

Optimal H.264/AVC Video Transcoding System

Fatma Al-Abri⁽¹⁾, Eran Edirisingh⁽¹⁾, Jan De Cock⁽²⁾, Stijn Notebaert⁽²⁾ and Rik Van de Walle⁽²⁾

⁽¹⁾ Digital Imaging Research Group, Department of Computer Science, Loughborough University, UK

⁽²⁾ IBBT, Department of Electronics and Information Systems, Multimedia Lab, Ghent University, Belgium

Abstract — This paper presents an efficient receiver-aware video transcoding system that systematically chooses the optimal transcoding operation from multiple options while meeting network and user constraints. Multi-objective optimization is used to select the best **transcoding method** that minimizes transcoding complexity and memory usage while ensuring the client constraint of bitrate and requested quality are fulfilled.

Index Terms — Joint Optimization, H.264 AVC, Transcoding, Transrating.

I. INTRODUCTION

Most existing multimedia systems provide services to various consumer electronic devices (PCs, laptops, PDAs, smart phones, etc.) interconnected via heterogeneous wired and wireless networks. Adaptation of the properties of multimedia content, such as video streams in digital video services, to constraints imposed by the network and the client is essential. Transcoding is one of the solutions used to adapt a video stream by reducing the temporal resolution, lowering the spatial resolution, decreasing the visual quality, or changing the coding format [1]. The optimal selection of the transcoding operation that reduces the complexity of the transcoder while maintaining high quality output within the target bit-rate is one of the challenges in video transcoding research. An optimal adaptation framework for streaming multiple video objects was presented in [2] which considered the selection of the most appropriate versions for the video objects and the transcoding method for the video objects given the constraints of both network bandwidth and the cost of transcoding. A general framework, called utility-based video adaptation was proposed in [3]. This approach is a simple extension of conventional R-D framework that allowed the incorporation of diverse types of resources (e.g., complexity and bandwidth) and adaptation operations. In this paper, we propose an optimized video transcoding system where the transcoding methods, for different types of macroblocks in the encoded bitstream, are optimally selected according to the available system resources in terms of computational power and memory space in the transcoder. In addition to that, the client requirements of minimum video quality and supported bitrate are satisfied.

II. SYSTEM OVERVIEW

In the context of this system as shown in Figure 1, a separate H.264 AVC encoding entity is assumed and the transcoder retains the coding decisions of the incoming video stream. The received bitstream from the H.264 AVC encoder is parsed and relevant coding information is extracted from it. This information is necessary to estimate the total system

memory and total computational power required to transcode the bitstream. The estimated values are compared to the current available system resources in the transcoding server. The decision of which transcoding method to use is reached by evaluating the required transcoding resources, available resources, the output bit-rate and PSNR value for the transcoding options.

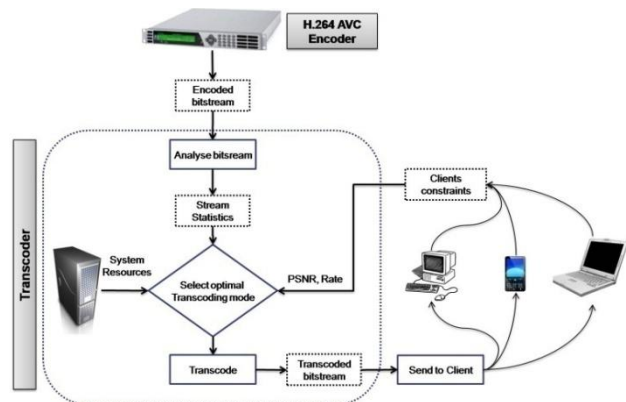


Figure 1: Framework for the Optimal Transcoding System

A. Transcoding Options

The set of transcoding options is based on the mixed architectures for H.264/AVC transrating presented on [4] and [5]. These architectures combine different transrating techniques, which are applied depending on the picture/macroblock type. The mixed combinations are based on three well-known transcoding methods; Open Loop (OL) transrater, Fast-Pixel domain (FP) transrater and Cascaded Pixel-domain (CP) transrater. Detailed explanations of the three methods in the context of H.264 AVC is presented in [4]. Table 1 shows a list of different combinations of the applied transcoding methods for each block and frame type.

Table 1: List of Mixed Transcoding modes

Transcoding Scheme	I	P/I	P/P	B/I	B/B
0	OL	OL	OL	OL	OL
1	-	OL	OL	OL	OL
2	FP	FP	FP	FP	FP
3	FP	OL	OL	OL	OL
4	CL	CL	CL	CL	CL
5	FP	CL	OL	CL	OL
6	FP	CL	CL	CL	OL
7	FP	CL	CL	CL	CL
8	FP	-	CL	-	CL
9	FP	OL	CL	OL	CL
10	FP	-	OL	-	OL
11	-	-	OL	-	OL

III. Optimal Transcoding

The decision on which transcoding mode to use is based on satisfying the following objectives:

- Minimum use of memory in the transcoder
- Minimum use of computational complexity
- Maximum PSNR within client range
- Maximum bitrate within client's network capabilities.

The computational power is estimated by analyzing the total number of mathematical, logical and comparison operations required by the transcoding scheme implementation. The memory usage is also estimated in the same manner but by counting the total number of memory accesses required by the transcoder. The memory and complexity are added up for the whole video sequence from the different combinations of transcoding schemes for different blocks in different types of frames. The PSNR and bitrate are calculated as described in the H.264 AVC standard [6].

IV. EXPERIMENTAL RESULTS

The optimal transcoder was tested on a Pentium 4 PC clocked at 3.19 GHz with 1 GB RAM. The results presented in this section are for 100 frames of the Foreman QCIF (30fps) video sequence that was initially encoding using the JM 15.1 H.264 AVC reference software using an intra period of 15, 2 reference frames and IBBP GOP with I frame the quantisation parameter QP=22, P frame QP=23 and B frame QP=24. In the transcoder the re-quantisation is done with $\Delta QP=4$ for all frame types.

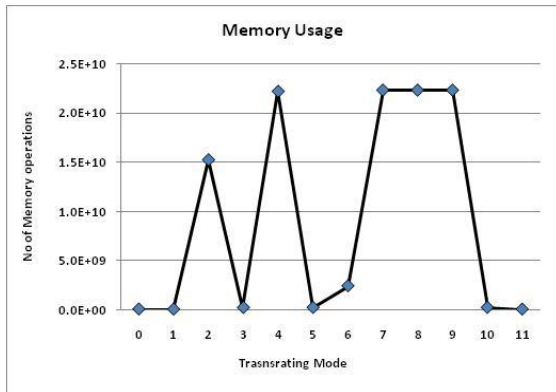


Figure 2: Memory usage in the for each transcoding mode

Figure 2 shows the changes in memory usage that correspond to each transcoding mode. As seen in the graph, mode 4, 7, 8 and 9 are the highest in memory access operations. This is expected as these modes use the cascade pixel transcoder for some or all the frames, which include the full decode-encode process. However, mode 2 in Figure 3 was the most complex in terms of computation operations due to the intensive use of motion compensation for all frames.

A summary of the results of all modes for all objectives for the Foreman QCIF sequence is given in Table 2. The transcoder will choose a mode depending on the available system resources and on the client's requirement for quality in terms of PSNR and bandwidth in terms of bitrate. For example if a client requested a video and it indicated that the network

bandwidth it is connected through does not exceed 200 kbps, mode 2, 4, 6, 7, 8 and 9 will satisfy the condition. Selecting the modes of the best PSNR will exclude mode 4 from the options and further refining the selection by minimizing the complexity and memory will result in mode 6 being the optimal mode for transcoding.

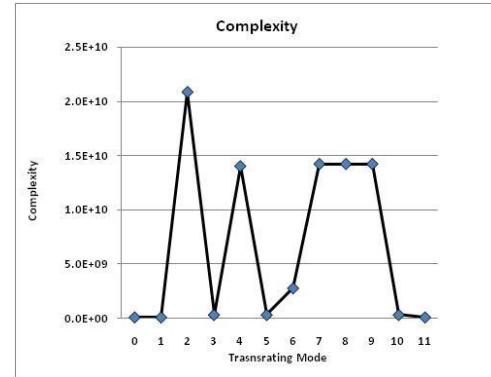


Figure 3: The complexity changes associated with the selection of different transcoding options.

Table 2: Results for Foreman QCIF with $\Delta QP=4$

Encoding scheme	PSNR (db)	Bit rate kbps	Complexity CUnit	Memory MUnit
0	27.66	217.49	5.669E+07	9.482E+07
1	41.98	246.19	5.170E+07	8.562E+07
2	39.96	193.21	2.078E+10	1.525E+10
3	39.08	200.79	3.495E+08	3.114E+08
4	35.72	188.61	1.450E+10	2.387E+10
5	39.12	200.58	3.498E+08	3.121E+08
6	39.04	192.00	1.889E+09	1.682E+09
7	39.00	189.80	1.477E+10	2.403E+10
8	39.03	190.47	1.477E+10	2.403E+10
9	38.92	189.87	1.477E+10	2.403E+10
10	39.16	201.41	3.494E+08	3.113E+08
11	42.23	246.80	5.164E+07	8.551E+07

V. Conclusion

We proposed a jointly optimized H.264 transcoder where the transcoding mode for a video sequence is optimally selected from a set of options depending on the available resources in the transcoder and receiver requirements on quality and bitrate. Such a system will have wide applicability in the area of consumer electronics where consumer devices with different capabilities co-exist within heterogeneous networks having multitude of practical constraints.

REFERENCES

- I. Ahmad, X. Wei, Y. Sun, and Y.-Q. Zhang, "Video transcoding: an overview of various techniques and research issues," *IEEE Trans. Multimedia*, vol. 7, no. 5, pp. 793-804, Oct. 2005.
- Jeongyeon Lim, Munchurl Kim, "An Optimal Adaptation Framework for Streaming Multiple Video Objects", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 18, no 5, 699-703, May 2008
- Y.Wang, J.-G. Kim, S.-F. Chang, and H.-M. Kim, "Utility-based video adaptation for universal media access (UMA) and content-based utility function prediction for real-time video transcoding," *IEEE Trans. Multimedia*, vol. 13, pp. 213-220, Sep. 2007.
- Notebaert S, De Cock J, Beheydt S, De Lameilleure J, Van de Walle R (2009) Mixed architectures for H.264/AVC digital video transcoding. *Multimedia Tools and Application*.
- De Cock J, Notebaert S, Lambert P, Van de Walle R (2010) Requantization transcoding for H.264/AVC video coding. *Signal Process Image Commun* 25(4):235-254.
- Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard, and Ajay Luthra, "Overview of the H.264/AVC Video Coding Standard", in *IEEE Transactions on Circuits and Systems for Video Technology*, 2003