University Visvesvaraya College of Engineering,
Bangalore University,
Bengaluru, Karnataka, India
Email: vishwakirana@gmail.com
*Corresponding author

S. Raghuram

Pushkala Technologies Pvt Ltd.,
Bengaluru, Karnataka, India
Email: raghushivram@gmail.com

J. Thriveni and K.R. Venugopal

Department of CSE,
University Visvesvaraya College of Engineering,
Bangalore University,
Bengaluru, Karnataka, India
Email: drthrivenij@gmail.com
Email: venugopalkr@gmail.com

Abstract: Stereoscopic video transmission in telemedicine application requires the data to be transferred with a minimal jitter. It is not possible to send stereoscopic video at full HD rate on single internet service providers (ISPs) as the bandwidth becomes a bottle-neck and congestion can lead to packet drops, eventually leading to jitter in a video. This could be a circumvented by employing multiple ISPs to stream stereoscopic video utilising multiple real-time packets (RTPs) sessions. Usage of multiple ISPs results in multiple network paths between the video streaming device and video consumers. This concept effectively involves aggregation of bandwidth, delay, jitter, packet loss, and other qualitative network attributes with respect to every ISP participating in the video transmission process. This article analyses through simulation collective delay and jitter which affects the video reconstruction process and concludes with the estimation of minimum qualitative network parameters required.

Keywords: 3D video; bandwidth; cloud aggregation server; CAS; discrete event simulator; ISP; jitter; multipath; multiple ISPs; simpy simulator; stereoscopic video.

Reference to this paper should be made as follows: Vishwa Kiran, S., Raghuram, S., Thriveni, J. and Venugopal, K.R. (2023) 'Reduction of jitter in 3D video by transmitting over multiple network paths', *Int. J. Advanced Intelligence Paradigms*, Vol. 24, Nos. 1/2, pp.110–125.

Biographical notes: S. Vishwa Kiran has 15 years of experience in software development and training. He has handled a corporate training assignments for Cisco, Texas instruments, L&T, Nokia, Samsung, Sasken, Siemens, Honeywell, NSN, HCL, Wipro and Infinite Solutions through Aprameyah Technologies Pvt Ltd. He received his Bachelor of Engineering in Electronics and Communication from the Bangalore University and Master of Technology in Computer Science and Engineering from the Visvesvaraya Technological University and presently associated as a PhD Research Scholar at University Visvesvaraya College of Engineering, Bangalore University and his research interest is in mobile cloud computing and internet of things.

S. Raghuram has completed his BE in Electronics and Communication and MTech in Computer Networks. He has more than 15 years of experience in embedded domain as hardware designer and firmware developer. He is a Director at Pushkala Technologies Pvt Ltd. and pursuing his PhD at Visvesvaraya Technological University.

J. Thriveni has completed her Bachelor of Engineering, Masters of Engineering and Doctoral in Computer Science and Engineering. She has four years of industrial experience and 20 years of teaching experience. Currently, she is an Associate Professor in the Department of CSE, University Visvesvaraya College of Engineering, Bangalore. She has over 50 research papers to her credit. She has produced three doctorate student and guiding seven PhD students. Her research interests include networks, data mining and biometrics.

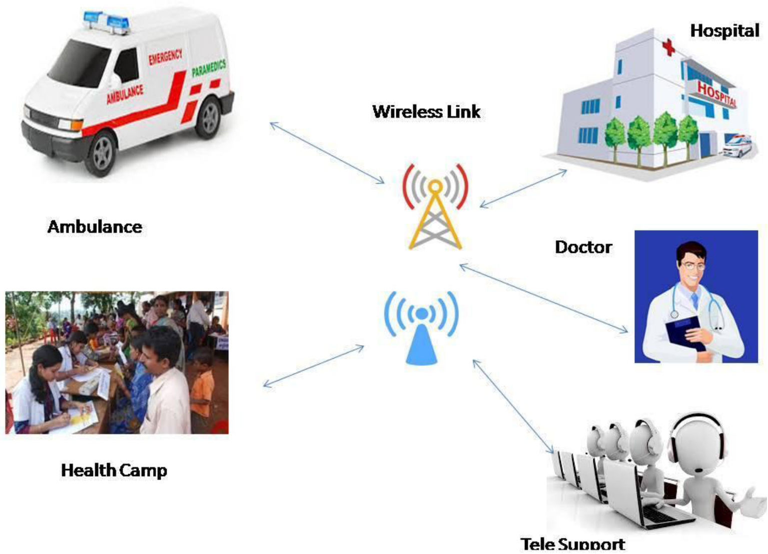
K.R. Venugopal has been in Bangalore University for the last four decades. He has 11 degrees with PhD in Computer Science Engineering from the IIT-Madras, Chennai and another PhD in Economics from the Bangalore University. He has degrees in Law, Mass Communication, Electronics, Economics, Business Finance, Computer Science, Public Relations and Industrial Relations. He has authored and edited 64 books and published more than 600 papers in refereed international journals and international conferences. He has supervised 630 ME dissertations, 22 PhDs and filed 101 patents. He was a post doctoral research scholar and a Visiting Professor at University of Southern California, USA. He has been conferred Fellow of IEEE, USA and ACM Distinguished Educator for contributions to computer science engineering and electrical engineering education.

This paper is a revised and expanded version of a paper entitled 'Efficient stereoscopic 3d video transmission over multiple network paths' presented at CIC2016, Puducherry, India, 19–20 October 2016.

1 Introduction

There is a growing demand for telemedicine where online doctors carry out remote diagnosis and treatment. Doctors can examine patients remotely through video conference. In the present day scenario, there is a disparity in upload and download speeds of Internet connectivity due to asymmetric digital subscriber line (ADSL) and similar technology implementation. Vishwa Kiran et al. (2014) have proposed mobile cloud application for medical applications, which uses 3D tablets that are well suited for remote diagnosis. 3D Tablets help doctors to examine a patient remotely and can be used for guiding doctors during surgery. Figure 1 demonstrates applications of 3D tablets for medical instances.

Figure 1 Use of 3D tablets in telemedicine (see online version for colours)



2 Problem statement

2.1 Background

As discussed in Section 1, HD video is required for live transmission and storage of telemedicine applications. The required stereoscopic video bit rates (video encoding settings for H.264 excellence) are listed in Table 1. To transmit full HD stereoscopic video at 1,080 p resolution, data needs to be transmitted at 9,984 kbps.

Table 1 Video bitrates derived from video encoding settings for H.264 excellence

Name	*Resolution*	*Mono video (kbps)*	*Stereo video (kbps)*
240p	424 × 240	576	1,152
360p	640 × 360	896	1,792
432p	768 × 432	1,088	2,176
480p	848 × 480	1,216	2,432
480p HQ	848 × 480	1,536	3,072
576p	1,024 × 576	1,856	3,712
576p HQ	1,024 × 576	2,176	4,352
720p	1,280 × 720	2,496	4,992
720p HQ	1,280 × 720	3,072	6,144
1,080p	1,920 × 1,080	4,992	9,984
1,080p HQ	1,920 × 1,080	7,552	15,104
1,080p Superbit	1,920 × 1,080	20,000	40,000

Table 2 Measured speeds of various ISPs at random day and time intervals

Service provider	*Sample set 1 – Mbps*		*Sample set 2 – Mbps*		*Sample set 3 – Mbps*	
	Downlink	*Uplink*	*Downlink*	*Uplink*	*Downlink*	*Uplink*
ISP – 1 4G	16.16	2.60	14.16	1.54	19.34	2.68
ISP – 1 3G	4.38	0.86	3.10	0.39	4.43	1.36
ISP – 1 ADSL	1.88	0.25	2.20	0.42	2.15	0.39
ISP – 2 fibre net	19.47	12.53	20.89	19.79	12.76	6.62
ISP – 3 VDSL	12.28	2.84	14.46	1.67	6.34	0.79

Uplink speeds of various internet service providers (ISPs) are listed as shown in Table 2. These samples are taken at random day and time intervals. We observed that none of them meet the required speed of 9,984 kbps. To overcome these limitations Vishwa Kiran et al. (2016) have proposed a novel architecture which uses multiple Internet connections to transmit packet/data to a cloud aggregation server (CAS). The CAS, instead of streaming to a single ISP, makes use of more than one ISP to achieve jitter-free data transmission.

Uplink speeds of various ISPs is listed as shown in Table 2. These samples are taken at random day and time intervals. We observed that none of them meet the required speed of 9,984 kbps. To overcome these limitations, Vishwa Kiran et al. (2016) have proposed a novel architecture which uses multiple Internet connections to transmit packet/data to a CAS. The CAS, instead of streaming to a single ISP, makes use of more than one ISP to achieve jitter-free data transmission.

2.2 Contribution

Transmitting sequentially correlated data (e.g., video stream) over multiple networks often results in jitter and packet loss during the reconstruction of data. This problem is evident due to inherent network delay as modelled in equation (1) and the variations in bandwidth. Consequently, data packets may arrive out of order. The current work analyses the contribution of delay towards jitter and proposes a solution wherein, the reconstruction or consumption of video data stream starts after a delay defined by equation (3); ToS is an abbreviation for Time to Start. Mean delay calculation is modelled as in equation (2) by considering the effect of jitter to be minimised or overcome altogether. The proposed hypothesis is tested through a simulator called simpy.

$$\text{NWdelay}_{\text{max}} = \max\{\text{delayNW1, delayNW2, delayNWn}\} \tag{1}$$

$$\text{NWdelay}_{\text{mean}} = 1/n\sum_{i=1}^{n}\text{delayNWi} \tag{2}$$

$$\text{ToS} = \text{NWdelay}_{\text{max}} + \text{NWdelay}_{\text{mean}} \tag{3}$$

2.3 Organisation of the paper

The rest of the paper is organised as follows:

Section 3 focuses on literature survey which provides insight into related work, Section 4 presents proposed architecture and algorithm, Simulation results are discussed in Section 5 and finally, Section 6 concludes the paper.

3 Literature survey

Restrictions of 4G (Martin et al., 2011) are clear as far as spectrum distribution and are subject to various nations. Due to administrative reasons the accessible range is not used to the fullest extent. Considering all chances still, 4G systems happen to be encouraging and productive for medical applications (Hewage et al., 2011).

Stereoscopic video encoding (Karim et al., 2008; Tseng et al., 1994; Hewage et al., 2007; Merkle et al., 2009a, 2009b; Balasko, 2009; Vetro et al., 2011; Jin et al., 2012; Zund et al., 2013; Schwarz et al., 2007; Stockhammer et al., 2003) is one of the principal methods in 3D video encoding techniques, there are numerous research endeavours toward proficiently encoding stereoscopic 3D video content and transmitting over Internet and mobile networks. H.264/MVC is one of the most extensively acknowledged encoding standard formats for 3D video transmission applications, and extremely reasonable for low latency and high speed networks (Jassal, 2016; Wenger, 2003; Micallef and Debono, 2010; Kordelas et al., 2012; Micallef et al., 2010; Liu et al., 2011; Zhao et al., 1997; Seo et al., 2010).

There have been endeavours to effectively transfer/stream 3D video for health check-up related applications through 4G systems (Hewage et al., 2007). Teleoperation by specialists is a reality now and the coming of 3D transmission for examining a patient's condition has given a gigantic lift to this section.

Hewage et al. (2007) have identified certain difficulties in 3D video capturing in remote gadgets particularly engaged towards medical applications. One of the difficulties is system data transfer capacity limit when transmitting high-quality 3D video. A similar work claims that 4G data transfer capacity accessibility can offer just in part towards 3D video gushing prerequisites. For this they recommend utilising asymmetric 3D stereoscopic encoding strategies and other related strategies. It is evident that a stereoscopic 3D video consumes twice the bandwidth when compared to a standard video transmission. H.264/MVC (Liu et al., 2011) happens to be the most preferred multi-view coding technique for commercial 3D applications and devices.

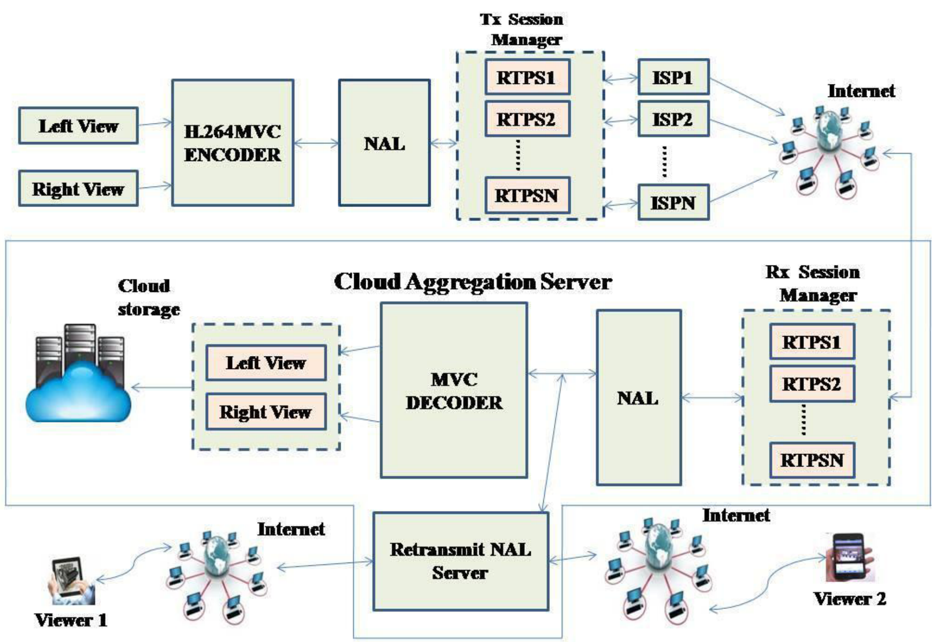
An alternate approach to capturing, streaming and viewing of stereoscopic video incorporates older H.264 AVC, where data is encoded individually, independent of right and left channels. This is referred to as Simulcast (Merkle et al., 2009a). We believe that futuristic 3D video HEVC techniques will be based on H.264/MVC (Kovacs, 2014) which encodes both right and left views concurrently, resulting in two interdependent bit streams (BS), followed by a multiplexer, which interleaves frames of each channel culminating into a single transport stream (TS) (Merkle et al., 2009a). Further, the TS are packetised by network abstraction layer (NAL) in various formats to suit the network need (Hewage et al., 2011; Kordelas et al., 2012; Wang et al., 2012). Real-time transport protocol (RTP) (Hannuksela et al., 2012; Wenger and Wang, 2011) running over user datagram protocol (UDP) is most widely used approach for streaming audio and video data. These proposed techniques are good, but are still limited by additional video processing algorithm and again limited by capacity and implementation of 4G networks.

To overcome these limitations and to cut computational load on battery operated mobile handheld devices, we propose a novel idea of splitting the streaming process of right and left channels into at least two available Internet connections. For example, say one is 4G and the other is WiFi network. WiFi network connected to the internet through ADSL links does not have high upload bandwidth capacity by itself, hence a combination of both WiFi/ADSL 4G is deemed to give high bandwidth capacity. 3D video processing is a challenging task (Merkle et al., 2009b), there are variations in video processing and encoding techniques and each algorithm or standard have its own advantages and limitations. In general, H.264 is one of the major encoding and decoding industry video standard.

4 Architecture

Proposed architecture depicted in Figure 2 uses the likelihood by multi-session transmission (MST) (Martin et al., 2011; Wang et al., 2011) characterised for H.264/MVC RTP sessions. Transmit session manager (TSM) and receive session manager (RSM) interfaces with NAL of both H.264/MVC encoder and decoder to distribute and combine RTP sessions on to several networks.

Figure 2 Proposed architecture of multiple network transmission technique process (see online version for colours)

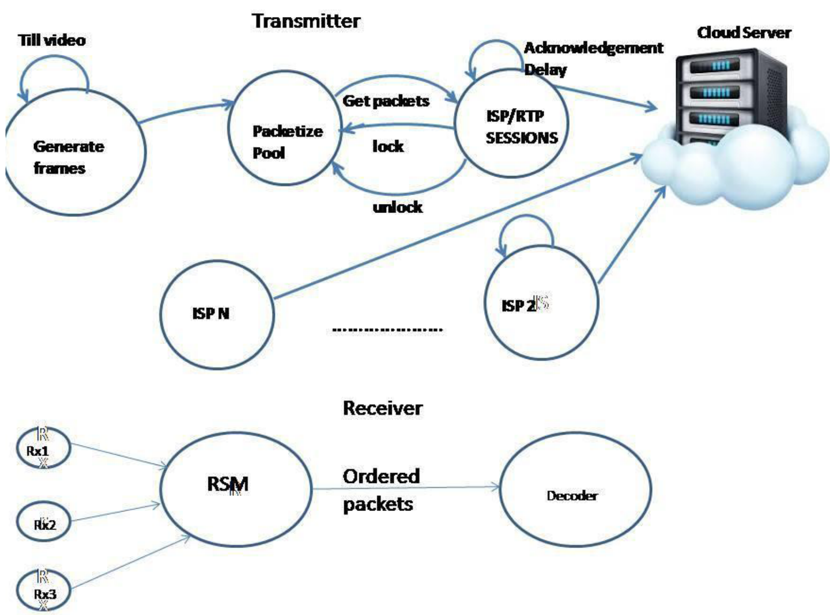


TSM at the encoder side is the key element which identifies several of RTP sessions required. This is based on a number of networks that are available and the transfer limit in each network. TSM conveys to RSM to open the same number of RTP accepting sockets. Interleaved cross-session decoding order number (CS-DON) packet based model

(Wang et al., 2012) of MST in NAL is arranged in the proposed design. This model empowers interleaving and thereby beats the impediment of few latency networks. Another preferred standpoint of utilising this mode is to use the component of CS-DON, this strategy encourages simple and productive decoding of all RTP session packets.

This approach gives flexibility of storing captured video for further analysis and more importantly a single cloud-based viewing endpoint. In our proposed architecture, a stereoscopic 3D video that is captured is transmitted to CAS, instead of transmitting to one or many viewers directly. This approach gives flexibility to analyse the captured video. The uplink ability of every network connection is estimated using round trip time.

Figure 3 Process sequence for algorithm 1 for transmission and reconstruction (see online version for colours)



Process sequences depicted in Figure 3 corresponds to Algorithm 1, specifically referring to transmission at video recording source and represents the sequence of video reconstruction operation at CAS. Our simulation model is implemented utilising the same algorithm.

Algorithm 1 Pseudo code for transmitting 3D video over multiple networks

Input: nwBandwidth (each ISP's bandwidth), nwDelay (each ISP's average delay)
Output: Jitter statistics

for each ISP:
    create a transmission thread
    start thread
for each simulation **clock** tick:
    Thread – 1: //Video Generation
        generate video content
        write to buffer

```
Thread – 2: //Multi threaded transmission
    for each ISP
        lock buffer access read buffer
        unlock buffer access
        if data to transmit
            transmit
        while not acknowledgment
            wait
        if negative acknowledgment
            retransmit
        repeat till end of data
Thread – 3: //Receive and Video reconstructing
    while not initial delay expired: wait
    read received buffer
    if buffer has data to reconstruct video reconstruct video
    else: Buffer has insufficient data due to delay wait for data
Thread – 4: //Simulation control
    if transmission over and reception over:
            stop simulation
            save
            results exit
    else:
            continue
```

5 Simulation and results

Multiple trials are run under various bandwidth and delay conditions to test the proposed hypothesis. Video stream data of 3 seconds and 3 different ISPs or network paths are considered uniformly for all the trials. Figure 3 shows the sequence diagram for the proposed network architecture and its algorithm which is being simulated. Python and SimPy (Simpy Discrete Event Simulator, http://simpy.readthedocs.io/en/latest/) which is a discrete event simulator is used to implement Algorithm 1. Each simulation clock tick is set to be at 10ms real time.

Theoretically, the video data generation process and reconstruction process will take 300 steps uniformly. Referring to Table 1 and considering 1,080p stereoscopic video bit rate requirement of 9.984 Mbps is considered throughout the simulation trials. If the available network bandwidth caters to the required 9.984 Mbps, simulation cycle will end shortly after 3 seconds. Otherwise, it will consume much more time to end. It is quite obvious that if a network path between source and destination is available to cater this bandwidth and no latency, then video is decoded or streamed without any jitter. Unfortunately, this is not the case always; Table 2 presents measured upload speed samples of various ISPs measured.

Table 3 Consolidated simulation results

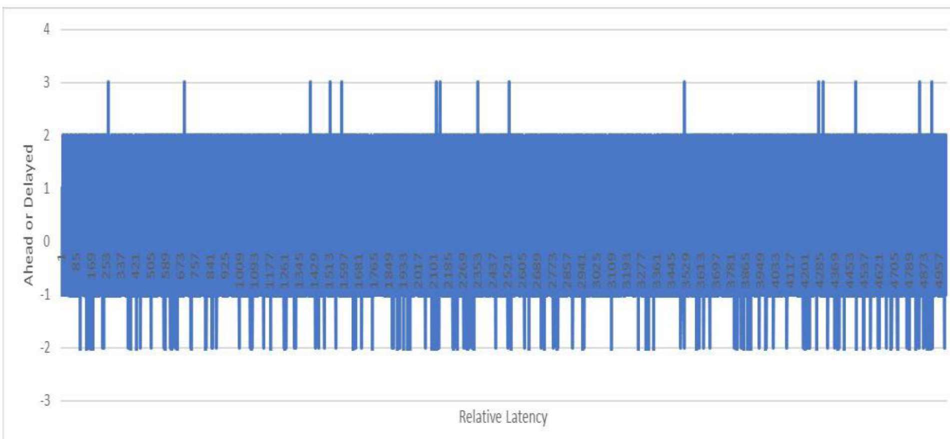
Simulation	Delay in seconds					Bandwidth in Mbps				Receive Buffer Access				Time to start in ms	Simulation duration in sec
	NW1	NW2	NW3	AVG	AVG + max	NW1	NW2	NW3	AVG	Miss	Hit	Total	Start		
1	0.07	0.07	0.13	0.09	0.22	1	0.25	0.5	0.583	9,441	300	9,741	3.08	0	97
2	0.01	0.01	0.01	0.01	0.02	1	0.25	0.5	0.583	1,397	300	1,697	17.68	0	17
3	0.01	0.01	0.01	0.01	0.02	1	1	1	1	637	300	937	32.02	0	9
4	0.01	0.01	0.01	0.01	0.02	1	1	1	1	635	300	635	47.24	700	9
5	0.01	0.01	0.01	0.01	0.02	3	3	3	3	35	300	335	89.55	0	3
6	0.01	0.01	0.01	0.01	0.02	3	3	3	3	4	300	304	98.68	0	3
7	0.01	0.01	0.01	0.01	0.02	3	3	3	3	0	300	300	100.00	0	3
8	0.07	0.07	0.07	0.07	0.14	3	3	3	3	1,649	299	1,350	22.15	0	16
9	0.03	0.03	0.03	0.03	0.06	3	3	3	3	530	300	830	36.14	60	10
10	0.03	0.03	0.03	0.03	0.06	3	3	1	2.333	815	300	1,115	26.91	60	13
11	0.03	0.03	0.03	0.03	0.06	3.328	3.328	3.328	3.328	430	299	729	41.02	60	9
12	0.04	0.05	0.06	0.05	0.11	3.328	3.328	3.328	3.328	1,147	299	1,446	20.68	110	14
13	0.05	0.06	0.07	0.06	0.13	3.328	3.328	3.328	3.328	1,458	300	1,758	17.06	130	17
14	0.06	0.07	0.08	0.07	0.15	2.328	2.328	2.328	2.328	2,650	300	2,950	10.17	150	29
15	0.06	0.07	0.08	0.07	0.15	3.328	3.328	3.328	3.328	2,051	299	1,752	17.07	150	20
16	0.06	0.07	0.08	0.07	0.15	4.328	4.328	4.328	4.328	1,282	300	1,582	18.96	150	16
17	0.06	0.07	0.08	0.07	0.15	5.328	5.328	5.328	5.328	981	300	1,281	23.42	150	13
18	0.06	0.07	0.08	0.07	0.15	6.328	6.328	6.328	6.328	778	300	1,078	27.83	150	11
19	0.06	0.07	0.08	0.07	0.15	7.328	7.328	7.328	7.328	630	300	930	32.26	150	9
20	0.06	0.07	0.08	0.07	0.15	8.328	8.328	8.328	8.328	518	300	818	36.67	150	8
21	0.06	0.07	0.08	0.07	0.15	9.328	9.328	9.328	9.328	429	300	729	41.15	150	7
22	0.06	0.07	0.08	0.07	0.15	10.328	10.328	10.328	10.328	359	300	659	45.52	150	6

Referring to Table 3 '% success' parameter is the percentage of jitter-free reconstruction of video when compared to given three seconds of total actual video content. This parameter is the measure of successful jitter-free video reproduction at CAS. Similarly 'time to start' is the time delay at CAS to start reproducing or decoding or reconstructing the video. This delay will provide the receiver data buffer at CAS to accumulate some amount of data to help jitter-free operation or reduction in jitter.

Due to fluctuations in delay, bandwidth and packet loss effects, ordered multisession video packets traversed over multiple network paths tend to arrive at varying time intervals in a disordered fashion. This effect is simulated and the resultant packet arrival relative latency information is presented below. Four simulation trials 1, 6, 7 and 22 are considered; these respectively are related to worst, best, perfect and mediocre video reconstruction success rates.

Since the data presented is relative to another packet, two samples per packet are recorded. Negative value with respect to an ordered packet number is considered as delayed arrival. If there is no latency observed, then the difference value would be zero.

Figure 4 Worst case scenarios (see online version for colours)



It could be seen that the density of relative latency is very high in worst case scenario depicted in Figure 4. Unlike in this scenario, in best and mediocre cases, the latency of arrival is sparse. It should be noted that average bandwidth and delay are different in trial 6 and 22 corresponding to best and mediocre scenarios. It can be observed that in the perfect scenario depicted in Figure 6, initially, the packets arrive in a disordered fashion and over a period, network flow stabilises and 100% jitter-free reproduction is achieved with initial latency 600 ms. Also note the overall bandwidth of the network is around 9 Mbps and average network delay is 10 ms. Figure 4 represents the packet arrival latency for trial 1 for worst case scenario and we can observe that there is more latency. Packet arrival relative latency for trial 6 is recorded as shown in Figure 5 for the best case scenario and we can observe less jitter in the graph. Figure 6 represents the packet arrival relative latency for trial 7 which depicts perfect scenario where there is no jitter. Figure 7 represents the packet arrival relative latency for trial 22, which depicts the mediocre scenario.

Figure 5 Best case scenarios (see online version for colours)

Figure 6 Perfect scenarios (see online version for colours)

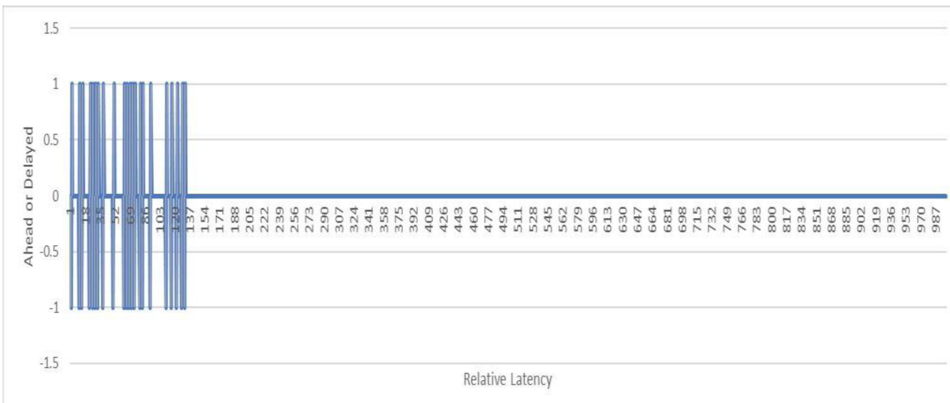


Figure 7 Mediocre scenarios (see online version for colours)

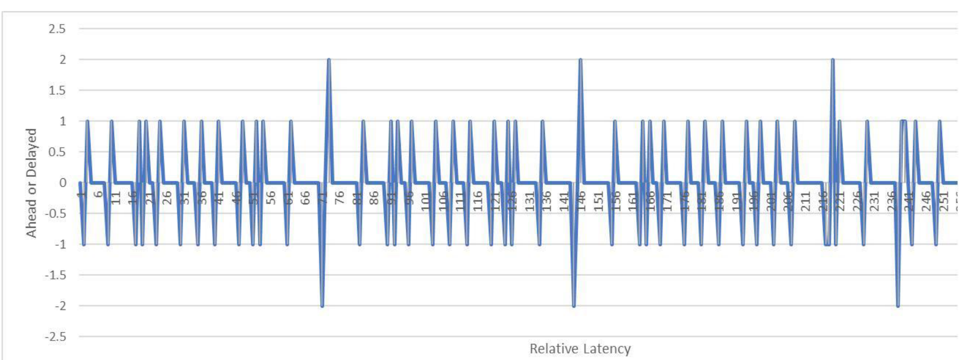


Figure 8 corresponds to simulation trials where network delay average and average bandwidth parameters are kept constant, these values are tabulated in Table 5. Eight simulation trials are conducted with ToS delay being varied as in Table 4. It may be noted that simulation time remains constant, and only the buffer read miss rate varies. It may be observed from Table 4, that success rate increases as the Time to Start delay increases, but after 900 ms success rate falls back to first value and pattern repeats.

Maximum success rate observed in these simulation trials is at 35%. Time to start is limited to 1,800 ms to restrict to 3 s of simulation video data and considering practical live video streaming applications, wherein higher time gap greater than 1 s between video capturing and reproduction is not appreciated.

Figure 8 Percentage of successful jitter free reconstruction based on time to start (see online version for colours)

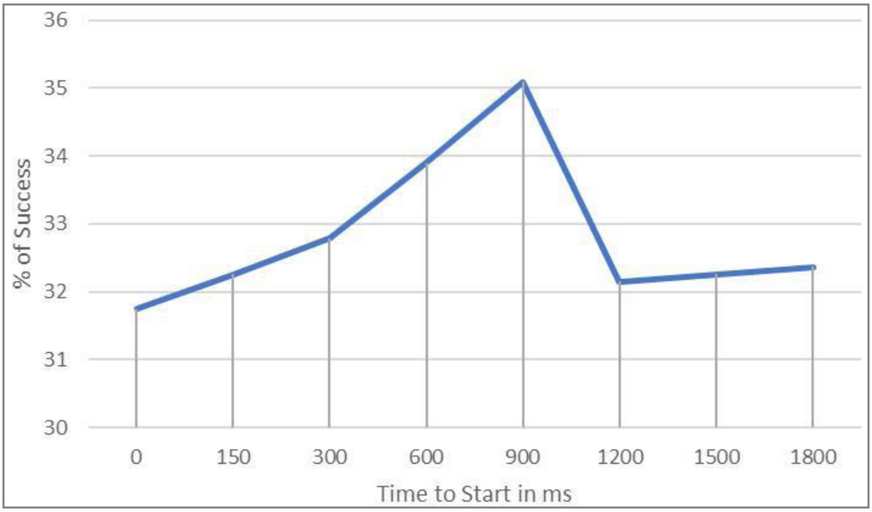


Table 4 Success percentage v/s time to start

Simulation trial	*Receive buffer access*				*Time to start in ms*	*Simulation duration sec*
	Miss	*Hit*	*Total*	*% success*		
1	645	300	945	31.75	0	9
2	630	300	930	32.26	150	9
3	615	300	915	32.77	300	9
4	585	300	885	33.89	600	9
5	555	300	855	35.08	900	9
6	633	300	933	32.15	1,200	9
7	630	300	930	32.25	1,500	9
8	627	300	927	32.36	1,800	9

Table 5 Constants for success percentage v/s time to start trials

Delay in seconds					*Bandwidth in Mbps*			
NW1	*NW2*	*NW3*	*AVG*	*AVG + Max*	*NW1*	*NW2*	*NW3*	*AVG*
0.06	0.07	0.08	0.07	0.15	10.328	7.328	3.328	6.995

Figure 9 depicts the variations in resulting successful jitter-free reproduction of video at CAS with respect to variable average bandwidth. Time to start and other parameters are kept constant. Simulation results have yielded % success growth rate of about 4.4% for every increase in 1 Mbps of average bandwidth. Data corresponding to these results is tabulated in Table 3.

Figure 9 Variation in average bandwidth v/s % of success based (see online version for colours)

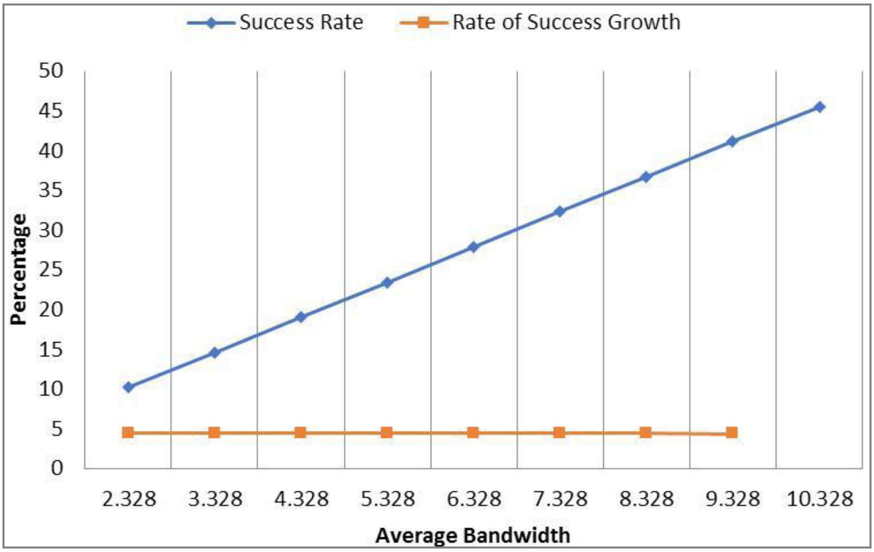
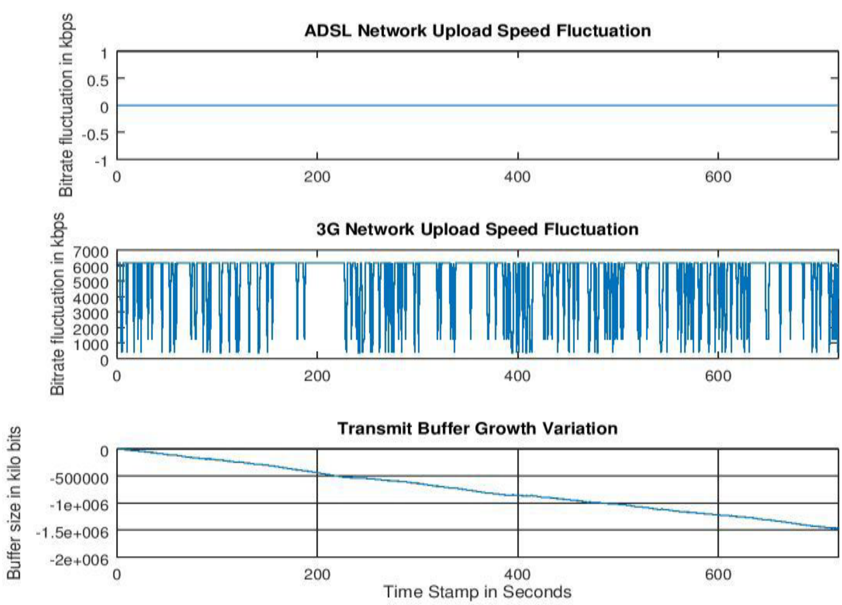


Figure 10 Single network with 6.144 Mbps bandwidth (see online version for colours)



Based on the work carried out by Vishwa Kiran et al. (2016). Simulation comparison results for a single network with a bandwidth of 6.144Mbps and two networks with a combined bandwidth of 6.144 Mbps are tabulated in Table 6. Figure 10 depicts Case 1 and Figure 11 depicts Case 2. It can be observed from Table 6, time to empty is the additional time required in seconds to send the buffered 3D video's RTP packets. Since time to empty is negative indicates there is no latency in the transmission of packet. Case 1 would be preferred over Case 2 as Case 2 would need more time for construction and reconstruction of packets at CAS end.

Figure 11 Two networks with 3.072 Mbps bandwidth each (see online version for colours)

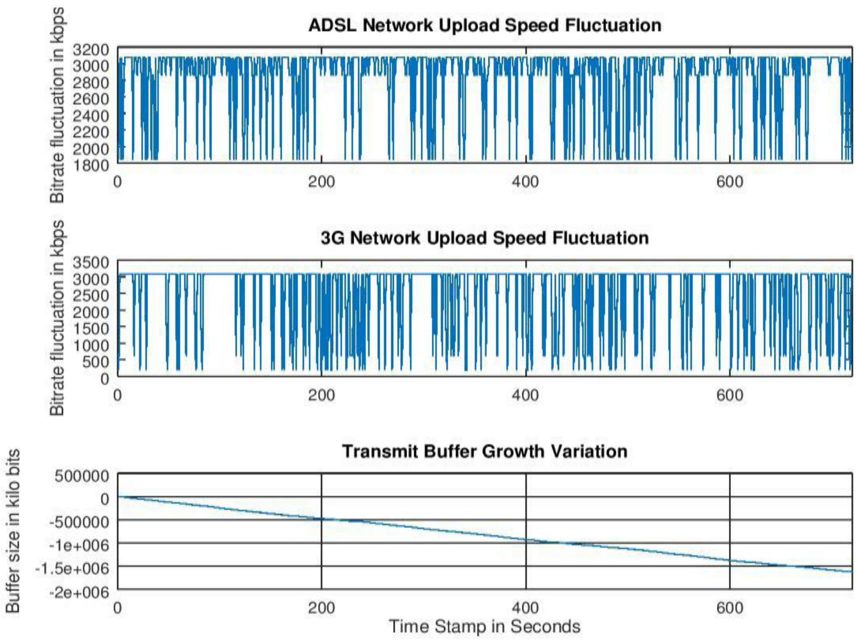


Table 6 Simulation test results of RTP packet buffering and time to empty

Case	*Peak ADSL upload rate in kbps*	*Peak 3G/4G Upload Rate in kbps*	*Mean ADSL upload rate in kbps*	*Mean 3G/4G Upload Rate in kbps*	*Consolidated mean upload rate in kpbs*	*Balance in buffer in kbps*	*Time to empty or latency in seconds*
1	0	6,144	0	5,104	5,104	–1.4635e+006	–287
2	3,072	3,072	2,835	2,492	5,327	–1.6242e+006	–305

6 Conclusions

Extensive simulation trials have been conducted to measure the effect of jitter due to multiple network paths and contribution of time to start factor in minimising the jitter effect or in other words improving the video reproduction quality. It is observed that hypothesis with respect to time to start parameter will improve the video playback performance. Variations in network bandwidth and delay are also significant contributors to jitter. Increase in average bandwidth availability and significant network delay has conclusively yielded in improving jitter-free reproduction.

References

Balasko, H. (2009) *Comparison of Compression Algorithms for High Definition and Super High Definition Video Signals* [online] http://vista.fer.hr/_download/repository/KDI_Balasko_(1).pdf (Accessed 10 January 2016).

Hannuksela, M., Chen, Y., Sullivan, G., Suzuki, T. and Hattori, S. (2012) *Study Text of ISO/IEC 14496-10:2012/DAM2 MVC Extension for Inclusion of Depth Maps* [online] http://phenix.it-sudparis.eu/jct2/doc_end_user/documents/2_Shanghai/wg11/JCT3V-B1001-v1.zip (accessed 22 December 2016).

Hewage, C.T.E.R. et al. (2007) 'Comparison of stereo video coding support in MPEG-4 MAC, H.264/AVC and H.264/SVC', *Proceedings of IET Visual Information Engineering-VIE07*.

Hewage, C.T.E.R., Martini, M.G. and Khan, N. (2011) '3D medical video transmission over 4G network', *Proceedings of the 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies*, ACM.

Jassal, A. (2016) *H.265/HEVC Video Transmission over 4G Cellular Networks* Doctorial dissertaion, The University of British, Columbia [online] https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.022389 (Accessed 20 December 2016).

Jin, L., Boev, A., Gotchev, A. and Egiazarian, K. (2012) '3D-DCT based multi-scale full-reference quality metric for stereoscopic vide', *Proceedings of 6th International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, Arizona State University, Arizona

Karim, H.A., Hewage, C.T.E.R. and Worrall, A.M.K. (2008) 'Scalable multiple description video coding for stereoscopic 3D', *IEEE Transactions on Consumer Electronics*, Vol. 54, No. 2, pp.745–752.

Kordelas, A., Dagiuklas, T. and Politis, I. (2012) 'On the performance of H.264/MVC over lossy IP-based networks', *Signal Processing Conference (EUSIPCO)*.

Kovacs, P.T. (2014) 'Overview of the applicability of H.264/MVC for real-time light-field applications', *Proceedings of 3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video*.

Liu, Y. et al. (2011) 'A novel rate control technique for multiview video plus depth based 3D video coding', *IEEE Transactions on Broadcasting*, Vol. 57, No. 2, pp.562–571.

Martin, J., Amin, R., Eltawil, A. and Hussien, A. (2011) 'Limitations of 4G wireless systems', in *Proceedings of Virginia Tech Wireless Symposium*, Blacksburg, VA.

Merkle, H.B., Dix, K., Muller, K. and Wiegand, T. (2009a) 'Stereo video compression for mobile 3D service', *Proceedings of IEEE 3D TV Conference: The True Vision- Capture, Transmission and Display of 3D Video*.

Merkle, W.Y., Muller, S. and Wiegand, T. (2009b) 'Video plus depth compression for mobile 3D services', *Proceedings of IEEE 3DTV Conference*, Germany.

Micallef, B.W. and Debono, C.J. (2010) 'An analysis on the effect of transmission errors in real-time H.264-MVC bit-streams', *15th IEEE Mediterranean Electrotechnical Conference (IEEE MELECON)*.

Micallef, B.W., Debono, C.J. and Farrugia, R.A. (2010) 'Error concealment techniques for H.264/MVC encoded sequences', *IEEE Proc. of Int. Conf. of Electrotechnical and Computer Science (ERK)*, Portoroz, Slovenia.

Schwarz, H., Marpe, D. and Wiegand, T. (2007) Overview of the scalable video coding extension of the H.264/AVC standard', *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 17, No. 9, pp.1103–1120.

Seo, K-D. et al. (2010) 'A practical RTP packetization scheme for SVC video transport over IP networks', *Etri Journal*, Vol. 32, No. 2, pp.281–291.

Simpy Discrete Event Simulation for Python [online] http://simpy.readthedocs.io/en/latest/ (Accessed 5 January 2017).

Stockhammer, T., Hannuksela, M.M. and Wiegand, T. (2003) 'H.264/AVC in wireless environments', *IEEE Transactions On Circuits and Systems for Video Technology*, Vol. 13, No. 7, pp.657–673.

Tseng, B.L. and Anastassiou, D. (1994) 'Compatible video coding of stereoscopic sequences using MPEG-2's scalability and interlaced structure', *Proceedings of International Workshop on HDTV*, Vol. 94.

Vetro, A., Tourapis, A.M., Muller, K. and Chen, T. (2011) '3D-TV content storage and transmissio', *IEEE Transactions on Broadcasting*, Vol. 57, No. 2, pp.384–394.

Video Encoding Settings for H.264 Excellence [online] http://www.lighterra.com/papers/videoencodingh264 (accessed 20 December 2016).

Vishwa Kiran, S., Raghuram, S., Thriveni, J/ and Venugopal, K.P. (2016) 'Efficient stereoscopic 3D video transmission over multiple network paths', *International Conference on Advances in Computational Intelligence and Communication*, Puducherry, India, October.

Vishwa Kiran, S., Prasad, R., Thriveni, J., Venugopal, K.R. and Patnaik, L.M. (2014) 'Mobile cloud computing for medical applications', in *Proceedings of the 11th IEEE India Conference on Emerging Trends and Innovation, INDICON*, Pune, India.

Wang, Y.K., Even, R., Tandberg, T.K. and Jesup, R. (2011) *RFC-6148: RTP Payload Format for H.264 Video*, May [online] https://tools.ietf.org/html/rfc6184 (accessed 20 December 2016).

Wang, Y.K., Schier1, T., Skupin, R. and Yue, P. (2012) *RTP Payload Format for MVC Video* [online] https://tools.ietf.org/html/draft-ietf-payload-rtp-mvc-02 (accessed 20 December 2016).

Wenger, S. (2003) 'H.264/avc over IP', *IEEE Transactions on Circuits and Systems for Video Technolo*gy, Vol. 13, No. 7, pp.645–656.

Wenger, Y.K. and Wang, S. (2011) *RTP Payload Format for Scalable Video Coding*, Internet Engineering Task Force.

Zhao, M.L. et al. (1997) *Experimental Evaluation of H.264/Multiview Video Coding over IP Networks*, arXiv preprint alg-geom/9705028.

Zund, F., Pritch, Y., Sorkine-Hornung, A., Mangold, S. and Gross, T. (2013) 'Content-aware compression using saliency-driven image retargeting', *Proceedings of IEEE International Conference on Image Processing*, pp.1845–1849.