

Investigation and Evaluation of high quality video streaming for mobile devices and web portals

Ms. Prema P. Gawade.

(M.E. Student)

Asst. Prof. P.B. Kumbharkar.

(Head Of Computer Engineering Department.)

Siddhant College Of Engineering, Pune

I. ABSTRACT

Presently the use of TV and mobile video is emerging and hence people can able to access the digital contents like TV shows, sports videos, TV shows online etc. from anyplace and this present methods of video streaming however having the limitations like shared resources among many end users, means bandwidth is shared among many of the users which is resulted into videos access with limited resolutions for the end users. There are many new devices presented in market recently in order to support the high resolution mobile devices by Google, Sony, Apple etc. but video streaming resolution is very poor which cannot support these new mobile devices. Hence this is resulted into the introduction to visual artefacts. Thus, to overcome this problem and provide the high quality video streaming for mobile devices, in this paper we will investigate the proposed method. In this paper the method proposed to bridge the resolution gap between end user mobile screen and video streaming. This investigated approach is based on novel upsampling-based system architecture to enable high-quality video streaming onto mobile devices. This proposed approach is considering the mobile video with higher resolution and then only sending over the web portal or wireless network. We consider applications in which a mobile video has a counterpart with the higher resolution for transcoding before sending it over the Internet or synchronizing it with a mobile device. And hence this resulted into less computation time and efficient video streaming as compared to previous approaches. We evaluated this proposed approach using the Java Video Streaming API over Internet and analyzed the results for proposed approach.

Index Terms— *Audiovisual Quality, Framerate, Bitrate, Picture Ratio, Video, Quality, Mobile Device.*

II. INTRODUCTION

The era of TV and mobile video now a day is growing rapidly. The mobile television is already adopted and used in many countries, however its proper adaptation has to be

taken. Multimedia messaging is also expected to increase in popularity. Although these services look appealing, the end-users' subjectively perceived quality of them is a critical factor for their success. Consequently, subjective quality evaluation tests during product development are necessary if the service is to reach an acceptable quality level. Subjective audiovisual parameters related to mobile television have not been widely studied. There are only a few published studies about subjective evaluation of low frame rates, bitrates and modern codec's used with devices with small displays. The relation between the audiovisual contents has, until recently, also been a relatively unexplored area [1].

In this particular research paper, we are presenting and investigating one of the recently presented methods for delivering the high quality video streaming for mobile devices and web portals. This method bridges the resolution gap between videos and the definition of mobile device screens [1]. Current image and video upsampling techniques are subject to a tradeoff between the quality and computation time because the information about high-resolution images and videos is not given *a priori* in the corresponding applied scenarios [2]. And hence objects details and their boundaries information for the high resolution version is not known whenever we up-sample the low resolution version that increases the overall computation time. In opposite to this, we take applications those having the mobile videos with higher resolution for transcoding before sending it over the Internet or synchronizing it with a mobile device. Therefore, the recent method which we are discussing in this paper prevent this dilemma because we extract the information about high-definition videos during the transcoding and send the metadata to clients to facilitate upsampling on mobile devices, which do not have sufficient computation resources for complex upsampling techniques. The metadata enables low-complexity video upsampling on the client side, as the bulk of the computation is shifted to the server side [1].

There are two main points those we are discussing over here, first is the literature review over the role played by codecs in the quality of video streaming as well as relation of audio and video streams on audiovisual quality. Second

thing, we will discuss the recent method for high video streaming.

Below section III presents the literature review over the video quality, section IV presents the discussion over the research methods for evaluation. Section V presents system architecture for new method for high quality video streaming along with its work done in section VI. Finally in Section VII we will discuss the conclusion over the study.

III. LITERATURE REVIEW

A. Perceiving and Producing Quality

This research is focusing on the video streaming quality for mobile devices and web portals. Hence in the literature it's necessary to us to first present the study of what makes the quality of video.

Perceived Quality and Produced Quality is a combination of Video Quality. It is an acceptable to pleasing levels and maximum perceived quality and also integrated set of perceptions of overall excellence of audiovisual material that unacceptable vary from noticeable. The produced quality is usually between an acceptable and a low level of quality in mobile video [3].

Quality perception varies according to the sensory channel. It can be reduced to sensory thresholds and modulation transfer functions (MTF), if the quality experience is considered at the sensorial level. However, human quality interpretation depends on the relationship between higher level processing and sensory processing in which knowledge, emotions, expectations and situational schemas are included [4].

In audiovisual perception, auditory and visual information are integrated into a unified perceptual experience just add the two perceptual channels in a more complicated way. For example, in the McGurk effect mismatched acoustics material and visual are integrated into a single audiovisual experience. It also correct for unified perceptual experience to bind the image and sound together but there is synchronization is necessary.

Good audio quality can enhance visual quality and vice-versa by showing recent studies. Importance of auditory and visual information depends on content... For example, Winkler & Faller found that the importance of the auditory channel increases when the complexity of an audiovisual scene increases at very low bitrates. Changing bitrates out of the accepted ranges of the different modalities (audio 16-24 kbps; video 32-40 kbps) lowered quality. Multimedia quality model, talking head content as both modalities have approximately the same affect the quality judgment. And the video quality is weighted more towards high-motion content.

2. Quality production

There are limited bandwidth and the limitations of the devices in mobile video production (e.g. display size, processing capabilities) are the main issues. Frame rate and frame quality are the factors that affect video quality and also spatial resolution, with 3G mobile networks, low framerates questions of low bitrates and small screen sizes are the most crucial.

Codec reveal bitrate and the frame quality of video. Latest video compression standard, H.264, network adaptation thus enabling a significant reduction in bitrate compared to other released standards and excellent coding efficiency such as H.263, MPEG-4 [5]. For example H.264 has features for decreasing visible coding errors (e.g. de-blocking filters and small block sizes).

The relation between the perceived quality and framerate is not linear and is dependent on the content used. By recent study sufficiently high and the content was personally significant was sports fans tolerated framerates as low as 6fps as long as single frame quality. As distinct snapshots at low level objective framerate as a temporal resolution of video appears to the perceived as natural motion at high level. (6-24fps/QCIF) These sports videos were taken on handheld devices. At low bitrates (24-48 kbps), a frame rate of 15kbps at the same bitrate from a PC with QCIF frame size, but a frame rate of 8fps gives better video quality.

Modern audio encoding e.g. MP3 and AAC is based on efficient perceptual coding. Low bitrates mono is more pleasing than the spatial parameters stereophonic sound; the most common impairment appears as preceding noise, sound of double recording typical unpleasant distortions with the use of headphones and also at low bitrates. Main parameters affecting quality are the temporal parameters and stereophonic sound relevant to the sampling rate or spatial parameters relevant to the monophonic [6].

IV. LITERATURE REVIEW

Firstly, the video study investigated the video quality factors and secondly, the audio-video study explored the combination of audio and video. These research methods were based on subjective evaluation of video and ITU-T recommendations for audiovisual quality and it used in two different studies.

1. Participants:

60 in the audio-video study and 75 subjects participated in the video study. They were stratified according to age (18-65 years) and sex and the number of professional evaluators was restricted to 20%.

2. Test procedure:

Before the test, participants were shown the lowest and highest quality samples as an example of the quality scale and they use for the test. In test, the material was shown

using the single stimulus method where the clips are viewed one after one and rated independently. Participants marked the quality score of a clip on an answer sheet using a discrete scale from 0 to 10. The test was followed by an interview concerning content recognition, interest in the content and an open interview on the participants' evaluation criteria during the test. Participants were screened for color vision and hearing, visual acuity (20/40). And also surveyed for demographics and briefed about the test procedure [7].

3.3 Selection of Test Material:

The richness of spatial resolution and temporal was the criteria for clip selection for the content of each category selected. The contents presented in Table 1 these tele-text was used only in video test Popularity and Resolution in which the main criteria in the selection of the content. The test materials were chosen according to Finnish TV - broadcast ratings from popular TV-programmers from the Finnish broadcasting network. The selected materials were also suitable for mobile TV broadcasting [7].

Table 1. The content selection for the test

		Spatial details	Temporal motion	
	Evening news	Low	Low	Speech
Sport	Ice hockey	Medium	High	Speech
Series	CSI	Medium	Medium	Speech
Animation	Simpsons	Low	High	Speech
Tele-text	Newsfeed	High	Low	-
Music video	Rock music	High	High	Music

4. Test Material Production Process:

Original material for clips was sourced from midi DV-tapes and DVB MPEG-2 and converted to PAL format AVI frames (Interview Win Producer (3.0B001.111C2A). these AVI frames used as the input to produce the sample clips. These original audio samples (stereo, 32 kHz) were normalized and converted to 16 kHz sampling rate mono. The parameters for 10 second -long sample clips are showing in Table 2 and the encoding tools in Table 3.

5. Presentation of Test Materials:

For audio playback the headphones supplied with the devices used. These devices attached to a stand and the viewing distance was set to 440mm. Set according to ITU recommendations general viewing conditions for the

laboratory. Two devices were used in both studies and the starting devices were randomly selected. The loudness of audio signal from the headphones to ear was adjusted to 75dB by using a human ear simulator [8]. All these clips were played from the device memory. The Sony-Ericsson a P800 Smart Movie and Nokia 7700, 6600 used a Real One Player.

VIDEO TEST						
Device	Nokia 6600			S-E P800		
Video codec	H.263	Real Video 8	H.264	XviD		
Picture ratio	QCIF	QCIF	QCIF	QCIF	QCIF	SIF-SP
Bitrate	80	80	80	80	128	128
Framerate	12,5					
Contents	6					
AUDIO-VIDEO TEST						
Device	Nokia 7700		Nokia 7700 S-E P800			
Codecs	H.263-AAC		H.263-AAC XviD -MP3			
Total bitrate	100		160			
Audio bitrate	24	16	64	32	16	
Video bitrate	76	84	96	128	144	
Framerate	6	12,5	12,5	12,5	12,5	
Contents	5					
Picture ratio	QCIF					

Table 2: Parameters for both tests

Table 3. The codecs and programs used in encoding

Codec	Program
H.263, AAC	Nokia Multimedia Converter Pro 2.0
RealVideo 8.0	RealProducer 8.0.0.45
H.264	Customized software
XviD SmartMovie, XviD Codec, Lame MP3 encoder	SmartMovie 2.20

V. INVESTIGATED METHOD ARCHITECTURE

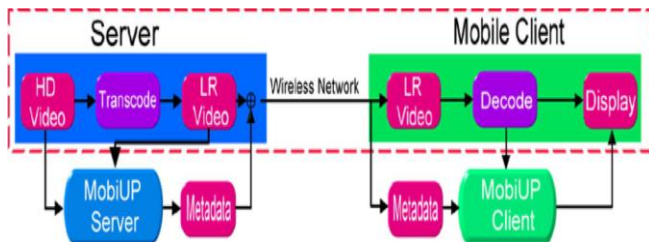


Figure 1: System Architecture for MobiWebUP

As in above figure 1, it shows the conceptual architecture of MobiWebUP [2]. The part enclosed by the dashed line is the framework of the existing video streaming process. That is, on the server side, a high-definition (HD) video is first transcoded into a lower resolution (LR) version with appropriate bit rates. It is then delivered over the Internet to the target mobile client. When receiving the bitstream, the mobile client decodes the bitstream into raw frames and displays them on the screen. In contrast, MobiWebUP enables good quality conversion from LR to HD videos in real time on the client side, because we first summarize specific metadata on the server side and then leverage the information on the client side for upsampling [2].

In case of server side, the extraction of metadata from the HD video is done and sends it with very less bit rates with transcoded LR video. To extract the metadata, we first segment the HD video into shots and label each shot with the upsampling method that yields the best visual quality. We also identify the boundaries of objects and important details in video frames because those parts are the main sources of visual artifacts. Next, we send a summary of the above information to clients to facilitate good quality upsampling in real time. It is worth noting that MobiWebUP does not need to modify existing codecs. Indeed, the proposed upsampling architecture can be regarded as a complement to existing schemes. Therefore, MobiWebUP is generic and flexible, and it can be implemented facily for practical use.

MobiWebUP can improve the user perception of the current mobile devices with any codecs, because we leverage the available computation resources to upsample decoded frames. Compared with scalable video coding (SVC), MobiWebUP computes the high-resolution version from the low-resolution video, instead of directly decoding the video with a higher resolution and a higher bit rate, since current mobile handheld devices and laptops are equipped with powerful processors, such as iPhone 4G (Apple A4 1 GHz), Google Nexus One (Qualcomm QSD8250 1 GHz), HTC HD2 (Snapdragon™ 1 GHz), and FUJITSU S6510 (Intel Core™ 2 Duo T7700 2.4 GHz).

To the best of our knowledge, the proposed MobiWebUP system is the first integrated video streaming architecture that attempts to improve the video experience of current mobile users by employing real-time video upsampling with good user perception.

VI. WORK DONE

As per given in paper [2], the implementation of MobiWebUP approach is done. For investigation purpose we calculated the visual quality for proposed method and compared its performance against the existing methods. Figure 2 below showing the same results from [2]. From this result we can say that this investigation approach avoiding the conspicuous artifacts on the boundaries of objects like sun and river. In addition, MobiWebUP can identify the important details and textures of the land in the figure. The clarity of the MobiWebUP video is better than that of the video upsampled by the comparing approaches under the same data rate. As shown in the figure, the resolution of the LR video is too small for mobile devices; hence, many jagged and blurred artifacts are generated by bilinear upsampling. On the other hand, proposed methods prevent conspicuous artifacts and yield a better quality video.

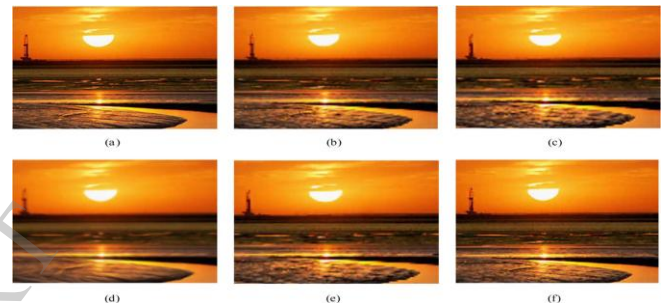


Figure 2: Comparisons of different approaches given the same frame of a 384-kbps LR video: (a) ground-truth frame in an HD video, (b) bilinear-upsampled result, (c) bicubic-upsampled result, (d) NEDI-upsampled result, (e) IENE-upsampled result, and (f) MobiUP-upsampled result. Note that, in our MobiUP approach, the total bit rate 384 kbps is constituted by both the data rate of the LR video and the associated metadata.

Regarding to performance metrics, we considered PSNR and SSIQ. Fig. 3(a) and (b) compares MobiUP and existing approaches in terms of the PSNR and SSIQ, respectively. The results demonstrate that MobiUP outperforms the existing approaches because it leverages the information about HD videos to ensure good quality upsampling in real time.

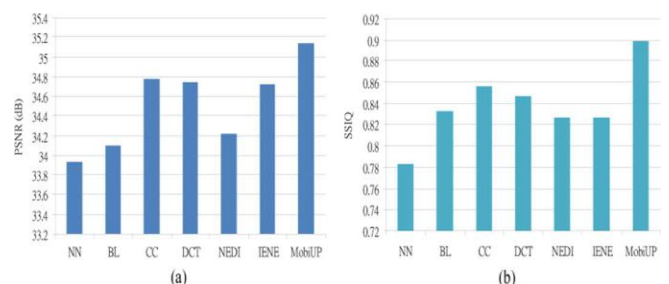


Figure 3: Comparison in terms of (a) PSNR and (b) SSIQ with a scaling factor of 4.

Fig. 4(a) measures the ratio of the amount of metadata to the total amount of transmitted data. The results show that the ratio varies with different types of video content. Whereas

fig.4 (b) compares the frame rates of MobiWebUP and the existing approaches. Four kinds of test videos in different content categories are used in our experiments, i.e., music video, animation, sport, and news. This shows again that our proposed method is far better as compared to existing methods.

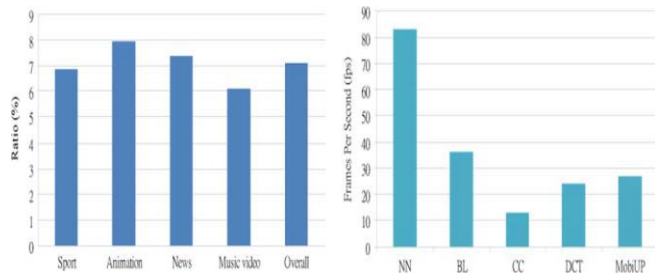


Fig. 4. Comparison in terms of (a) the ratio of generated metadata to the total transmitted data and (b) efficiency.

Apart from this, there are some more performance metrics are considered for the performance evaluation of proposed approach. Out of this, one can also need to present in this investigation paper. The battery life time of mobile device is also major factor to consider. Below table 4 compare the battery life of different approaches. For fairness, the brightness of the LCD backlight is set to 100% in all approaches. The test videos are the same as those used in the frame rate experiment. We upsample videos in the loop mode until the power are depleted. The results show that displaying videos with MobiWebUP consumes slightly more battery power than displaying videos with bilinear upsampling.

COMPARISON IN BATTERY LIFETIME

Displaying schema	Lifetime (seconds)
Do nothing	15760
Original LR size	13428
Bilinear upsampling	12394
MobiUP	10627

Table 4: Comparison in Battery Lifetime

VII. CONCLUSION AND FUTURE WORK

During this research study, we presented the literature review over the video quality measurement approaches along with research methods for quality measurement of video streamlining. After that our main motive was to discuss architecture and results of new recently present method called MobiUP in [2], which we renamed to MobiWebUp because we will now produce the same method for web based video streaming applications as well. From the above discussed results, this proposed method outperforms the existing approaches for the generation of high quality video streaming. The metadata generated by MobiWebUP account for less than 8% of the total transmitted data, but it reduces conspicuous artifacts

significantly compared with the existing approaches. In addition to this, proposed approach in [2] doesn't need to Modify current codecs for video streaming. This investigation method also provides the avenues for further exploration. With the improvement of CPU speeds, more complex upsampling techniques will become feasible. Therefore, MobiWebUP can incorporate more parts of the existing SR techniques to enhance the video quality and reduce the amount of metadata. In future work we will explore this approach in details with improving the base results.

VIII. REFERENCE

- [1] "Evaluation of Subjective Video Quality of Mobile Devices", Satu Jumisko-Pyykk and Jukka Häkkinen, 2011.
- [2] "MobiUP: An Upsampling-Based System Architecture for High-Quality Video Streaming on Mobile Devices", Hong-Han Shuai, De-Nian Yang, Member, IEEE, Wen-Huang Cheng, Member, IEEE, and Ming-Syan Chen, Fellow, IEEE, 2011.
- [3] Beerends, J.G. & de Caluwe, F.E. The influence of video quality on perceived audio quality and vice versa. *Journal of the Audio Engineering Society*, 47 (5), 355-362, 1999.
- [4] Brandenburg K, MP3 and AAC explained, AES 17th International Conference on High Quality Audio Coding. Italy: September 1999.
- [5] Casey, B., Casey, N., Calvert, B., French, L., Justin, L. *Television Studies – The Key Concepts*. London: Routledge, 2002.
- [6] Finnpanel. Television audience measurements. <http://www.finnpanel.fi>. (visited 05/2004).
- [7] Fukuda K. Integrated QoS Control Mechanisms for Real-Time Multimedia Systems in Reservation-Based Networks, PhD Thesis. Osaka University; 2000.
- [8] Hands, D. S. A Basic Multimedia Quality Model. *IEEE Transactions on Multimedia*, Vol. 6, No. 6, December (2004). pp. 806-816.
- [9] I. J. Cox, S. Roy, and S. L. Hingorani, "Dynamic histogram warping of image pairs for constant image brightness," in *Proc. ICIP*, 1995.
- [10] R. Dugad and N. Ahuja, "A fast scheme for downsampling and upsampling in the dct domain," in *Proc. ICIP*, Oct. 1999, vol. 2, pp. 903-913.

[11] R. Dugad and N. Ahuja, "A fast scheme for image size change in the compressed domain," *IEEE Trans. Circuit Syst. Video Technol.*, vol. 11, no. 4, pp. 461–474, Apr. 2001.

[12] V. Essen, C. Anderson, and D. Felleman, "Information processing in the primate visual system: An integrated system perspective," *Science*, vol. 255, pp. 419–423, 1992

IJERT