

MPEG-4 VIDEO TRANSCODER FOR MOBILE MULTIMEDIA TRAFFIC PLANNING

S. Dogan, A.H. Sadka and A.M. Kondo

Centre for Communication Systems Research (CCSR), University of Surrey, UK

ABSTRACT

Traffic planning for multimedia has become more and more significant as new network infrastructures with different characteristics have been introduced lately. The increasing popularity of mobile communications has also drawn much of the attention from the research point of view. Thus, this paper addresses the problem of traffic planning, initiated by this interest, for mobile video communications and proposes a video transcoder bank to resolve congestion and/or bandwidth limitation. The proposed architecture presents a layered structure of multiple video rates as required by various networks. Moreover, the paper also introduces an adaptive method for resolving congestion. The designed system monitors the congestion with a feedback loop within a network and adaptively produces necessary transmission rates whilst providing the best available service quality.

INTRODUCTION

Video transcoders that are particularly designed for video-on-demand (VoD) applications have been exhaustively presented in the literature for the past few years [1, 2]. These video transcoders have made use of the MPEG-2 digital video broadcasting (DVB) standard [3], bearing in mind that bit rate conversion of the previously encoded and stored VoD data might be needed for certain end-users. For the purposes of multimedia traffic planning, these video transcoders act as bit rate regulators in order that the congested or bandwidth-limited network(s) can accommodate the pre-encoded video stream [4].

The referred planning problem of the multimedia traffic for congested nodes or bandwidth-limited networks will also exist in the future as it does now since the number and diversity of users and their needs vary in a wide range. However, the scope of the existing problem has very much proceeded lately from VoD and broadcast applications towards two-way mobile video communications. The third generation (3G) mobile communications network infrastructure is promising to give support to video transmissions as well as voice and data communications. Hence, video transcoding algorithms have also become of importance as they provide the necessary bit rate conversions between two or more communicating parties.

Thus, in the light of these facts, this paper proposes a video transcoding algorithm which provides reasonable

quality of service (QoS) to each of the diverse bandwidth-constrained networks for a successful utilisation of the two-way video-telephony/conferencing application. The transcoder simultaneously produces varying outputs at different bit rates for different networks according to their bandwidth requirements as seen in Figure 1. The figure depicts a possible real-life scenario where a high quality, high bit rate (R_0) video data is converted to lower bit rates (R_1, R_2) via the video transcoder due to bandwidth and/or congestion problems in these networks. Here, the major achievement of the transcoder is to provide bit rate regulation whilst maintaining the QoS as much as possible. In the case that the two networks at either ends of the transcoder can support same or higher bit rates (R_3, R_4), then the video transcoder maintains this original rate, and hence, the original video quality. The video transcoder achieves its objective with the least possible delay and complexity for optimum simplicity. The proposed transcoder architecture carries out bi-directional operation since it is a two-way gateway structure to video-telephony/conferencing applications.

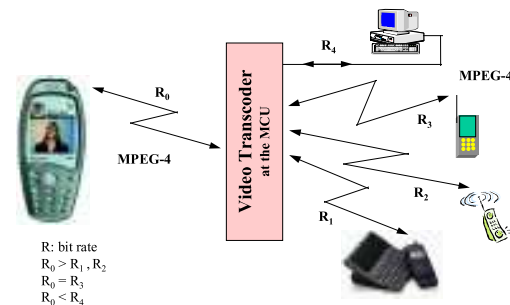


Figure 1: MPEG-4 multi-party video-telephony/conferencing scenario

A video transcoder of this kind allows the users of mobile communications to set up a connection regardless of their network and connection statistics whilst letting service providers employ much less complex encoder/decoder sets as opposed to complex layered scalable video transmission equipment [5]. The proposed video transcoder supports MPEG-4 [6] video transmission as MPEG-4 is believed to be the ubiquitous video compression standard for the next generation of mobile-wireless communications.

In the next section, brief information is given on video transcoding whilst the third section discusses the system architecture for the proposed MPEG-4 mobile multimedia traffic planning type of video transcoder. Furthermore, concluding remarks are provided

following the discussions of the results obtained from the simulations of the proposed traffic planner.

VIDEO TRANSCODING

Video transcoding is performed to convert one compressed video format to another. This conversion takes place at the multi-point control unit (MCU) of the multimedia network providing the necessary links between networks with different characteristics and constraints. With the use of a video transcoder, diverse video compression standards can be linked to each other [7] or incoming compressed video streams can be re-sized [1, 2]. Thus, a higher bit rate can be customised for a congested or bandwidth-limited network.

Re-sizing of a compressed video stream comprises partial decoding of the bitstream, re-quantisation of its transform coefficients and re-encoding. Re-quantisation is carried out using a different quantisation parameter from the incoming original one. A finer parameter than the original results in an increase in the outgoing bit rate whilst a coarser parameter has an opposite effect. Hence, in the case that a network node cannot support higher rates due to congestion problems, the video transcoder elaborates the stream coarsely to decrease the original rate whilst maintaining a certain QoS level.

However, direct re-quantisation introduces continuous picture drift in transcoded video streams. Drift is caused by the mismatch of reconstructed images during the transcoding as the transcoded pictures are re-quantised with a different parameter. Therefore, it is common sense to include a drift correction loop which adds the drift error back into the system removing the mismatch problems [1, 2]. However, often a correction loop is not enough solely for a better QoS, and a further motion vector (MV) refinement process also has to be integrated in the system [8]. MV refinement is a process whereby incoming non-optimal MVs are refined within a small search window around these original non-optimal MVs. This happens due to the fact that re-quantisation results in non-optimal MVs that were originally encoded using a different quantiser.

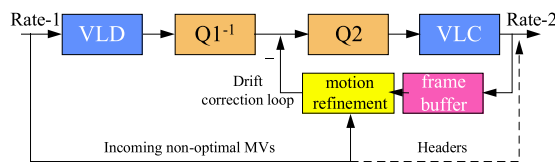


Figure 2: Typical video transcoder block diagram

A typical video transcoder is illustrated in Figure 2. The upper direct path without the feedback loop achieves the direct re-quantisation for rate conversions. The feedback loop (the lower path) corrects the drift error introduced by the upper direct path as stated above. Finally, the motion refinement block within this particular loop refines the original MVs to select the optimal ones for the new quantiser.

The Drift Effect and Its Correction

Figure 3 depicts the effects of drift and its correction together with the feedback and motion refinement. The simulated results were obtained for a 300-frame standard video test sequence, namely “Akiyo-with-Crowd”, originally encoded at 368 kb/s on average and 25 fr/s, giving a peak signal-to-noise ratio (PSNR) of 37.279 dB on average.

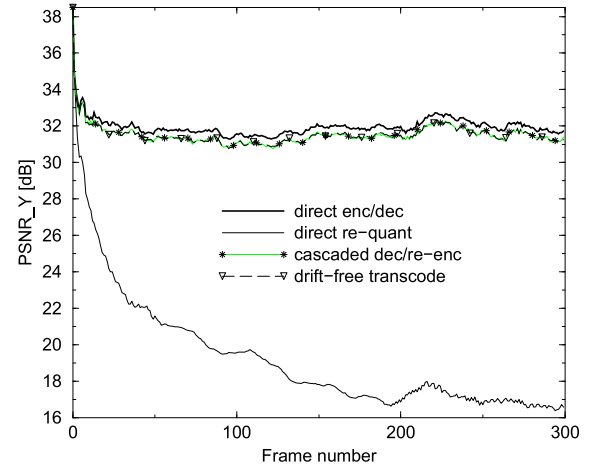


Figure 3: PSNR of transcoded “Akiyo-with-Crowd”, originally encoded at 368 kb/s on average and 25 fr/s

<i>Sequence type</i>	<i>av. Bit rate [kb/s]</i>	<i>av. PSNR Y [dB]</i>
direct enc/dec	128.75	31.956
direct re-quant	87.05	19.272
cascad dec/re-enc	125.48	31.514
drift-free trnsd	125.57	31.523

Table 1: Average bit rates and luminance PSNR values for the sequences in Figure 3

The bold line in the figure represents the quality level of the video sequence if it is directly encoded and decoded at 128 kb/s. This pattern has to be taken as a reference to other transcoding results. One extra decoding and re-encoding cycle degrades the quality by 0.442 dB on average as observed from the dotted line. Meanwhile, the dashed line for drift correction with MV refinement case presents 0.01 dB on average better quality than the cascaded fully decoding/re-encoding sequence with the dotted line. Here, the refinement search window size was set to ± 5 pixels. On the other hand, the performance of feedback with motion refinement is also significantly better than direct re-quantisation with straight line as picture drift continuously degrades the QoS. The difference is 12 dB on average.

Table 1 presents the bit rates obtained from different conditions of transcoded “Akiyo-with-Crowd” of Figure 3. As expected, the lowest bit rate was achieved for direct re-quantisation method since it comprised the simplest algorithm with the least number of added bits to the stream. On the contrary, the other results showed slightly increased bit rates as they included drift error

correction and motion re-evaluation. Table 1 also shows the detailed average PSNR values of Figure 3.

The Significance of MV Search Window Size on Refinement Quality

The refinement is accomplished within a MV search window around the blocks that are pointed to by the incoming original MVs. These MVs are known as non-optimal ones as they refer to incoming original rate. During the conversion of video rate, the original MVs point at wrong blocks within a video frame due to the mismatch caused by different quantiser levels. Hence, original MVs are needed to be refined in a small search window around them.

The MV search window size has conventionally been chosen as small due to prevention of high complexity factor of the video transcoder [8]. Indeed, small window sizes give reasonable quality levels during the refinement of incoming non-optimal MVs achieving lower complexity.

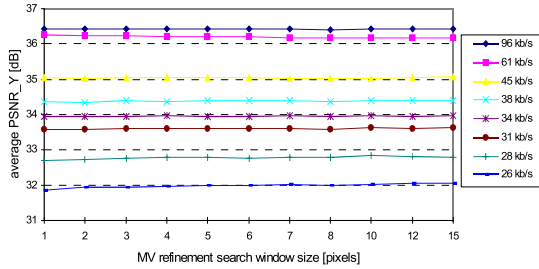


Figure 4: Average PSNRs of transcoded “Suzie”, originally encoded at 64 kb/s on average and 25 fr/s, versus various refinement window sizes

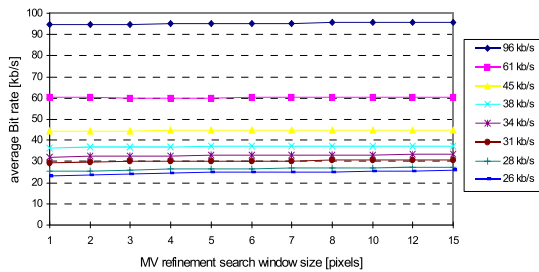


Figure 5: Average bit rates of Figure 4 versus various refinement window sizes

Enlarging the refining window size, meanwhile, gives a small degree improvement in quality. This is indeed an expected result since larger window size allows better block matching. However, whilst increasing search range gives better performance, it also results in an increase in already reduced bit rate at the transcoder. Therefore, an optimum point is necessary to be found for reduced bit rate and complexity with improved quality. Figure 4 demonstrates the average PSNR levels versus different refinement window sizes whilst Figure 5 presents the results of average bit rate variations for corresponding cases of Figure 4.

Figure 4 shows that increasing refinement search range also increases the PSNR slightly, but at the cost of increased bit rate as observed from Figure 5. Yet only the increase in PSNR is noticeable particularly when there is high asymmetry in bit rate difference, i.e., from very high rates to very low rates (from 64 kb/s to 28/26 kb/s cases in the figure). Conversely, when there is not much difference, neither the PSNR nor the bit rate changes in significant amounts.

SYSTEM ARCHITECTURE

In this paper, the video transcoder has been exploited to regulate the transmission bit rates for mobile multimedia traffic planning. For this purpose, a bank of video transcoders within the MCU has been designed for the proposed aim. The bank of video transcoders, as illustrated in Figure 6, comprises a series of transcoders combined with each other in parallel, each of which carries out similar operations for the provision of different layers of a video stream. Different layers have different bit rates and quality levels to suit all various conditions of destination networks. If a network has a congested node or if any user within a network cannot handle the incoming rate due to bandwidth constraints, such as a mobile-wireless network, the transcoder bank reduces the original rate. For networks that can support higher rates or at least the incoming rate, the transcoder bank can also maintain the incoming rate with the original quality without the need for transcoding the input video data, as shown in Figure 6 with the direct path.

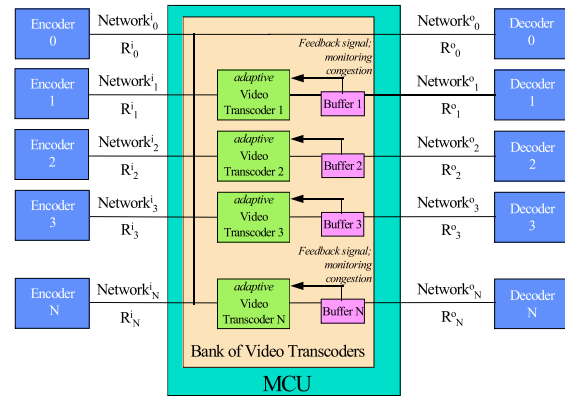


Figure 6: System architecture; the video transcoder bank (superscripts i and o representing inputs and outputs, respectively, and subscripts representing network types)

It has also been considered to employ only one video transcoder producing simultaneous multi-rate outputs, for a typical point-to-multi-point video-conferencing scenario. However, this has not been implemented in this paper due to its high complexity requirement with long processing delays, which cannot be tolerated by two-way video transmissions. As an alternative, a design of more than one video transcoder operating in parallel, as in Figure 6, is thought to be a more efficient and less

delay-prone implementation as the processing delay will only be one transcoding cycle.

With the use of a video transcoder bank, separate layers of individual compressed video data copies can be produced for a particular point-to-multi-point operation. In the case that more inputs are considered, as in a multi-point-to-multi-point video transmission scenario, then the transcoder bank is made capable of handling more than one input at a time. This is more like a time-shared networking issue between different inputs, and hence, it should be considered by the MCU itself.

Even though it seems like certain transcoders are dedicated to fixed output bit rates from Figure 6, a further flexible and adaptive operation has also been introduced to the system for time varying bandwidth requirements. For instance, if network 2 is congested at a time as stated earlier, and if the congestion is resolved before the video transmission ceases, then there is no need to continue to supply the lower rate to that particular network. In such a case, network congestion feedback can tell the associated video transcoder to increase the rate, hence also increasing the QoS. Conversely, if the congestion conditions persist, the feedback report is signalled to the transcoder to reduce the output rate further. Thus, each individual video transcoder within the bank operates in an adaptive way to the new conditions and relays the most suited output with the new rate. This feedback signalling, which reports the varying network congestion conditions, can simply be incorporated using a buffering mechanism. The buffers at the end of each transcoder estimate the congestion characteristics of the host networks by using the occupancy factor. Thus, according to the buffer occupancy, the feedback signal reports back to the particular video transcoder to increase or decrease the output rate.

Moreover, as the original rate output (upper path in the figure) can be sent without any transcoding operation and its associated processing delay, the synchronisation of this particular stream has to be maintained with the other layers by the MCU. Thus, the MCU is required to accomplish a synchronisation process between different outputs of the transcoder bank.

COMPUTER SIMULATIONS

Simulations were performed using the “Foreman” video test sequence with 200 frames. “Foreman” was originally encoded at 116 kb/s on average with a frame rate of 25 fr/s in I-P-P-P-... format. The sequence was encoded, transcoded and decoded in compliance with the MPEG-4 standard with the use of unrestricted MVs and advance prediction modes. The frame resolution was chosen as quarter common intermediate format (QCIF) with 176×144 pixels. The results given in the next section comprise luminance (Y) PSNR and bit rate variations versus the number of transcoded frames.

RESULTS AND DISCUSSIONS

Figure 8 demonstrates the layered quality structure of the video transcoder bank outputs at different bit rates. As clearly observed from the figure, as the output rate decreases, so does the video quality. However, it is also notable that for an approximately 46 kb/s decrease from R_1 to R_5 , the transcoded video quality only degrades for about 2.5 dB on average, which is fairly imperceptible to the human eye. The following Table 2 presents the detailed bit rate and PSNR results for individual rates. As proved one more time with this table, higher rate reduction asymmetry during transcoding requires larger MV refinement window sizes for better QoS provision.

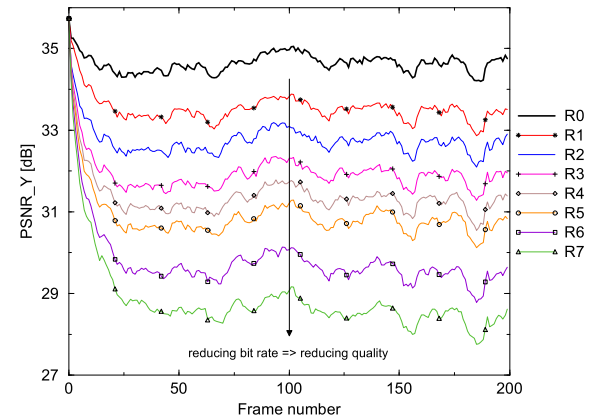


Figure 8: PSNRs of layered transcoder bank outputs

Network type	av. Bit rate [kb/s]	av. PSNR Y [dB]	MV refine. win. [pels]
R_0	116.063	34.642	N/A
R_1	99.208	33.531	± 1
R_2	82.088	32.755	± 1
R_3	66.038	31.947	± 1
R_4	57.803	31.380	± 3
R_5	52.778	30.891	± 4
R_6	45.218	29.792	± 6
R_7	41.655	28.827	± 7

Table 2: Average bit rate and PSNR values for Figure 8

Finally, Figures 9 and 10 present the bit rate and PSNR results of the proposed adaptive congestion control technique, respectively. As noticed from the figures, channel congestion characteristics varied throughout the video transmission. The particular channel already had congestion. Thus, the video transcoder bank produced a reduced rate, $R_2=82$ kb/s on av., from the input rate, $R_0=116$ kb/s on av. However, due to the additional congestion, the supported rate within the network reduced to 66 and 58 kb/s on av., respectively. This is evident from the figures for periods of the 50th-100th and the 150th-175th video frames. For these particular cases, video transcoder produced reduced rates as required by the congestion bottleneck yet at the cost of quality degradation. On the other hand, between the 110th and 125th frames, even the already existing congestion

condition was resolved for a short time. Hence, the adaptive transcoder bank increased the bit rate providing the best achievable quality for this short time period. As clear from the figures, the transcoder bank follows the variation of congestion conditions adaptively with the help of feedback signalling.

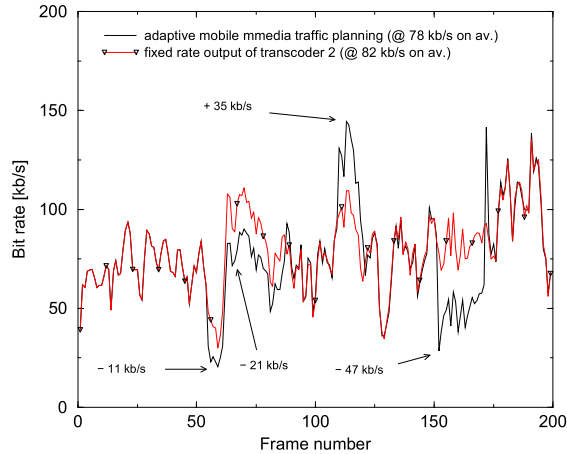


Figure 9: Rate variation for adaptive congestion control

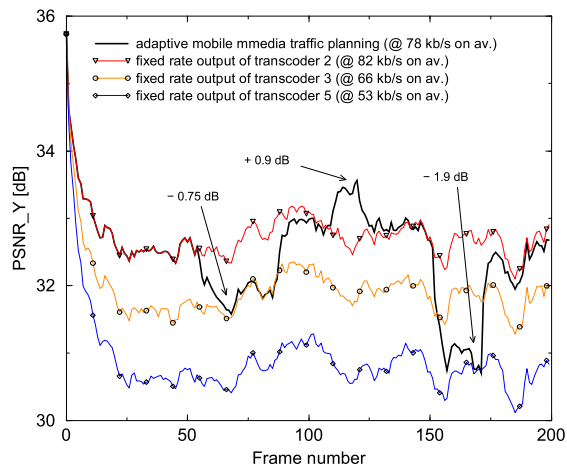


Figure 10: PSNR variation for adaptive scheme

Instantaneous bit rate and PSNR changes from the default values are given within the figures, (-) sign representing the decrease and (+) sign showing the increase from the fixed rate levels. PSNR figure also presents the results for certain fixed rate outputs of the transcoder bank, which fit in the varying congestion requirements, for comparison purposes. The proposed adaptive scheme gave better overall PSNR (32.483 dB at 78 kb/s on av.) showing that the video quality degraded only for a short time with this technique until the congestion was resolved. On the contrary, fixed low rates exhibited continuous low quality measures.

CONCLUSIONS

In this paper, it has been demonstrated that an MPEG-4 video transcoder bank can accomplish the necessary bit

rate regulations for mobile multimedia traffic planning. This transcoding operation has been achieved with the best available QoS and minimal complexity. The results yielded necessary bit rates for demanding networks at reasonable video qualities. Moreover, an adaptive scheme to monitor and control the varying congestion conditions with a feedback loop has also been employed. This kind of achievement has given better overall service quality as it exploited the changing conditions and available bandwidth more efficiently than a fixed rate output.

Furthermore, preliminary results showed that the significant QoS degrading drift effect could be overcome by a drift correction loop and MV refinement algorithm. Video transcoding with this method has presented an average of 12 dB better quality than the drift-prone algorithm. Meanwhile, MV refinement search window size increase resulted in slightly improved performance yet at the expense of higher bit rate and complexity.

REFERENCES

1. Keesman G., Hellinghuizen R., Hoeksema F. and Heideman G., 1996, "Transcoding of MPEG bitstreams", *Sig. Proc.: Image Comm.*, **8**, (6), 481-500
2. Kan K.S. and Fan K.C., 1998, "Video transcoding architecture with minimum buffer requirement for compressed MPEG-2 bitstream", *Sig. Proc.*, **67**, 223-235
3. ISO/IEC CD 13818-2, Nov. 1993, "Generic coding of moving pictures and associated audio"
4. Assuncao P.A.A. and Ghanbari M., June 8-12 1997, "Congestion control of video traffic with transcoders", *in ICC'97 conf.*, **1**, 523-527, Montreal, Canada
5. Ghanbari M., 1989, "Two-layer coding of video signals for VBR networks", *IEEE J-SAC*, **7**, (5), 771-781
6. ISO/IEC JTC1/SC29/WG11 N2802, July 1999, "Information technology - Generic coding of audio-visual objects - Part 2: Visual, ISO/IEC 14496-2", MPEG Vancouver meeting
7. Dogan S., Sadka A.H. and Kondo A.M., 1999, "Efficient MPEG-4/H.263 video transcoder for interoperability of heterogeneous multimedia networks", *IEE Electron. Lett.*, **35**, (11), 863-864
8. Youn J., Sun M.T. and Lin C.W., 1998, "Motion estimation for high performance transcoding", *IEEE Trans. on Cons. Electron.*, **44**, (3), 649-657