International Journal on Recent and Innovation Trends in Computing and Communication
Volume: 3 Issue: 8
ISSN: 2321-8169
5246 - 5251

# Analysis of Real Time Video Communication Systems

Mr. Hitendra S. Patil
P.G.Student, Computer Department
P.I.C.T, Dhankawadi, Savitribai Phule University,
Pune, Maharashtra, India.
*email: hitendrapatil2010@gmail.com*

Mr. Amar R. Buchade
Professor, Computer Department
P.I.C.T, Dhankawadi, Savitribai Phule University,
Pune, Maharashtra, India.
*email: amar.buchade@gmail.com*

***Abstract***—The most of the existing real time video communication systems mainly focus their work on providing better video quality throughout session. In quest of maintaining video quality they succeed in doing so at the cost of broken sessions, blocky video or sound disturbances when network bandwidth drops below required rate. The system described in this paper mainly concentrates on analysis of input parameters to audio and video encoder which affects the quality of communication. The input parameters to video encoder are altered such that a balance is maintained between video quality and continuity in communication. The input parameters to video encoderused for analysis are video frame size, and frames per second and target encode bitrate used for encoding video frame. The input parameters for audio encoder used for analysis are sampling frequency, bits per sample and no of audio channels used for recording sound. The input parameters to video encoderare changed frequently depending upon various factors such as bandwidth variations, and encodetime required on hardware used. In extreme low bandwidth situation the video is stopped. The communication should always keep alive throughout the session by keeping audio session connected always, so that users should not feel disconnected. The other important factors required for real time video communication to work smoothly are transport protocols used to carry media data and control data across peers. The protocols discussed in this paper are Real Time Protocol (RTP) and Real Time Control Protocol (RTCP). The media data generated at peers is transported using RTP and the control data describing the media data is transported using RTCP.

***Keywords—****Real Time Video, Audio, RTP, RTCP, Bandwidth Adaptation.*

*****

## I. INTRODUCTION

A Real Time Video Communication system mainly deals with sending and receiving audio and video in real time. The Real Time Video Communication Systems has to do many tasks in parallel to achieve a feel of liveliness. The tasks a Real Time Video Communication System has to doin parallel are capturing audio and video frames, encoding those frames, sending this encoded data to other peer, receiving encoded media data from other peer and decode it, and all these tasks are to be done in parallel. To achieve the feel of liveliness the system has to capture at least ‘n’ number of video frames, encode these frames and send them across the network and that too in restricted time frame. Also the system has to receive media encoded data sent from Remotepeer, decode this encoded data and display it at local peer. Simultaneously a system has to do the same process with audio. The similar process is followed at Remotepeer. All these Real Time activities are measured in specific time frame particularly a time frame of 1 second is used. For Real Time Video a measurement is done for number of image frames captured per second and for audio the number of audio frames captured per second. The other attributes associated with captured video frames is the parameters used to encode the frames so that the generated data is of lesser volume and can be transported over a restricted network bandwidth. The video parameters used for encoding are video frame resolution (size of each video image), the number of video frames the system has supposed to capture per second (frame rate) and a cumulative estimate of number of bytes to be transported per second after encoding captured frames (bit rate) (this is the input to video encoder so that it can use that much compression over the frames to achieve target bandwidth available for transportation) [3] [13].

Many Video Communication systems does not maintain balance in using video encoding parameters, hence they fail to provide uniform quality video over entire session. The excessive use of higher video resolution to provide good quality video does not sustain connectivity all the time in fluctuating network bandwidth conditions. The drawback of using higher resolutions all the time is that it will generate larger volume data and also wastes CPU and Hardware resources time for encoding bigger input data due to the bigger frame size. Using video frame rate more than 15 frames per second is not required for a face to face session. Using frame rate of 30 frames per second will unnecessarily waste CPU and Hardware resources in encoding extra 15 frames and also contribute to extra data bytes for transportation. Using higher bit rate for encoding a frame will expose more details of each frame but will generate equivalent amount of larger volume data for transportation. Usually in face to face session image details up to certain amount is only necessary.

In the proposed system a balance in video encoding parameters is maintained all the time. VP8 Video codec is used for video encoding and decoding. VP8 is an open source video codec comparatively equivalent to H.264 video codec [6] [9]. The video resolution is altered in frequent time gaps depending upon the availability of bandwidth for last time gap and also the cumulative encoding time required for per second. Generally a constant frame rate of 15 frames per second is used to reduce volume of data generated for transportation and also to reduce encoding time required [8]. The bitrate is also periodically altered depending upon data loss in transportation calculated for previous time gap. The parameters are altered frequently to maintain connectivity. A compromise is made in quality whenever network bandwidth drops. Quality will regain when bandwidth catches up.

The audio parameters used in encoding sound are sampling frequency, bit depth per sample and number of audio channels used for recording sound. For end to end communication systems sampling frequency of 8000 samples per second is sufficient to generate good quality mono sound. Bit depth of 16 bits per sample is sufficient to describe reasonable up downs in sound wave. Single audio channel is sufficient for recording verbal communication. Also a good audio codec like iLBC (Internet Low Bitrate Codec) will cut down the generated audio for transportation up to 15.2 kilobits per second. So using iLBC and minimum audio encoding parameter combination will provide good quality audio with saving in target bitrate for transportation [1].

## II. ANALYSIS OF TRANSPORT PROTOCOLS

For Real Time Video Communication to happen smoothly a Transport Protocol is required which can efficiently and speedily transport large amount of media data between end users. Transmission Control Protocol (TCP) cannot be used for Real Time Communication, because it is connection oriented and involve acknowledgement which consumes lot of time. The only benefit of TCP is guaranteed delivery of frame data which is not a requirement of Real Time Communication Systems as missing of certain amount of frame data is tolerable. Hence UDP is used in the system discussed in this paper [4].

UDP is connectionless hence the frame data can be sent across peers irrespective of knowing about its arrival at opposite peers. This benefit in speeding up the process of transportation of large amount of frame data especially generated from Video. But doing so is useless in fluctuating network bandwidth condition. Adequate control is needed over sending and receiving of data. Hence some type of reliability is needed in support of connectionless ness of UDP. Hence UDP alone is not sufficient in achieving reasonable quality of Audio and Video communication.

Hence RTP is used to format the data in fixed size packets to be transported over UDP. By using fixed size packets data can be efficiently distributed and also can be uniquely identified by giving it a unique Time Stamp. This way an order is maintained in sending and receiving of frame data. The generated raw data for each video frame is huge due to storing colour information for each pixel of image. Even after using Encoder for compressing a video frame a sufficiently large amount of data is generated for transportation. There is a limit of packet size sent for a single packet called Maximum Transmission Unit (MTU). Large size packets occupy a slow link for more time than a smaller packet, causing greater delays to subsequent packets, and increasing lag and latency. Generally size of 1500 bytes is used for every packet. The frames having size greater than MTU is fragmented in packets of MTU size minus fixed packet header size. The fragmented packets are given sequence number to keep the order of delivery. All the packets of same frame excluding last packet has set the marker bit to 0 while the last packet is set with marker bit of 1, so that at other end the packets can reassembled to form a complete frame. This way a fixed shape is given to data to be transported over UDP.

Still the connectionless ness of UDP is not fully utilized, the functionality is lacking in knowing the status of data reaching at opposite end, different parameters related to data delivered and the User details of sender. Hence in addition to RTP packets periodic packets are sent in schedule interval which is calculated depending upon the bandwidth availability and number of senders and receivers. Generally interval greater than 2 second is used to send such control packets. These packets follow specific protocol named RTCP. By using RTP in combination with RTCP will make the UDP transmission reliable [5] [7]. This way a speedy and reliable data transmission is achieved through UDP.

## III. SYSTEM REVIEW

In this section the design of proposed system is elaborated. Proposed system uses RTP with UDP for transportation of Media data. The control data is sent using RTCP with UDP. Packet size of 1400 bytes is used for each packet. RTCP packets are sent at regular intervals of more than 2 seconds. RTCP packet interval is calculated depending upon number of senders and receivers.

The packet loss between two RTCP packets is calculated by taking the difference between highest sequence numbers arrived at peer and the base sequence number [2]. A threshold of number of packets lost is considered for altering the video encode parameters to minimize the packet lost for subsequent RTCP intervals. A higher video frame resolution used at sender will bring down the sending frame rate by more than 70% which causes glitter in video at other end. A higher video resolution single frame generally (1280x720) takes more time by encoder to encode hence in restricted time period of 1 second only few frames are encoded. As the system works in Real Time after generation of frame it is transported over network, hence only 30–40% of target frame rate is achieved which result in glitter at other end. Hence to avoid frame rate loss a threshold is maintained for cumulative time required by encoder to encode frames per second, and whenever the encoding time exceeds threshold the capturing video resolution is brings down. The CPU and the memory resources of particular system decides the encoding time required, hence the threshold figure assumed works for particular machine doesn’t work for the other. Hence a percentage of per second time is taken as threshold which represents average CPU and memory capacity.

In fluctuating network bandwidth conditions where a bandwidth is shared between numbers of users, bandwidth changes frequently as the users enter and leave the network. The effect of frequent changes in bandwidth causes the system to alter encoding parameters frequently which affects the video quality. Hence the frequent alteration to video parameters is also not desirable. To avoid frequent alteration to video encoding parameters and also to sustain the system to lower bandwidths a lower parameter combination is used which is one step below the sustainable combination. The problem with such alteration is that when bandwidth availability increases the system does not step up and continue with giving low performance. To avoid such condition after the system stabilizes over some parameter combination, a fixed time interval of x seconds is decided where after each interval the system steps up and tested for its sustainability. If the system is unable to sustain for this combination it steps down again. In stable network bandwidth the system gives best performance, while in changing network bandwidth conditions system gives average performance in terms of quality. The benefit with this approach is that connectivity is maintained all the time.

## IV. BANDWIDTH REQUIREMENT ANALYSIS

In this section bandwidth requirement analysis is done for different video encoding parameters. The analysis is done for four video resolutions 160x120, 320x240, 640x480, and 1280x720 [12]. Frame rates of 5, 15 and 30 frames per second are considered for analysis.

Generally a resolution is needed to be encoded with its desired bitrate and not smaller than that to get a minimum clear visible image. The video target bitrate parameter of encoder decides how much each frame should compress so that target bitrate is achieved approximately. For example if video resolution is 320x240, frame rate is 15 and bitrate is 128kb/s then the encoder is bound to apply that much compression to 15 frames of 320x240 resolutions so that the total of encoded data generated from these 15 frames should be approximately 128kilobits. Hence if a higher resolution frame is tried to apply with lower target bitrate then the quality of video is very unclear. Hence to get a good quality video for different video frame resolutions it is necessary to use a specific minimum target bitrate for encoding depending upon the target frame rate. In other terms it can be said that to produce a reasonably good video for a particular video frame resolution and target frame rate combination a required minimum network bandwidth is needed. Hence the target bitrate for encoder should not be less than the minimum required to produce reasonably good video. The estimation of minimum required bandwidth needed to show a video with particular frame resolution and frame rate combination is described in Table 1.

| **Minimum Required Bandwidth** | **Resolution** | **Frame rate** |
|---|---|---|
| **kbps** | **w x h** | **fps** |
| 128 | 320x240 | 15 |
| 384 | 320x240 | 30 |
| 512 | 640x480 | 15 |
| 768 | 640x480 | 30 |
| 1000 | 1280x720 | 15 |
| 2000 | 1280x720 | 30 |
| 3000 | 1280x720 | 60 |
| > 4000 | 1920x1080 | 30 |

**Table 1. Minimum Bandwidth Requirements for Different Resolution and Frame Rate Combination**

```
Procedure StepDownEncode
    If packetslost >packetslostthreshold Then
        bitrate ←stepdown(bitrate)
        If bitrate <bitratelowerthreshold Then
            frameresolution←stepdown(frameresolution)
        End If
    End If
    If encodetime>encodetimethreshold Then
        frameresolution←stepdown(frameresolution)
    End If
End Procedure
```

**Algorithm 1**

```
Procedure StepUpEncode
    bitrate←stepup(bitrate)
    If bitrate >bitrateupperthreshold Then
        frameresolution←stepup(frameresolution)
    End If
End Procedure
```

**Algorithm 2**

Even if a required minimum bitrate should be used for encoding to get a reasonable quality video, it is not a compulsion to do so. A lower bitrate can be used with higher resolution to catch with the low bandwidth, at the cost of reduced quality video. The system explained in this paper uses this technique to sustain the connectivity in network bandwidth fluctuations. Using lower bitrate for encoding higher resolution image is inefficient and useless too. Hence to cope with lower bandwidth both video frame resolution and target bitrate are brought down to maintain relatively good quality for the lowered resolution.

The algorithm for altering encode parameters for lowering bitrate and resolution is explained in Algorithm 1 and algorithm for stepping up bitrate and resolution periodically is explained in Algorithm 2. Generally nowadays a bandwidth of at least 128 kb/s is available for everybody who is having network connection, and a fixed bandwidth approximately up to subscribe limit is available for personal users. The network fluctuations are observed where a bandwidth is shared among number of users. The proposed algorithm works for both conditions, the difference is fixed bandwidth users experience good quality all the time, while users with bandwidth up downs observed minor quality differences whenever bandwidth changes.

## V. RESULTS

The proposed system is tested on machines with different processing capabilities and different bandwidth fixed and wireless connections. The system is also tested in private standalone as well shared bandwidth conditions. The different processing capabilities include CPU capability and RAM capability. The system is tested over Pentium 4 CPU with 1.5 GB RAM, Intel i5 machine with 6 GB RAM, Intel Dual Core machine with 4 GB RAM and over Intel i7 machine with 8 GB RAM.

Testing the system over low configuration machine involved certain issues regarding encoding the video frames in short period due to its limited capability to perform complex operations involved in encoding process. For bigger resolutions the encoding process takes higher time even on larger capability machines. Hence for successful working of the system over lower capability machines the resolution is usually kept low all the time. In Real Time environment to get the feel of continuity the frames need to be up to certain numbers and that to with shorter gaps of nearly 70 ms. Lower capability machines are not able to produce high resolution frames encoded within shorter period.

On standalone systems bandwidth is not shared hence the bandwidth fluctuations are low. If a standalone system is having larger processing capability it is able to produce higher resolution encoded frames at larger frame rate. Also due to reasonably stable bandwidth they are also able to send those frames over network. But if the machine at other end is having stable bandwidth but lower processing capability it is able to receive those higher resolution encoded frames but is unable to reproduce them by decoding them at higher speeds due to their lower capability. This issue is to be handled at opposite end by lowering resolution and frame rate at its end. The drawback of such approach is average quality at both ends but connectivity is maintained all the time without frequent break ups.

The drawback of altering encode parameters frequently is that some amount of flicker is observed for some time whenever a resolution is changed and the encoder is adjusting itself with the new resolution. Hence the alteration is stopped after the system confirms over some resolution, frame rate and bitrate combination which is suitable for both ends processing capabilities and bandwidth availability. Generally a video communication system starts with lower resolutions and bitrate and gradually steps up with higher resolutions and higher bandwidth. The same policy is used in proposed system. The results in different situations depending upon machine capabilities and available bandwidth is described in tables. The system gives comparatively low bitrate than other systems, but the proposed system is not compared on bitrate, but a priority is shown in maintaining connectivity of the system all the time. The graph of various parameters altered periodically is shown in Graph in Fig. 1.

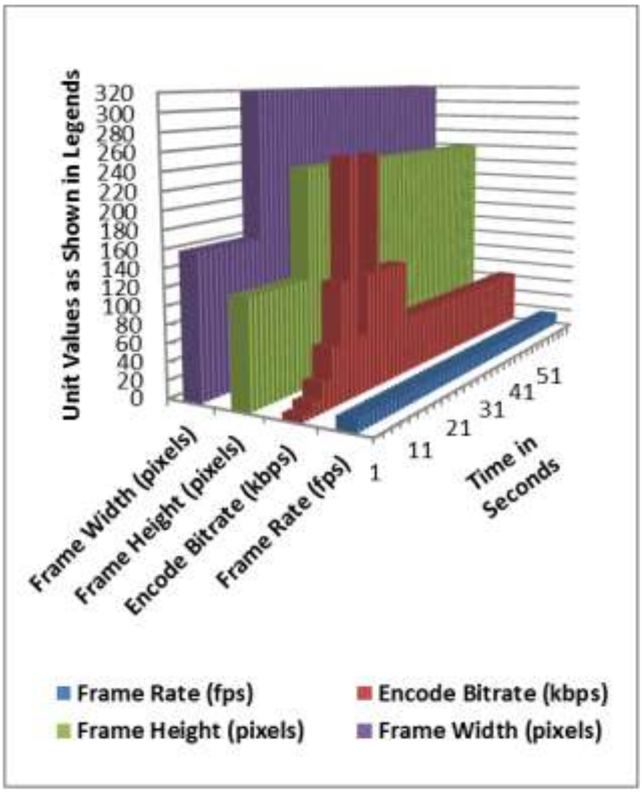


**Fig. 1. Encode Parameters Alteration**

The Graph in Fig. 1 is plotted for width, height, frame rate and bitrate altered periodically in seconds. The results are plotted for system having capabilities of Intel i5 CPU and 6 GB RAM and having total bandwidth of 256 kb/s shared with other machine in communication. The system stabilizes on 320x240 frame resolution, bitrate of 64 kb/s and frame rate of 15 fps. The system is having 128 kb/s upload as well 128 kb/s download bandwidth. The bandwidth is also shared with other machine in communication.The resolution is finalized on 320x240 because the other machine in communication having lower capabilities such as Pentium 4 CPU and 1.5 GB RAM .The encoding and decoding time required for 640x480 resolution on lower capability machine is 80% of total unit time. Hence to cope up with the lower capability machine the other machine has put constraint over sending resolution to 320x240. The resultant video quality is sufficiently good for face to face discussion not requiring minute details to be shared. The benefit with this approach is that the connectivity is sustained for average bandwidth fluctuations. The system fails only when bandwidth drops below 40 kb/s for a longer period or network is disconnected.

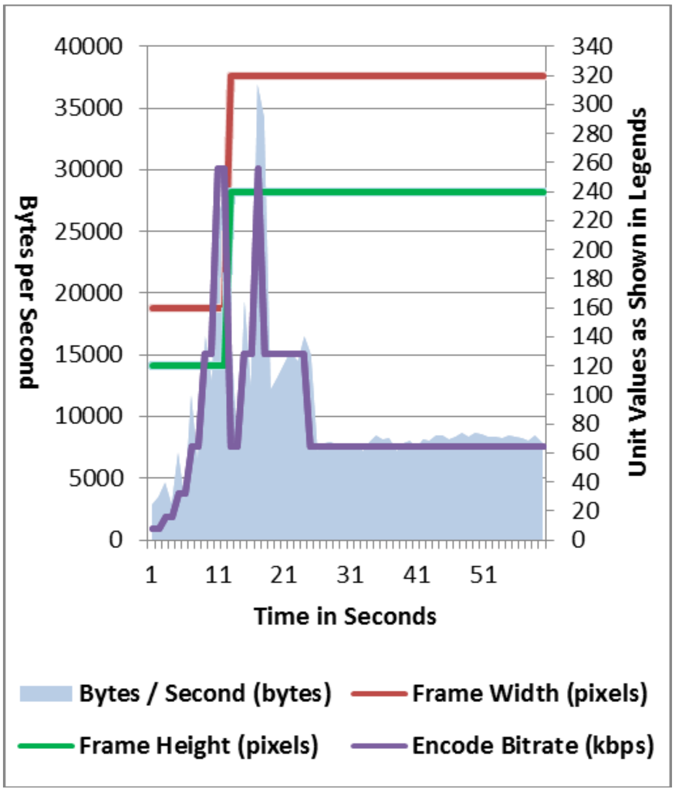


**Fig. 2. Encode Parameters with Bytes/Sec**

The Graph in Fig. 2 is plotted for data bytes transported per second, against width, height and bitrate used per second. The machine capabilities used are same as for Graph in Fig. 1. As per the graph a data transported per second is aligned with the bitrate used per second. Hence the data generated per second is directly proportional to the bitrate used for that slot. Even though due to the varying bitrate used initially and at the end when system stabilizes over 64 kb/s the data generated over total length of communication remains low compared to systems which concentrates over good quality video only. Hence by using proposed approach a sufficiently good quality is maintained with lower transportation cost.

The Graph in Fig. 3 is plotted for width, height, frame rate and bitrate altered periodically in seconds. The results are plotted for system having capabilities of Intel i5 CPU and 6 GB RAM in communication with system capabilities of Intel i7 CPU and 8 GB RAM. The communication is sharing network bandwidth of 1 mbps. The resolution of 1280x720 is adapted for few seconds and system stabilizes over 640x480 resolutions in the end.

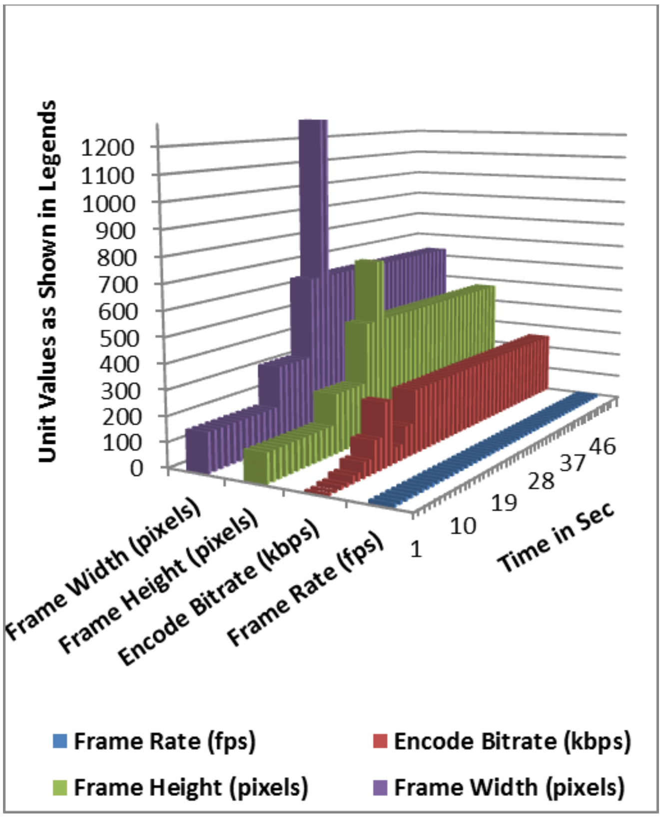


**Fig. 3. Encode Parameters Mapping**

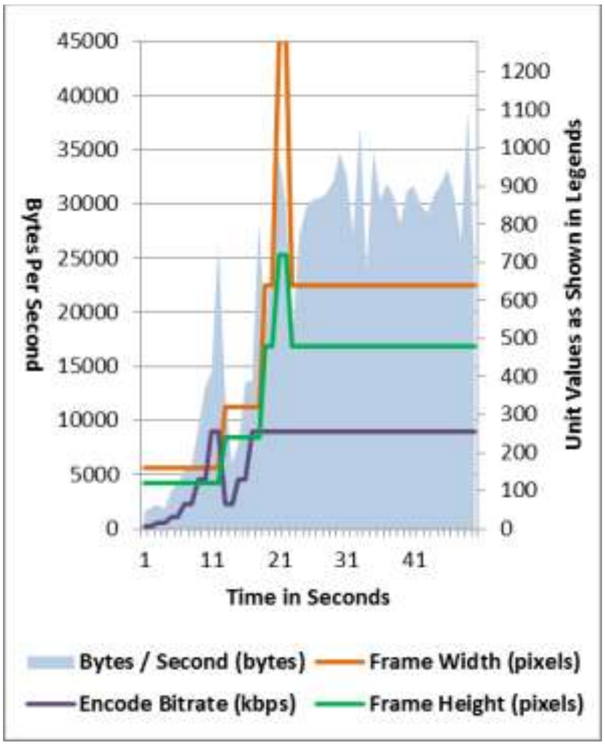


**Fig. 4. Parameter wise Data Transported Mapping**

The Intel i5 capability machine requires more encoding time for 1280x720 resolutions per second than threshold encode time, hence a lower resolution is adapted to stabilize the system [10].It is not a requirement to step down resolution, but doing so also provide reasonably good quality and that to with lower bitrate. The visual difference observed between 1280x720 and 640x480 resolutions is practically negligible. In face to face discussion such compromise is very useful in terms of saving in bandwidth cost at the same time with negligible quality difference. The bitrate is also settled over 256 kbps which is generally suitable to encode a higher resolution frame of 640x480 dimensions as well 1280x720 dimension. Unnecessary use of higher bitrate only increases the transportation requirements.

The Graph in Fig. 4 is plotted for data bytes transported per second, against width, height and bitrate used per second. The machine capabilities used are same as for Graph in Fig. 3. As per the graph a data transported per second is aligned with the bitrate used per second.

Hence the data generated per second is directly proportional to the bitrate used for that slot. Even though due to the varying bitrate used initially and at the end when system stabilizes over 256 kb/s the data generated over total length of communication remains low compared to systems which uses higher bitrates to get good quality. The quality difference is negligible as discussed in Graph in Fig. 3, hence nearly a similar quality is achieved and that to with low bitrate [11].

## VI. CONCLUSION AND FUTURE WORK

In this paper a focus is maintained over how the video encode parameters are intelligently altered such that a discontinuity in communication is avoided as much as possible. The second objective achieved through this approach is that a lower bitrate is maintained for a sufficiently good video quality. Hence a conclusion is drawn from the overall analysis is that very rarely a high resolution and high bitrate is required in Real Time Video communication, a certain amount of low resolution and low bitrate is also sufficient to achieve a nearly good quality video communication.

Currently the system is supported for video communication over Desktop machines and Laptop machines. The system also works within wire line as well wireless network environments. Future enhancements to the system are to be planned so that the system will support mobile devices too. Due to the restricted environment and lower device capabilities associated with mobile devices this type of approach is very much useful for achieving a maintained video communication in mobile devices.

## REFERENCES

[1] A. Duric, S. Andersen, "Real-time Transport Protocol (RTP) Payload Format for internet Low Bit Rate Codec (iLBC) Speech," Network Working Group, RFC 3952, Dec. 2004.

[2] A. Selvan, R. Raman, "MinimizingPacket Loss Using Congestion Control Scheme for Video Streaming," in International Journal of Science and Research (IJSR),Vol. 3, Issue 3, pp. 432-435,.Mar. 2014.

[3] A. Mishra, H.Agarwal, R.Matharu, "NAT Boxes and Firewalls, A Look at VoD and Skype,".

[4] Colin Perkings, "RTP: Audio and Video for the Internet," Addison Wesley,Jun. 2003.

[5] D. Wu, Y. Hou, and Y. Zhang, “Transporting Real-time Video over the Internet: Challenges and Approaches,” in Proceedings of the IEEE, Vol. 88, Issue. 12, pp. 1855-1877, Dec. 2000.

ISSN: 2321-8169
5246 - 5251

[6] J. Bankoski, J. Koleszar, L. Quillio, J. Salonen, P. Wilkins, Y. Xu, "VP8 Data Format and Decoding Guide," Google Inc., RFC 6386, Nov. 2011.

[7] J. Baillieul and P. J. Antsaklis, “Control and Communication Challenges in Networked Real-Time Systems,” in Proceedings of the IEEE, Vol. 95, Issue. 1, pp. 9-28, Jan. 2007.

[8] Z. Wen, Z. Liu, M. Cohen, J. Li, K. Zheng, T. Huang, " Low bit-rate video streaming for face-to-face teleconference," in IEEE International Conference on Multimedia and Expo, Vol. 3, pp. 1631-1634, Jun. 2004.

[9] R.Addu and V.Potuvardanam, "Effect of Codec Performance on Video QoE for videos encoded with Xvid, H.264 and WebM/VP8," Master Thesis, School of Computing Blekinge Institute of Technology.Karlskrona, Sweden, Aug. 2014.

[10] S. F. Chang and A. Vetro, “Video Adaptation Concepts, Technologies, and Open Issues,” in Proceedings of IEEE, Vol. 93, Issue 1, pp. 148-158, Jan. 2005.

[11] TresorMvumbi, Flora Kundaeli, ZafikaManzi, Kyle Williams and Hussein Suleman, "An Online Meeting Tool for Low Bandwidth Environments," in proceedings of SAICSIT, pp. 226-235, Oct. 2012.

[12] W.Geisler and J. Perry, "A real-time foveatedmultiresolution system for low-bandwidth video communication," in Proceedings of SPIE, Vol. 3299, pp. 1-13, 1998.

[13] Z. Hu, "NAT Traversal Techniques and Peer-to-Peer Applications,"Seminar on Internetworking, Apr. 2005.