Influence of network awareness on perceived video quality.

Dmitri Jarnikov*, Peter van der Stok*‡, Johan Lukkien*

** Dept. of Math. and Computer Science, Eindhoven University of Technology P.O. Box 513, 5600 MB Eindhoven, the Netherlands {d.s.jarnikov, j.j.lukkien}@tue.nl*

‡ Philips Research Prof Holstlaan 4 5656 AA Eindhoven, the Netherlands peter.van.der.stok@philips.com

Abstract

In this paper, we present a strategy to control a decoder that processes scalable video on a networked terminal. The controller maximizes user perceived quality taking into account available computing resources and condition of the network.

1. Introduction

We consider the problem of video streaming via a wireless medium. The medium is characterized by frequent and unpredictable quality fluctuations. Moreover, resource constrained terminals are not always capable of processing all video data that is transmitted by a sender. Scalable video can be used to overcome these problems.

Scalable video partitions video data into a Base Layer (BL) and one or more Enhancement Layers (EL). The transmission and decoding of the BL is enough to reconstruct video of recognizable quality, while the ELs are needed only for an incremental improvement of the quality of the received video sequence.

The variations in the number of received layers, resulting from instabilities of the network, along with resource limitations of the terminal require continuous decisions on the number of layers to be processed. A dynamic control mechanism for the networked terminal is proposed that uses a strategy created offline by means of a Markov Decision Process (MDP) [1]. The assumption of the model in [1] is that indifferent of the number of layers received for the current frame the probability to receive x layers for the next frame is $1/N*100\%$ (where N is maximal number of layers). Changing network conditions is not taken into account. Thus, the chosen strategy is equal for bad and good state of the media.

In this paper we present a method to create controller strategies for various network conditions. The network aware solution (Solution 2) is compared to the unaware one proposed in [1] (Solution 1) and it is shown that the network awareness of solution 2 results in fewer quality level changes in the system.

2. Model

A scalable video decoder [3] enables trade-off between resource consumption and video quality. Decoding BL and 1 to i ELs corresponds to quality level i. A change in a quality level results in a change of the video quality and resource consumption. Frequent changes of picture quality are not appreciated by user. The control strategy proposed in [1] minimizes the number of quality level changes while maximizing the average quality level.

The control strategy is a result of solving a MDP, which is constructed as follows. Decoding a frame has a deadline. The deadlines for the successive frames are strictly periodic. In each period the decoder is guaranteed a fixed amount of processing time, called *budget*. After decoding a frame, we calculate the relative progress, defined as the fraction of budget left until the deadline of the next frame. There is an upper bound on the relative progress that defines the maximal number of frames that can be decoded in advance. When decoding of a frame is finished, a decision must be taken about the quality level of the next frame. The maximal quality level is given by the number of layers received for the next frame as it is not possible to decode more layers than is available.

The decision of the controller depends on the number of layers available for the next frame and the amount of terminal resources, expressed in a value of the relative progress. Thus, we need to estimate how many layers will be available for the frame that will come after the next one. Solution 1 assumes that with equal probability it could be any number of layers. The proposed network-aware Solution 2 calculates separate probabilities for different combinations of number and size of layers as well as for various network conditions.

The probabilities of receiving a particular number of layers for the next frame are calculated by a network simulation based on the Gilbert-Elliot channel model [2]. The simulation takes into account layers configuration and network conditions. The network conditions are expressed in bit error rate (BER) and burstiness (μ).

3. Evaluation

We used a CE device with a hardware decoder to receive and decode video. Parsing and summation of the video layers are done in software. Processing BL only takes, on average, 4.4 ms. Processing BL and EL or BL and two ELs takes 12 and 24.8 ms respectively. As the influence of the content of a video on the hardware decoder is negligible, the time needed for processing a particular number of layers has a low deviation from the average.

We have chosen an 802.11b network as streaming medium, as it is the most widespread standard for in-home network nowadays. The network was used exclusively for our streaming application. There was no significant interference on the network, so the chosen parameters for the channel conditions were $BER=10^{-6}$ and $\mu=0.3$. The useful throughput of such a channel is around 6Mbps.

For the MDP we used the following parameters. The upper bound on relative progress is set to 2, which assumes that we can process one frame in advance. The utility function, which defines the reward for being on the particular quality level, is set to 2, 4 and 8 for quality levels 0,1 and 2 respectively. We set the deadline penalty to 100000. The quality change penalties for increasing the quality level are set to 5 and 50 if the quality level is increased by 1 or 2 respectively. For decreasing the quality level, the penalties are set to 50 and 500 for going down by 1 or 2 levels respectively. Solution 1 assigns no penalties for decreasing quality level if the decrease is caused by network conditions. For Solution 2 penalties for decreasing quality level are the same for changes caused by controller as well as by network.

For the comparison we've chosen 70000 frames long scalable video consisting of three layers: a BL and two ELs. We considered three different configurations with respect to the sizes of the three layers. In the first configuration all layers are 1Mbps, they consume only half of the available bandwidth, so ideally the controller strategy should be defined only by terminal resource limitations. The second configuration, with all layers of 2Mbps, requires all the bandwidth, so the network condition is as important as the terminal resources. The third configuration (all layers are 3Mbps) needs more bandwidth than is available, so the network condition should play the most important role in the controller decisions. Probabilities of receiving exactly one, two or three layers are given in Table 1. These probabilities are calculated offline by the network simulation.

Configuration	BL only	$BL+EL_1$	$BL+EL_1+EL_2$
1	0	0	100
2	0.08	51.36	48.56
3	71.29	28.71	0

Table 1 Probability (%) of having the given layers

We made pairwise comparison of the solutions, looking at the average quality level and the number of quality level changes. The comparison was made for all three above-mentions configurations. We considered budgets from 4 to 40 ms, with step of 1 ms.

For configuration 1, both solutions behave in the same way, delivering equal quality and making almost the same amount of quality changes. The reason is that in this configuration all three layers are constantly available for processing. Consequently, both solutions take into account only terminal resources (which are equal) resulting in nearly the same strategies.

The behavior of controllers for configuration 2 differs significantly (Figure 1,Figure 2). Starting from a budget of 12 ms, which allows successful decoding of two layers, the controller based on the strategy of Solution 2 does not attempt to increase the quality level. The reason is that according to the network conditions (see Table 1) the second EL is not available in, roughly, half of the cases. Thus, choosing quality level 2 will lead to frequent quality level changes. On the other side, the Solution 1 does not take network changes into account, resulting in higher average quality for the price of extremely high quality fluctuations.

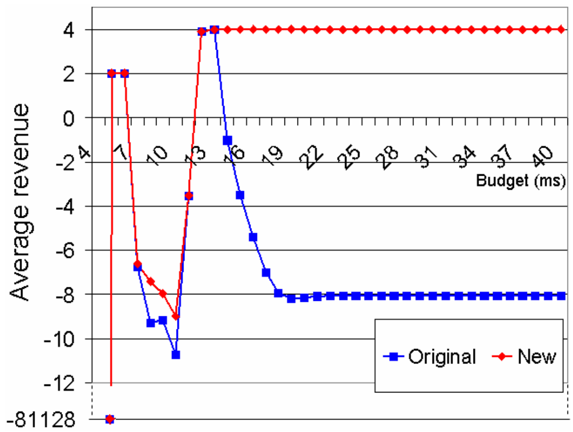

Figure 1 Average revenue for Configuration 2

The reason that the average revenue of both solutions is extremely low for the budget of 4 ms is a large number of deadline misses. The deadline misses occur because the budget given to the decoder is significantly lower than the average processing time for BL (4.4 ms). Starting with a budget of 5 ms, the number of deadline misses is 0, as there is always enough time to decode at least BL.

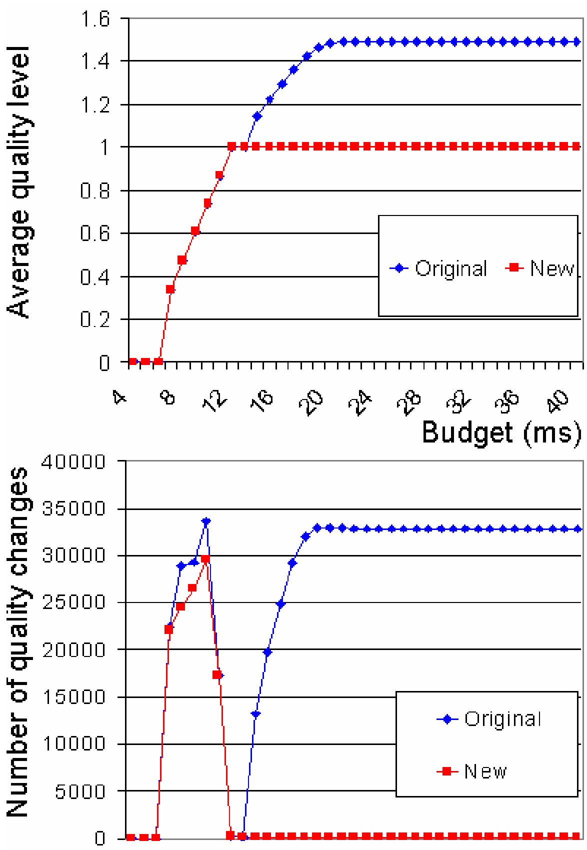

Figure 2 Comparison of average quality level and quality level changes for Configuration 2

The results for configuration 3 are again the same. Since receiving BL and two ELs is not probable, the controller is left with the choice between processing one or two layers. However, in view that most of the time (71.29%) only BL is available, both controllers behave conservatively, trying not to choose quality level 1. Thus, the strategy of the controllers is fully defined by the network conditions, as there are enough resources to decode BL for budgets higher than 4.4 ms (average decoding time for BL).

4. Conclusions

In this paper we presented improvements to the creation of the controlling strategy for scalable video decoder, presented in [1]. We've shown that the knowledge of the network conditions allows decrease of quality level changes, thus delivering a video with higher objective quality.

5. References

[1] Jarnikov, D.; van der Stok, P.; Wust, C.C., "Predictive Control of Video Quality under Fluctuating Bandwidth Conditions". Multimedia and Expo, 2004. ICME '04. 2004 IEEE International Conference on, Volume: 2 , pp. 1051 – 1054, June 27-30, 2004

[2] J. R. Yee and J. Edward J. Weldon, "Evaluation of the performance of error-correcting codes on a Gilbert channel", IEEE Trans. on Communications, pp. 2316-2323, Aug. 1995

[3] C. Hentschel, R.J. Bril, M. Gabrani, L. Steffens, K. van Zon, and S. van Loo, Scalable video algorithms and dynamic resource management for consumer terminals, In Proc. International Conference on Media Futures (ICMF), pp. 193–196, Florence, Italy, May 2001