QUANTIZER DESIGN FOR EXPLOITING COMMON INFORMATION IN LAYERED CODING

Mehdi Salehifar, Tejaswi Nanjundaswamy, and Kenneth Rose

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA, 93106
E-mail:{salehifar, tejaswi, rose}@ece.ucsb.edu

ABSTRACT

Today's diverse data consumption devices and heterogeneous network conditions require content to be coded at different quality levels. Conventional scalable coding, which generates hierarchical layers that refine quality incrementally, introduces a performance penalty, as most sources are not successively refinable at finite delays for the distortion measure employed and the combination of rates at each layer. On the other hand encoding different copies at required quality levels is clearly wasteful in resources. We previously proposed a common information based framework with a relaxed hierarchical structure to generate common and individuals bit-streams for different quality levels, to provide the flexibility of operating at points between conventional scalable coding and independent coding. In this paper we propose a quantizer design technique for this layered coding framework, which enables extracting common information between two quality levels with negligible performance penalty. Experimental results for Laplacian sources, which are prevalent in practical multimedia systems, substantiate the effectiveness of our proposed technique.

***Index Terms*—** Laplacian Sources, Dead-Zone Quantizer, Scalable Coding, Common Information.

1. INTRODUCTION

Technological advances ranging from multigigabit high-speed Internet to wireless communication and mobile, limited resource receivers, have created an extremely heterogeneous network scenario with data consumption devices of highly diverse decoding and display capabilities, all accessing the same content over networks of time varying bandwidth and latency. The primary challenge is to maintain optimal signal quality for a wide variety of users, while ensuring efficient use of resources for storage and transmission across the network. The simplest solution to address this challenge is storing and transmitting independent copies of the signal for every type of user the provider serves. This solution is highly wasteful in resources and results in extremely poor scalability. In an alternative solution, conventional scalable coding [1, 2] generates layered bit-streams, wherein a base layer provides a coarse quality reconstruction and successive layers refine the quality, incrementally. Depending on the network, channel and user constraints, a suitable number of layers is transmitted and decoded, yielding a prescribed quality level. However, it is widely recognized that there is an inherent loss due to the scalable coding structure, with significantly worse distortion compared to independent (non-scalable) encoding at given receive rates [3, 4, 5], as most sources are not successively refinable at finite delays for the distortion measure employed and the combination of rates at each layer. Moreover for fixed receive rates, non-scalable coding and conventional scalable coding have the highest and the lowest total transmit rate, respectively. Thus, non-scalable coding and conventional scalable coding represent two extreme points in the tradeoff between total transmit rate and distortions at the decoders, with fixed receive rates.

In our previous work we proposed a novel layered coding paradigm for multiple quality levels [6] inspired by the information theoretic concept of common information of dependent random variables [7, 8, 9], wherein only a (properly selected) subset of the information at a lower quality level is shared with the higher quality level. This flexibility enables efficiently extracting common information between quality levels and achieve intermediate operating points in the tradeoff between total transmit rate and distortions at the decoders, in effect controlling the layered coding penalty. Our early paper [6] established the information theoretic foundations for this framework and a later paper [10] employed this framework within a standard audio coder to demonstrate its potential. In this paper we tackle the important problem of designing quantizers for this layered coding framework, specifically for two quality levels with fixed receive rates. We need to design three quantizers, one for the common layer, whose output is sent to both the decoders, and two other quantizers refining the common layer information at two quality levels, whose output is sent individually to the two decoders. We first propose to employ an optimal quantizer for a given rate at the common layer. Given this quantizer, we design two other optimal quantizers at two different required rates, conditioned on each common layer interval. Finally the optimal common layer rate is estimated numerically by trying multiple allowed common rates and selecting the highest one amongst those with negligible loss in distortion compared to non-scalable coding. We then adapt this technique to the practically important Laplacian sources.

The rest of the paper is organized as follows: In Part 2, Laplacian sources and their optimal quantizers are discussed. In Part 3, common information based layered coding paradigm is introduced. In Part 4, quantizer design for layered coding of a general source and then specifically a Laplacian source is described. In Part 5, experimental results substantiating the proposed technique are presented. Finally, we conclude in Part 6.

2. LAPLACIAN SOURCES

In many practical applications, multimedia sources are modeled by the Laplacian distribution,

$$f_L(x) = \frac{\lambda}{2} e^{-\lambda|x|},$$

The work was supported in part by the National Science Foundation under grant CCF-1320599.

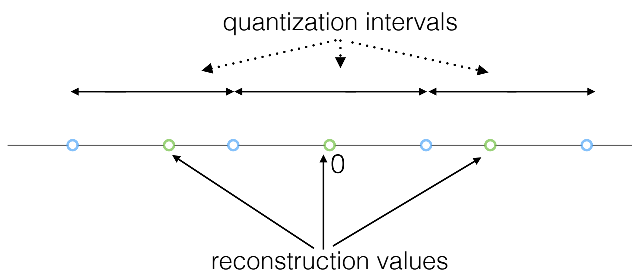

Fig. 1. The optimal scalar dead-zone quantizer for Laplacian sources with nearly uniform reconstruction rule.

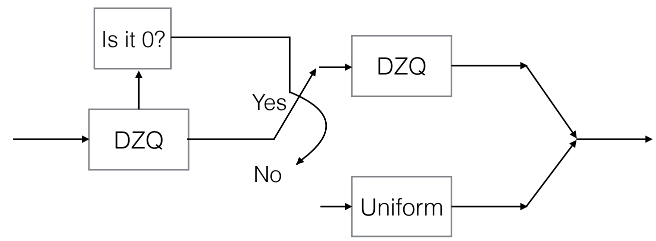

Fig. 2. Conditional enhancement layer quantizer for Laplacian sources. Based on the base layer DZQ interval, the enhancement layer quantizer is chosen.

where λ is Laplacian parameter.

Hence considerable attention has been focused on its optimal quantization, which is discussed in the following subsections.

2.1. Efficient Scalar Quantization of Laplacian Sources

In [11], the optimal entropy constrained quantizer for the Laplacian source was derived to be the dead-zone plus uniform threshold quantization classification rule and the nearly uniform reconstruction rule (as illustrated in Fig. 1). This dead-zone quantizer (DZQ) has uniform step size in all of the intervals, except the dead-zone interval around zero, which is wider than the other intervals.

2.2. Scalable Coding of Laplacian Sources

2.2.1. In Current Multimedia Standards

In current scalable coding standards such as, scalable HEVC [12] for video, and scalable AAC [13] for audio, the base layer employs DZQ for quantizing the source. Then, in the enhancement layer, a scaled version of the base layer DZQ quantizes the base layer reconstruction error.

2.2.2. Conditional Enhancement Layer Quantization (CELQ)

In [5], an efficient approach for scalable coding of Laplacian sources is proposed, wherein:

- The base layer employs a DZQ.
- The enhancement layer quantizers are conditioned on the base layer quantization interval: Use DZQ if a dead zone interval was established by the base layer, and use a uniform quantizer otherwise (as illustrated in Fig. 2).

This improved scalable coder still suffers from performance penalty compared to non-scalable coding.

3. COMMON INFORMATION BASED LAYERED CODING PARADIGM

In this section we explain the novel layered coding paradigm and importance of good quantizer design for this paradigm, with a toy example of uniform distribution. The best entropy constrained scalar quantization of a uniformly distributed random variable with rate $\log(N)$, where N is an integer, is a uniform quantizer with N quantization levels (as proven in [14]). Fig. 3(a) depicts the partition points for quantizing a uniform random variable, $U(0, 6)$, with rates $R_1 = 2$ and $R_2 = \log(6)$, resulting in distortion D_1 and D_2, respectively. Clearly, all the partition points of the quantizer 1, are not aligned with partition points of quantizer 2. This implies that scalable coding with base layer at rate 2, and enhancement layer at rate $\log(6) - 2$ (to achieve same receive rates at the decoders), will result in the enhancement layer distortion that is worse than independently quantizing at rate $\log(6)$. Hence, a uniformly distributed source is not successively refinable for mean squared error (MSE) distortion measure, and rates 2 and $\log(6) - 2$. The fact that a source is not successively refinable must imply that the information required to achieve D_1 is not a proper subset of the information required to achieve D_2. However, it is obvious that there is a considerable overlap in information required to reconstruct at the two distortion levels, thus independent encoding is wasteful. The common information based layered coding paradigm addresses this challenge by sending only part of the information required to achieve D_1 to the decoder reconstructing at lower distortion D_2. That is, the encoder generates 3 different packets (as illustrated in Fig. 4):

- At rate R_1, sent only to the decoder reconstructing at D_1
- At rate R_2, sent only to the decoder reconstructing at D_2
- At rate R_{12}, sent to both the decoders.

Conventional scalable coding is achieved in this paradigm when $R_1 = 0$, and non-scalable coding is achieved when $R_{12} = 0$. With appropriately designed quantizers, this framework provides the extra degree of freedom required to achieve rate-distortion optimality at both the layers with a total transmit rate lower than that of non-scalable coding. For our specific example of uniformly distributed source at receive rates of 2 and $\log(6)$, Fig. 3(b) depicts the quantizers for the layered coding paradigm, where rate $R_{12} = 1$ is sent to both decoders, and rates $R_1 = 1$ and $R_2 = \log(3)$ are sent to decoder 1 and 2, respectively. The overall quantizers with these partitions are same as the optimal individual quantizers, ensuring the same distortions are achieved at the decoders. However, we reduce the total transmit rate by 22% when compared to non-scalable coding. This example clearly demonstrates the utility of the proposed paradigm with appropriately designed quantizers extracting common information between different quality levels. In the following section we propose a design technique for quantizers of a general source distribution, to be employed within the common information based layered coding framework.

4. QUANTIZER DESIGN FOR COMMON INFORMATION BASED LAYERED CODING

For fixed received rates, $R_{r_1} = R_{12} + R_1 = c_1$ and $R_{r_2} = R_{12} + R_2 = c_2$, at decoder 1 and 2, respectively, there is a tradeoff

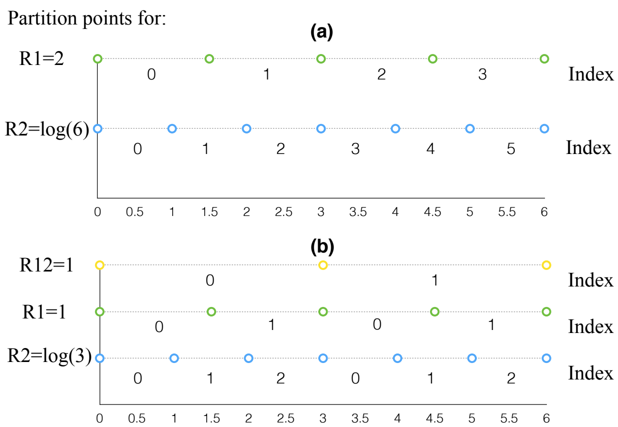

Fig. 3. Partition points for quantizing a $U(0,6)$. (a) Depicts individual quantizers at rates $R_1 = 2$ and $R_2 = \log(6)$. (b) Depicts quantizers for the common information based layered coding paradigm, where rate $R_{12} = 1$ is sent to both the decoders, and rates $R_1 = 1$ and $R_2 = \log(3)$ are sent to corresponding decoders.

between total transmit rate, $R_t = R_{12} + R_1 + R_2$, and sum of distortions at the decoders, $D_1 + D_2$. The two extremes of this tradeoff are: i) non-scalable coding, which uses the highest $R_t = c_1 + c_2$ while achieving the lowest distortions of $D^*(c_1) + D^*(c_2)$, where $D^*(\cdot)$ is the optimal distortion at a given rate; and ii) conventional scalable coding, which uses the lowest $R_t = c_2$, but significantly worse distortion than non-scalable coding at the enhancement layer. We would like to design our layered coding quantizers to optimize this tradeoff, thus we define our cost function as

$$J = R_t + \alpha \Delta D, \text{ s. t. } R_{r_1} = c_1, \ R_{r_2} = c_2,$$

where, $\Delta D = D_1 + D_2 - D^*(c_1) - D^*(c_2)$, and α controls the tradeoff. Minimizing this cost function gives us quantizers which achieve the best distortions at the decoders for a given total transmit rate and fixed receive rates. We design our quantizers in the following steps:

1. For the common layer, we design the optimal entropy constrained quantizer for the source distribution at a given rate, R_{12}.
2. For the two individual layers, we design optimal entropy constrained quantizers for each common layer quantizer interval, at their corresponding rates of R_1 and R_2.
3. We then numerically estimate the optimal common layer rate, by trying multiple allowed common rates and selecting the one that results in minimum cost J.

4.1. Quantizer Design for Laplacian Sources

For the practically important case of Laplacian source distribution the above generic design is adapted as below:

1. For the common layer, we estimate the best step size for the DZQ at a given rate, R_{12}.
2. For the two individual layers, we design optimal entropy constrained quantizers for each common layer quantizer interval, at their corresponding rates of R_1 and R_2. Specifically, we iteratively optimize the quantizer interval partitions and reconstruction points to minimize the entropy constrained distortion, with smart initializations of,
   - A DZQ for the dead zone interval, and
   - A uniform quantizer for other intervals,

   of the common layer quantizer.
3. We then numerically estimate the optimal common layer rate, by trying multiple allowed common rates and selecting the one that results in minimum cost J.

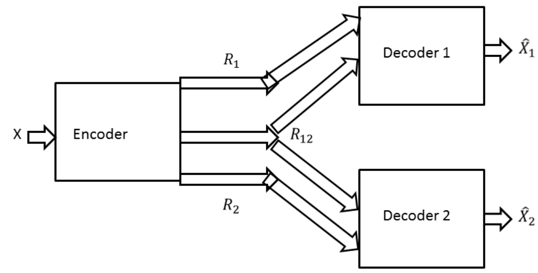

Fig. 4. Common information based layered coding paradigm. One packet at rate R_{12} is sent to both the decoders, and individual packets at rate R_1 and R_2 are sent to decoder 1 and 2, respectively.

Since the dead zone interval is a truncated Laplacian distribution and other intervals are a truncated exponential distribution, we select the initializations in Step 2 above to be the optimal entropy constrained quantizers of their corresponding non-truncated distributions.

Note that, we can achieve a non-zero common rate with negligible ΔD, if the DZQ at rate R_{12} is such that all its partition points align closely with partition points of DZQ at both rate c_1 and c_2. Conditions for such an alignment of partitions between two DZQ were derived in [15] as, the dead-zone of the coarser DZQ has to be divided into $2n+1$ intervals, and other intervals of this DZQ have to be divided into $m+1$ intervals, with $2n/m = z$, where, n and m are integers, and z is the ratio of the dead-zone interval length over other intervals' length. Our design technique numerically estimates the common layer DZQ which closely satisfies these conditions with DZQ at both rate c_1 and c_2.

Note that the proposed design technique does not ensure joint optimality of all the quantizers, particularly since we independently optimize the common layer quantizer (e.g., DZQ for Laplacian) without considering its effect on other layers. Despite this assumption we obtain considerable performance improvements (as will be discussed in Section 5). However, joint optimization of all the quantizers will be one of our future research directions.

5. EXPERIMENTAL RESULTS

In our experiments we used a Laplacian source with $\lambda = 1$ and the distortions at each decoder are measured in dB. For our first experiment we used fixed receive rates of $c_1 = 1.6$ and $c_2 = 2.8$. In Fig. 5 we plot, the R_t versus ΔD curve obtained by employing quantizers

	Non scalable total transmit rate $R_{12}+R_1+R_2=R_t^{NS}$	Proposed method total transmit rate $R_{12}+R_1+R_2=R_t^{P}$	Total transmit rate reduction $(R_t^{NS}-R_t^{P})/R_t^{NS}$
$R_{r_1}=1.6, R_{r_2}=2.8$	$0+1.6+2.8=4.4$	$0.4+1.2+2.4=4$	$(4.4-4)/4.4=9\%$
$R_{r_1}=1.5, R_{r_2}=2.3$	$0+1.5+2.3=3.8$	$0.3+1.2+2=3.5$	$(3.8-3.5)/3.8=8\%$
$R_{r_1}=1.4, R_{r_2}=2$	$0+1.4+2=3.4$	$0.3+1.1+1.7=3.1$	$(3.4-3.1)/3.4=9\%$

Table 1. Total transmit rate for non-scalable coding and proposed paradigm operating with negligible loss in distortion.

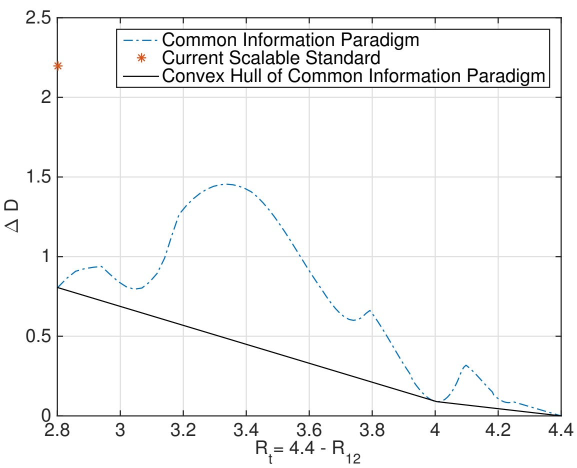

Fig. 5. Total transmit rate vs distortion deviations at the decoders.

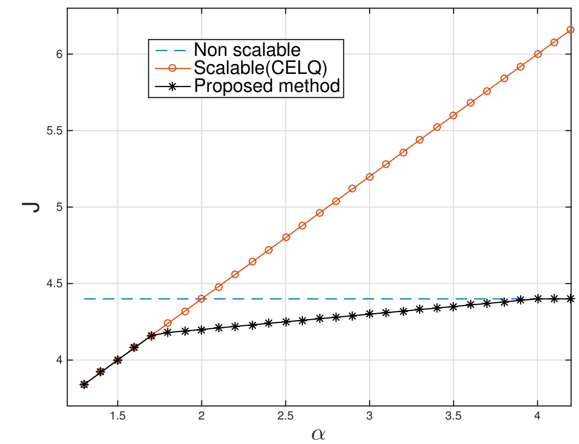

Fig. 6. Cost J vs tradeoff parameter α for non-scalable coding, scalable coding using CELQ, and proposed paradigm.

designed by our proposed technique at various common layer rates (R_{12}) ranging from 0 bits (i.e., non-scalable coding) to 1.6 bits (i.e., scalable coding using CELQ), and the convex hull for the proposed paradigm, which is obtained using the time sharing argument. Note that scalable coding employed in current standards has around 1.5 dB distortion loss compared to efficient scalable coding, which itself has around 0.8 dB distortion loss compared to non-scalable coding. The proposed technique can operate at all points along the convex hull and at considerably better performance compared to the scalable coding of current standards.

In Fig. 6, we plot J versus α for non-scalable coding, efficient scalable coding using CELQ, and proposed paradigm. We can see in this figure that the proposed paradigm bridges the non-scalable and scalable coding techniques, while performing at least as good as one of them and better than both of them at many operating points, which demonstrates the utility of the proposed technique in achieving better tradeoff between total transmit rate and distortions at the decoders.

Moreover, note that in Fig. 5 at $R_t = 4$ or equivalently $R_{12} = 0.4$, we obtain distortions that are very close to that of non-scalable coding at a 9% reduction in total transmit rate compared to that of non-scalable coding. Thus we conducted another experiment with different fixed receive rate combinations and obtained similar results of transmit rate savings with negligible distortion loss, which are shown in Table 1 to demonstrate the capability of the proposed technique to efficiently extract the common information between different quality levels.

6. CONCLUSION

This paper demonstrates a common information based layered coding framework with appropriately designed quantizers, which overcomes the limitations of conventional scalable coding and non-scalable coding, by providing the flexibility of transmitting common and individual bit-streams for different quality levels. The proposed quantizer design technique enables efficiently extracting common information between different quality levels with negligible performance penalty, and also enables achieving better operating points in the tradeoff between total transmit rate and distortions at the decoders. Experimental results for the practically important Laplacian sources, validate the superiority of the proposed approach.

7. REFERENCES

[1] J. Ohm, "Advances in scalable video coding," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 42–56, 2005.

[2] T. Painter and A. Spanias, "Perceptual coding of digital audio," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 451–515, 2000.

[3] J. Chow and T. Berger, "Failure of successive refinement for symmetric gaussian mixtures," *IEEE Transactions on Information Theory*, vol. 43, no. 1, pp. 350–352, Jan 1997.

[4] F. Wu, S. Li, and Y.Q. Zhang, "A framework for efficient progressive fine granularity scalable video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 3, pp. 332–344, Mar 2001.

[5] A. Aggarwal, S. Regunathan, and K. Rose, "Efficient bit-rate scalability for weighted squared error optimization in audio coding," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 4, pp. 1313–1327, July 2006.

[6] K. Viswanatha, E. Akyol, T. Nanjundaswamy, and K. Rose, "On common information and the encoding of sources that are

not successively refinable," *IEEE Information Theory Workshop (ITW)*, pp. 129–133, Sept. 2012.

[7] R. Gray and A. Wyner, "Source coding for a simple network," *The Bell Systems Technical Journal*, vol. 53, no. 9, pp. 1681–1721, Nov. 1974.

[8] A. Wyner, "The common information of two dependent random variables," *IEEE Transactions on Information Theory*, vol. 21, no. 2, pp. 163–179, Mar 1975.

[9] K. Viswanatha, E. Akyol, and K. Rose, "The lossy common information of correlated sources," *IEEE Transactions on Information Theory*, vol. 60, no. 6, pp. 3238–3253, 2014.

[10] T. Nanjundaswamy, K. Viswanatha, and K. Rose, "On relaxing the strict hierarchical constraints in layered coding of audio signals," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3680–3684, May 2014.

[11] G.J. Sullivan, "Efficient scalar quantization of exponential and laplacian random variables," *IEEE Transactions on Information Theory*, vol. 42, no. 5, pp. 1365–1374, Sept. 1996.

[12] Y. Yan and P. Andrivon, "The scalable extensions of hevc for ultra-high-definition video delivery," *IEEE MultiMedia*, vol. 21, no. 3, pp. 58–64, Sept. 2014.

[13] ISO/IEC JTC1/SC29 14496-3:2005, "Information technology - Coding of audio-visual objects - Part 3: Audio - Subpart 4: General audio coding (GA)," 2005.

[14] A. György and T. Linder, "Optimal entropy-constrained scalar quantization of a uniform source," *IEEE Transactions on Information Theory*, vol. 46, no. 7, pp. 2704–2711, November 2000.

[15] G.J. Sullivan, "On embedded scalar quantization," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, pp. 605–608, May 2004.