Asian Journal of Information Technology 15 (18): 3474-3479, 2016
ISSN: 1682-3915

Video Transcoding by Dynamic Frame Skipping Mechanism and Offloading Through Decision Logica

[1]B.L.Velammal, [2]P.Anandhakumar and [1]Varun Gowrishankar
[1]College of Engineering Guindy, Computer Science and Engineering, Chennai, India
[2]Institute of Technology, Computer Science and Engineering, Chennai, India

Abstract: In last few years, there has been a rapid development in mobile communication technology. Developments in this technology has grown in heterogeneous manner, meaning that there have been varied mobile devices with different architectures supporting versatile formats such as WMV,FLV,MPEG. Moreover, advents of high quality mobile services such as streaming high definition videos, on Demand TV, etc., caused a great strain in network bandwidth allocation and processing capability. Ourstudy proposes a transcoding scheme through frame skipping by analyzing frame details via a dynamic window selection mechanism for proficient usage of network bandwidth. Analysis of frame intricacies consists of two sectors one dealing with calculating the importance of each frame and other deals with computing the motion magnitude between frames. This transcoding operation is offloaded through a Decision Logica function to decrease the computation work in the host system. Experimental results demonstrate that the proposed mechanism for transcoding resulted better precision in selecting frames for encoding.

Key words: Video streaming, transcoding, dynamic frame skipping, decsion logica, multimedia

INTRODUCTION

Mobile technology is used in cellular communication which is not only used for voice communication but also used extensively in delivering multimedia contents like images, audio files, video streaming etc. The huge surge in electronics industry resulted in lowering the cost of mobile devices and enhancing its features; it leads to a direct consequence of different mobile devices supporting various multimedia formats. One such example is video, there are many number of video formats ranging from mp4 to fly each running in their respective smart mobiles which supports it (In-Stat, 2011). Video content providers have also moved their production framework to high definition video which resulted in a drastic increase in network bandwidth exponentially (Cisco, 2016). All these factors made transcoding an imperative necessity for any mobile devices/desktop to have a seamless video transmission and streaming. Transcoding is a direct analog to analog or digital to digital conversion of one encoding to another or the same.

This transcoding operation can also be used to reduce the bit rate of the video to adapt the different network and device environments. It also plays a key role in providing Universal Multimedia Access (UMA) (Xu *et al*., 2012). A various transcoding methods have been previously proposed for reducing the bitrate of the video. Some of them are requantization, frame skipping, spatial resolution reduction and combination of these (Shu and Chau, 2004).Phi phenomenon is being applied in frame skipping mechanism to create an optical illusion for displaying a quality output with less bitrate. In past various algorithms has been proposed to transcode the video for both rate adaptation and quality control. Huge number of experiments have already been done in regular frame skipping, dynamic frame skipping and it has been found that it produces jerky/choppy effect in transcoded O/P for regular frame skipping. Metrics used for evaluating the transcoded video are PSNR, VQA (VIDEO Quality Assessment), MOVIE (Motion Based Video Integrity) (Yeh *et al*., 2013). Dynamic frame skipping is a computationally complex process which makes a low end mobile device incompatible for transcoding. To eradicate this problem GPU (Graphical Processing Unit), Offloading server can be introduced to do the transcoding operation.

Literature review: Many tens of papers have been published in transcoding as well as computation offloading. Rudenko *et al*. (1998) first proposed computation offloading technique to lessen the computation overload. He concluded, even if some overhead was found, it was better to have an overloading server. He concluded overhead is negligible, if

Corresponding Author: B.L. Velammal, College of Engineering Guindy, Computer Science and Engineering, Chennai, India

computation performed is so high that made mobile device lower its time performance or battery power (Chen *et al*., 2004). Kumar and Lu (2010) checked whether offloading it, saves energy as claimed or merely a mirage. This is one of the important work, that helped to show the importance of offloading a system; it explains the importance of the offloading framework but also highlights the issues and solutions regarding the offloading framework like privacy and security of any cloud server which comes with the offloading framework (Rudenko *et al*., 1998).

Mobiles are resource constrained devices, so we need to optimize every bit of resource we use in it. One of innovative way namely Distributed Execution Transformer (DIET) was proposed by Kim *et al*. (2009). It provides application developers and application providers a transparent way to program and deploy (Ahmad *et al*., 2005).

Basics of transcoding framework were explained by Ishfaq Ahmad and Yu sun in their study which explain the basic issues of transcoding. Categorization of video transcoding was also put forwarded by them. Two main categorization of transcoding is homogenous and heterogeneous; sub categorization can also take place. In the homogenous transcoding there is temporal resolution adjustment, spatial resolution adjustment and in heterogeneous transcoding main part is between standard transcoding and interlaced and progressive transcoding (Kim *et al*., 2012). Haiyan Shu and lap puchan proposed a downsizing algorithm, itwas very much necessary to have a good video content delivery system in a wide heterogeneous environment. He proposed a system which works in a still image environment and also in a video content delivery system making it very much valuable in the field of multimedia transcoding (Shu and Chau, 2007). There has been various rate adaptation techniques introduced from the start of the mobile technology's cradle, some of them were frame loss based rate adaptation .The rate adaptation was based on the frame loss in the network being transmitted and in Signal strength based rate adaptation the signal strength of the wireless network will determine the rate of the transmission of the signal (Hsu *et al*., 2009).

In Bit rate reduction, the objective of bit-rate reduction is to reduce the bit ratewhile maintaining low complexity and achieving the highest quality possible, application of the bitrate reduction can be seen from television broadcast to internet streaming. The most straightforward way to achieve this is to decode the video bit stream and fully re-encode the reconstructed signal at the new rate (Kumar and Lu, 2010).

Guanguchen, Narayana vijayakumar studied the energy trade off in computation offloading of java enabled devices; java has been a forefront in mobile device technology making it synonym to multimedia mobile devices. Computing the computation energy in offloading is very important task, calculation of energy takes in to account the computation energy, compilation energy and overhead energy spent. This helps primarily in deciding whether offloading should be in client or server (Kim *et al*., 2006).

Offloading framework for transcoding will incur overhead in energy consumption and to avoid this Namgon Kim and Jae introduced a visual sharing switching device which supports in-network content adaptation. Visual sharing switching device was an extension of a residential gateway, which was equipped with GPU (Graphical Processing Unit).This switching device provide a flow based transcoding i.e., the transcoding will happen, once it gets the data from the client and won't wait for full data to be received; but it should wait for remaining full set of data before starting the transcoding process (Ni and Cao, 2012). Spatial resolution reduction and temporal reduction transcoding can be applied alternatively in the gateway itself and by combining these 2 important methods we can improve the QOS (Quality of Service) for the user. One can imagine following scenarios:

- Client-a remote user, has a small display panel such as mobile phone, the spatial resolution reduction transcoding can be selected
- Temporal Resolution Reduction transcoding method can be applied whenthe client has a sufficient size of display panel but the communication channel has low bandwidth (Biaz and Wu, 2008)

Ching-Ting,Chia-Hung, Chao-Yu worked with temporal and spatial complexity in a video transcoding mechanism. In their model, motion vector composition algorithm was designed to reduce the computation on motion estimation process (Hsu *et al*., 2009).

MATERIALS AND METHODS

Proposed system: In this study, we propose an offloading framework and a transcoding computation for a video input as shown in Fig. 1.

In the offloading framework, we propose a Decision Logica function to determine whether to execute the transcoding computation in host system or in an offloaded Server. In our transcoding computation mechanism , we propose a dynamic window selection

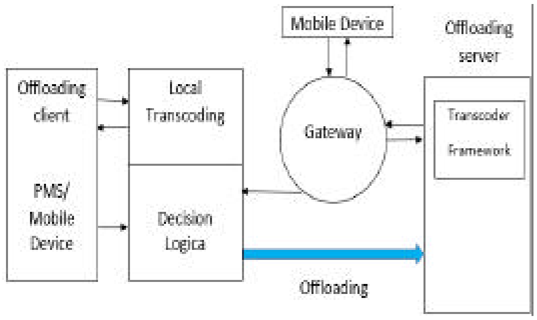

Fig. 1: System design

cheme which determines the window size for which our transcoding operation is performed. The proposed transcoding computation algorithm deals with frame skipping scheme.

Decision logica: Since transcoding is a computation intensive process we offload but offloading itself has a large overhead attached to it; making it unnecessary to transcode certain videos which can be transcoded in the host system itself in much more efficient way. This can be implemented by proposed Decision Logica function.

Decision Logica function computes P_i, i.e., power consumed by mobile device when the device is idle, P_c, i.e., power consumed when it does computation operation or processor operation and P_{tr} which is the power used for transmission of the video input when it transmits the video to the offloading server. Using above 3 major parameters, we compute the Decision Logica by following Eq. (1) which has following parameters as follows:

$$\text{Decision Logica}=Pc*(C/m)-P_i*(C/s)+P_{tr*}(D/B)) \quad (1)$$

- S: Speed /instruction per second for Offloading server
- m: Speed /instructions per second for mobile device
- C: No. of instruction to be computed
- D: Size of the data

When the value of the decision Logica is positive transcoding operation is offloaded, if not, it is done locally in the host system itself.

Window selection mechanism: There has always been the tradeoff between quality of the video to be transmitted and network bandwidth usage.

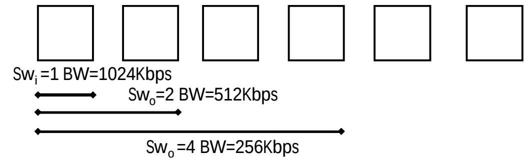

Fig. 2: Window selection mechanism

Proposed dynamic frame skipping mechanism considers the channel bandwidth as a key factor for determining number of frames to be encoded. This can be implemented by window selection mechanism. Window size is determined based on Eq. 2:

$$SW_o = (N_i / N_o)*SW_{i,} \quad (2)$$

where, SW_o is the new window size, SW_i initial window size and N_i, N_o, represents the original bitrate of the pre-coded video, bandwidth obtained from the feedback of the channel, respectively. Selection Mechanism dynamically changes the size of the window taking the current bandwidth of the network from the feedback component. Figure 1 explains the change in window size dynamically, initial window size is set to 1 where the bandwidth is 1024 Kbps (max) as Bandwidth decreases to 512 Kbps the Window size shifts to 2 and this process continues as bandwidth randomly alters (Fig. 2).

Frame skipping based in frame intricacies: The proposed frame skipping mechanism is based on frame intricacies which consist of 2 parts-one dealing with visual magnitude Impact (VMI) (deals with importance of each frame) and other part works with motion magnitude between frames (MMI).These two impacts are implemented after determining the window size by the proposed window selection mechanism that is explained in (Section B).For each frame (t), Frame Intricacies value (FI) is determined using the Eq. 3:

$$FI_t = VMI_t + MMI_t \quad (3)$$

Visual magnitude impact is calculated with the help of saliency image generator. The impact is based on the following phenomenon."The visual system provides us a huge amount of information for processing in real time; humans have developed an attention system that allosws them to filter out non important sections of the scene by just focusing on the most salient parts of what is being observed". The retina of the human eye contains two type of ganglion cells, one which respond to bright areas

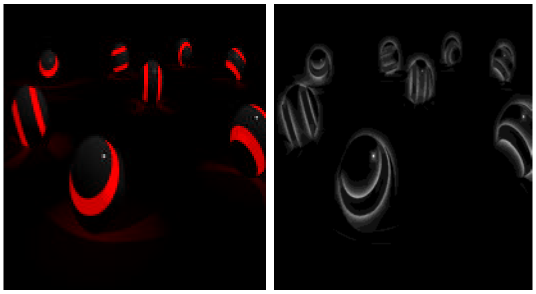

Fig. 3: Original image and its corresponding salient image

surrounded by a dark background and another which responds to dark areas surrounded by bright background. Salient image is using the Eq. 4:

$$S_t\text{=Intensity On}_t + \text{Intensity Off}_t \tag{4}$$

Intensity On and intensity Off correspond to the images obtained by each ganglion cells respectively. Figure 3 represents the salient image and its corresponding original image. Visual magnitude Impact and Visual Magnitude(VM) are represented mathematically following 2 equations (Eq. 5 and 6), respectively.

$$VMI_t \sum_{i=0}^{sw0-1} VMj \tag{5}$$

$$VM_t \sqrt{\frac{\sum_m^W \sum_n^H (M(m,n) - M)^2}{W * H - 1}} \tag{6}$$

Algorithm:
Input: Frame (type: color or grey scale)
Output: Salient image

Step 1: Original color image is converted in to grey scale image.
Step 2: The converted image is smoothened by applying 3*3 Gaussian filter (this process is applied twice)
Step 3: Integral Image is calculated for the given frame
Integral image will be used for 5th steps
I(x,y) =
Step 4: Image is subsampled, that subsampling is done 4 times resulting in(S0 to S4)

Following computation is done for subsampled images S2 to S4
Step 5: Intensity ON(x,y) =
Max{Center(x,y,s) – Surround(x,y,s,
Step 6: IntensityOFF(x, y) = max {Surround(x,y,s,
Step 7: Across scale sum is done, resulting in intensity onand intensity off
Step 8: Intensity on and intensity off are sued to get the salient image.

This algorithm is used for generating salient images. Where M (m,n) represents the contrast value at the coordinates(m,n) of the salient image which was obtained

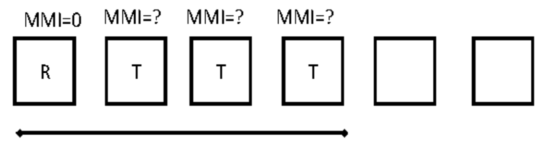

Fig. 4: Reference frame (R) and Target frame (T) in the Selection Window. Motion Magnitude Impact of reference frame is always zero.

using the Algorithm 1,M bar represents average of the contrast values in the saliency map. W and H represent the Width and Height of each frame.

Visual magnitude deals with importance of each frame with the help of saliency map and Motion Magnitude. It also helps to determine the frame needed to skip. This can be done by relying only on visual Magnitude to avoid jerkiness in the output video. Calculating the Motion Magnitude involves determining the motion vector magnitude between 2 frames in the sliding window.

Motion vector can be calculated by various ways; in our system we have used 2 approaches one using Optical Flow farneback Algorithm and other involving careful block matching algorithm. In the Sliding window first frame is referred as reference frame(R) and all other frames are refereed as target frames as show in the Fig. 4.

Motion vector calculation involves determining the motion vector for each MB in the target frame. Motion vector for target MB is determined by comparing the target MB with the best matched MB in the reference frame. Brute force approach involves comparing each MB in the target frame with all MB in the reference frame for finding the best match.

This is not only time consuming but also a computation intensive operation. To avoid this proposed careful block matching is applied, here it involves comparing the target MB only with MB of a particular region called Selected Region in Reference frame. As we know the coordinates of the MB in target frame we can determine the selected region as a region surrounding the same coordinates in the reference frame.

More efficient way of determining the motion vector between consecutive frames is optical flow Farneback algorithm which is applied. Motion Magnitude Vector is calculated by following formulas:

$$MMI_t = \frac{MM_t}{\sum_{j=0}^{SW} MM_t} \tag{7}$$

$$Mm \sum_{i=0}^{NumberofMB-1} TC(FramejandMBi) \tag{8}$$

$$MM(Frame^j and MB^I) = DET(MV_{MB,I,X}) + DET(MV_{MB,I,Y}) \quad (9)$$

Equation (7-9) are used to determine motion magnitude Impact using careful block matching method:

$$MM(Frame^j) = DET(MV_{j,X}) + DET(MV_{j,Y}) \quad (10)$$

Equation 7, 10 are used to determine motion magnitude impact using optical flow farneback algorithm. Frame Intricacy for each frame is the sum of visual magnitude impact and motion magnitude impact of that frame as per Eq. 1. The frame intricacy value is calculated for each frame in the sliding window. The frames with low intricacy values are skipped and the ones with high frame intricacy values are encoded.

RESULTS AND DISCUSSION

In this study, various experiments are conducted in order to evaluate the performance of offloading framework and the proposed dynamic frame skipping mechanism. The proposed transcoding mechanism is applied for certain benchmarks videos like Big Buck Bunny, Bridge (close), Forman, The coastguard and Tsunami.

Comparison of PSNR: Error metrics that are used to compare the various image compression techniques are the Mean Square Error(MSE) and Peak Signal to Noise Ratio (PSNR):

$$MSE = \frac{1}{MN}\sum_{y=1}^{M}\sum_{x=1}^{N} I[(x,y) - I'(x,y)] \quad (11)$$

$$PSNR = 20 * \log10 (255 / sqrt(MSE)) \quad (12)$$

where I(x,y) is the original frame, I`(x,y) is the transcoded frame and M,N are the dimensions of the frames. Higher value of PSNR is good because it means that the ratio of Signal to Noise is higher.

Table 1 and 2 represents comparison of PSNR for different input videos and transcoded videos obtained using different frame skipping mechanisms in a network having a bandwidth range of about 1024 -512 kbps and 1024-256 kbps, respectively.

Comparing the average of PSNR values from Table 1 and 2 for each Frame skipping mechanisms, we conclude that optical flow method of Frame skipping is efficient, as it has higher PSNR value.

Table 1:PSNR Comparison for1024-512 kbps for I/P videos

Sample/Method	Regular	Careful block matching	Optical flow
Frog	76.7781	71.64	73.01
Penguins	45.8625	40.24	43.82
Ping-Pong	41.974	31.52	38.09
Atom-Bomb	24.3194	21.85	25.56
Tsunami	26.1832	23.47	24.002
Average	42.82344	37.744	49.9578

Table 2: PSNR Comparison for1024-256 kbps

Sample/Method	Regular	Careful Block matching	Optical flow
Frog	63.9326	58.832	65.1691
Penguins	42.1728	41.674	42.9599
Ping-Pong	37.9946	37.568	38.875
Atom-Bomb	26.567	23.36	25.45
Tsunami	21.459	20.86	23.99
Average	37.864	36.45	38.68

Table 3:Comparison of number of coded frames for 1024-512 kbps

Sample/Method	Regular	Careful Blockmatching	Optical flow
Frog	86	105	116
Penguins	328	431	438
Ping-Pong	267	326	355
Atom-Bomb	441	377	386
Tsunami	563	375	380

Table 4: Number of coded frames for 1024-256 Kbps

Sample/Method	Regular	Careful block matching	Optical flow
Frog	43	67	71
Penguins	164	266	270
Ping-Pong	133	218	223
Atom-Bomb	232	238	240
Tsunami	352	223	235

Results of the careful block matching algorithm will always yield much more efficient output video when compared to direct brute force block matching approach which is not only time consuming, but also not necessary to have an efficient output.

By viewing the data from Table 3 and 4, we can figure out that, there is a wide difference in number of encoded frames between regular frame skipping mechanism and dynamic frame skipping mechanism (careful block matching or optical flow).

Additional analysis on the experimented outcomes, we can make a concrete point that number of skipped frames is inversely proportional to the bandwidth. From Fig. 5 and 6, we can conclude that number of coded frames will be very less in number in regular frame skipping which makes the video output that we produce to have jerky effect which reduces the readability of the output video.

The pictorial representations of the experimental results are provided in Fig. 4 and 5. The proposed algorithm, help in improving the transcoding process by reducing the frames through the usage of dynamic frame skipping algorithm.

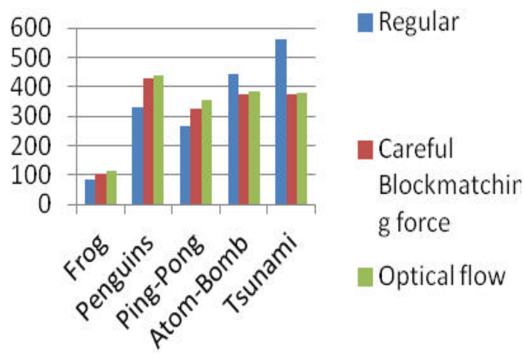

Fig. 5: Number of coded frames for bandwidth 1024-512 kbps

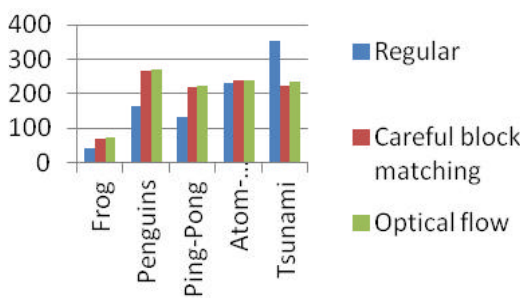

Fig 6: Number of coded frames for bandwidth 1024 to 256 kbps

CONCLUSION

The proposed Dynamic Frame Skipping and Offloading has improved the performance of transcoding video content suitable for the mobile devices with varied features and bandwidth requirements. It is evident that, the proposed algorithm adjusts the transcoding and transfer process as per the dynamic input criteria.

To calculate the transcoded video in much more effective manner, subjective assessment techniques can be applied .Our System has taken in to account the users vision and perspective in to account while transcoding. Future work can continue by including user's mood in the transcoding framework.

REFERENCES

Ahmad, I., X. Wei, Y. Sun and Y.Q. Zhang, 2005. Video transcoding: An overview of various techniques and research issues. IEEE Trans. Multimedia, 7: 793-804.

Biaz, S. and S. Wu, 2008. Rate adaptation algorithms for IEEE 802.11 networks: A survey and comparison. Proceedings of the IEEE Symposium on Computers and Communications, July 6-9, 2008, Marrakech, pp: 130-136.

Chen, G., B.T. Kang, M. Kandemir, N. Vijaykrishnan, M.J. Irwin and R. Chandramouli, 2004. Studying energy trade offs in offloading computation-compilation in java-enabled mobile devices. IEEE Trans. Parallel Distrib. Syst., 15: 795-809.

Cisco., 2016. Cisco visual networking index: Global mobile data traffic forecast update, 2015-2020. White Paper, Cisco Systems Inc., February 3, 2016, pp: 1-8.

Hsu, C.T., C.H. Yeh, C.Y. Chen and M.J. Chen, 2009. Arbitrary frame rate transcoding through temporal and spatial complexity. IEEE Trans. Broadcast., 55: 767-775.

In-Stat, 2011. More than half of US handset shipments will be smartphones by 2012: Worldwide smartphone shipments move toward 1 billion by 2015, Says in-stat. In-Stat, January 25, 2011. http://www.marketwired.com/press-release/more-than-half-us-handset-shipments-will-be-smartphones-2012-worldwide-smartphone-shipments-1384821.htm.

Kim, J.W., G.R. Kwon, N.H. Kim, A. Morales and S.J. Ko, 2006. Efficient video transcoding technique for QoS-based home gateway service. IEEE Trans. Consumer Elect., 52: 129-137.

Kim, N., J.Y. Yoo, N.L. Kim and J. Kim, 2012. A visual-sharing switching device supporting programmable in-network content adaptation. IEEE Trans. Consumer Elect., 58: 413-418.

Kim, S., H. Rim and H. Han, 2009. Distributed execution for resource-constrained mobile consumer devices. IEEE Trans. Consumer Elect., 55: 376-384.

Kumar, K. and Y.H. Lu, 2010. Cloud computing for mobile users: Can offloading computation save energy? Computer, 4: 51-56.

Ni, H.B. and S.X. Cao, 2012. Research and implementation of asynchronous video converter based on Linux. Applied Mech. Mater., 241: 2596-2600.

Rudenko, A., P. Reiher, G.J. Popek and G.H. Kuenning, 1998. Saving portable computer battery power through remote process execution. ACM SIGMOBILE Mobile Comput. Communi. Rev., 2: 19-26.

Shu, H. and L.P. Chau, 2004. An efficient arbitrary downsizing algorithm for video transcoding. IEEE Trans. Circuits Syst. Video Technol., 14: 887-891.

Shu, H. and L.P. Chau, 2007. Dynamic frame-skipping transcoding with motion information considered. IET Image Proces., 1: 335-342.

Xu, L., S. Kwong, H. Wang, Y. Zhang, D. Zhao and W. Gao, 2012. A universal rate control scheme for video transcoding. IEEE Trans. Circuits Syst. Video Technol., 22: 489-501.

Yeh, C.H., S.J.F. Jiang, C.Y. Lin and M.J. Chen, 2013. Temporal video transcoding based on frame complexity analysis for mobile video communication. IEEE Trans. Broadcast., 59: 38-46.