Signal Processing: Image Communication 63 (2018) 113–124

Contents lists available at ScienceDirect

Signal Processing: Image Communication

journal homepage: www.elsevier.com/locate/image

Optimal Lagrange multipliers for dependent rate allocation in video coding

Ana De Abreu [a,*], Gene Cheung [b], Pascal Frossard [c], Fernando Pereira [d]

[a] Trinity College Dublin, Dublin, Ireland
[b] National Institute of Informatics, 2-1-2, Hitotsubashi, Chiyoda-ku, Tokyo, 101–8430, Japan
[c] Signal Processing Laboratory (LTS4), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
[d] Instituto Superior Técnico - Instituto de Telecomunicações (IST-IT), Lisbon, Portugal

ARTICLE INFO

Keywords:
Lagrangian optimization
Video and multiview image coding
Rate–distortion (RD) optimization

ABSTRACT

In a typical video rate allocation problem, the objective is to optimally distribute a source rate budget among a set of (in)dependently coded data units to minimize the total distortion. Conventional Lagrangian approaches convert the lone rate constraint to a linear rate penalty scaled by a multiplier in the objective, resulting in a simpler unconstrained formulation. However, the search for the "optimal" multiplier – one that results in a distortion-minimizing solution among all Lagrangian solutions that satisfy the original rate constraint – remains an elusive open problem in the general setting. To address this problem, we are the first in the literature to construct a computation-efficient search strategy to identify this optimal multiplier numerically in the general dependent coding scenario. Specifically, we first formulate a general rate allocation problem where each data unit can be dependently coded at different quantization parameters (QP) using a previous unit as predictor, or left uncoded at the encoder and subsequently interpolated at the decoder using neighboring coded units. After converting the original rate-constrained problem to the unconstrained Lagrangian counterpart, we design an efficient dynamic programming (DP) algorithm that finds the optimal Lagrangian solution for a fixed multiplier. In extensive monoview and multiview video coding experiments, we show that for fixed target rate constraints, our algorithm is able to find the optimal multipliers in a distortion minimum sense among all Lagrangian solutions. Moreover, we show that our simple solution is able to compete with complex rate control (RC) solutions used in video compression standards such as HEVC and 3D-HEVC, which outlines the importance of the proper choice of the Lagrangian multipliers.

1. Introduction

In video coding, rate allocation is the problem of distributing a source bit budget B to a set of (in)dependently coded data units, $v \in \mathcal{V}$, in order to minimize the distortion of all units. For example, a data unit v can be a video frame predictively coded at a quantization parameter (QP) q_v, using a previous frame as a predictor. Coding at a larger (coarser) QP requires fewer bits in general but results in a higher quantization distortion. In some cases, leaving a unit uncoded at the encoder may be a better rate–distortion (RD) decision; the unit is subsequently interpolated at the decoder using neighboring coded units via techniques such as motion compensated interpolation (MCI) [1] in monoview video or depth-image based rendering (DIBR) [2,3] in multiview video when color and depth maps are available. In these cases, the more general rate allocation problem is to first select data units $\mathbf{v} \subseteq \mathcal{V}$ for coding, and then select a set of QPs $\mathbf{q} = \{q_1, \ldots, q_{|\mathbf{v}|}\}$ at which to code the selected units $\mathbf{v}$ to minimize the total distortion $D(\mathbf{v}, \mathbf{q})$ subject to a rate constraint $R(\mathbf{v}, \mathbf{q})$:

$$\min_{\mathbf{v} \subseteq \mathcal{V}, \mathbf{q}} D(\mathbf{v}, \mathbf{q}) \quad \text{s.t. } R(\mathbf{v}, \mathbf{q}) \leq B. \tag{1}$$

To address different variants of the rate allocation problem, Lagrangian approaches – where the lone rate constraint is first converted to a linear rate penalty in the objective scaled by a multiplier λ – are common in the literature [4–8]. This results in a simpler unconstrained problem:

$$(\mathbf{v}_\lambda, \mathbf{q}_\lambda) = \arg \min_{\mathbf{v} \subseteq \mathcal{V}, \mathbf{q}} D(\mathbf{v}, \mathbf{q}) + \lambda\, R(\mathbf{v}, \mathbf{q}) \tag{2}$$

which in general is easier to solve for a fixed multiplier λ [4–8]. However, the Lagrangian relaxed problem (2) is inherently not the same as the original rate-constrained problem (1); the difference in distortion between their respective optimal solutions is called a *duality gap* (see Appendix B and [4]).

To minimize this gap, it is imperative to find the "optimal" multiplier λ^* — one that results in a distortion-minimizing solution $(\mathbf{v}_{\lambda^*}, \mathbf{q}_{\lambda^*})$ among all Lagrangian solutions $(\mathbf{v}_\lambda, \mathbf{q}_\lambda)$ to (2) for different λ that satisfy

* Corresponding author.
E-mail addresses: deabreua@scss.tcd.ie (A. De Abreu), cheung@nii.ac.jp (G. Cheung), pascal.frossard@epfl.ch (P. Frossard), fp@lx.it.pt (F. Pereira).

https://doi.org/10.1016/j.image.2018.02.003
Received 4 June 2017; Received in revised form 11 December 2017; Accepted 4 February 2018
Available online 17 February 2018
0923-5965/

$R(\mathbf{v}_\lambda, \mathbf{q}_\lambda) \le B$. However, given empirical discrete rate and distortion functions $R(\mathbf{v}, \mathbf{q})$ and $D(\mathbf{v}, \mathbf{q})$, the search for this optimal multiplier numerically without resorting to continuous rate and distortion models [9,10] remains an open problem in the general setting.

To address this problem, we are the first in the literature to construct a computation-efficient search strategy to find this optimal multiplier numerically in the general dependent coding scenario. Specifically, we first formulate a general rate allocation problem, where each data unit can be dependently coded at different QPs using a previous coded unit as predictor, or left uncoded at the encoder for interpolation at the decoder using neighboring coded units as reference. After converting the original rate-constrained problem to the unconstrained Lagrangian counterpart (2), we design an efficient dynamic programming (DP) algorithm that finds the optimal solution to (2) for a fixed λ.

To find the optimal multiplier, we iteratively compute neighboring *singular multiplier values* [4] from the computed DP solution; each singular value λ results in multiple simultaneously optimal solutions with different rates, *i.e.* $R(\mathbf{v}_\lambda^{(1)}, \mathbf{q}_\lambda^{(1)}) \neq R(\mathbf{v}_\lambda^{(2)}, \mathbf{q}_\lambda^{(2)})$. We show that singular values alone lead to all Lagrangian solutions to (2). When we obtain solutions $(\mathbf{v}_{\lambda^*}^{(1)}, \mathbf{q}_{\lambda^*}^{(1)})$ and $(\mathbf{v}_{\lambda^*}^{(2)}, \mathbf{q}_{\lambda^*}^{(2)})$ corresponding to singular value λ^* with rates $R(\mathbf{v}_{\lambda^*}^{(1)}, \mathbf{q}_{\lambda^*}^{(1)}) \le B \le R(\mathbf{v}_{\lambda^*}^{(2)}, \mathbf{q}_{\lambda^*}^{(2)})$, we prove that $(\mathbf{v}_{\lambda^*}^{(1)}, \mathbf{q}_{\lambda^*}^{(1)})$ is the distortion-minimizing Lagrangian solution, and declare λ^* as the optimal multiplier. To the best of our knowledge, no previously proposed Lagrangian multiplier searches [4–8] provide this theoretical claim in our general setting.[1]

Experimental results illustrate the good performance of our proposed rate allocation algorithm with optimal Lagrange multiplier selection when data units are independently or predictively coded in both monoview and multiview video sequence compression. Specifically, we show that our algorithm is able to find the closest rate values given a fixed bit budget constraint, driven by the optimal selection of Lagrangian multipliers. Moreover, we show that our bit allocation strategy is able to compete with complex rate control (RC) solutions adopted in the reference softwares of current monoview and multiview video standards, namely HEVC [11] and 3D-HEVC [12], that do not skip frames in Y-PSNR. These illustrative results show that our novel and generic algorithm for optimal selection of Lagrange multiplier value can bring large benefits in complex rate allocation problems.

The paper is organized as follows. We first review related work in Section 2 and formulate our dependent rate allocation problem in Section 3. We then describe a DP algorithm that solves the rate-constrained problem optimally but in exponential time in Section 4. To reduce complexity, we convert the problem to the Lagrangian relaxed version and propose a corresponding polynomial-time DP algorithm for a fixed multiplier. In Section 5, we discuss an efficient search methodology to identify the optimal multiplier based on the computation of neighboring singular multiplier values. Finally, we present experimental results and conclusion in Sections 6 and 7, respectively.

2. Related work

2.1. Continuous RD function modeling

In order to obtain rate and distortion functions for different data units, one can simply apply different QPs to encode each data unit and observe the resulting rate and distortion values. Then, these empirical RD points are fitted into mathematical functions for each particular sequence to derive an RD model [13,14]. However, in these types of models some parameters have to be estimated for each given video sequence and therefore they cannot be easily generalized.

Alternatively, it is possible to theoretically derive the different parameters of the RD model under simplifying assumptions. Several RD models using well understood exponential functions have been proposed [9,10]. In [9], both a Laplacian and a Generalized Gaussian (GG) distribution have been considered in their RD model for a wavelet video coding. In [10], the Bernoulli Generalized Gaussian (BGG) model have been adopted for both rate and distortion functions. The work in [15] proposed a gradient based R-lambda (GRL) model for intra frame rate control and a new bit allocation scheme for the coding tree unit (CTU) rate control. Specifically, bit per pixel (BPP), gradient per pixel (GPP) and λ are modeled using a hyperbolic function. The authors then proposed a bit allocation scheme at three levels: GOP, frame and CTU. The authors in [16] proposed a pixel-wise unified rate-quantization (URQ) model working at the multi-level regardless of block sizes, where texture complexity is estimated using a linear MAD-based measurement. For the R–Q model, the authors assume that the ratio of distortion over bits is constant, while the parameter λ used to trade off distortion and rate is chosen using an empirical formula. The work in [17] proposed a rate control scheme for HEVC based on new rate models (Laplacian distributions) for texture and non-texture bits, taking into account different statistical characteristics at different depths of coding units (CU). Because there is no differential coding among the CU, coding parameters for each CU are selected independently. Bits for each frame are allocated ignoring inter-frame dependency (*i.e.*, a poorly coded frame F_i can negatively affect the coding performance of the following frame F_{i+1} that uses F_i as predictor during motion prediction.) In general, these approaches suffer from: (i) modeling errors due to idealized model inaccuracy; and (ii) continuous approximation error since the problem (selecting QPs from a finite set) is inherently discrete. In contrast, we take an empirical approach and solve the inherent discrete problem of selecting data units and QPs for coding directly, and thus do not suffer from modeling errors.

2.2. Constrained formulation via dynamic programming

Addressing directly the discrete rate-constrained bit allocation problem, one common approach is Dynamic Programming (DP). For example, assuming independently coded data units, [18] constructed a tree to represent all possible solutions: a node at a stage i of the tree represents a particular selection of QPs $\{q_1, \dots, q_i\}$ for data unit 1 to i. If two different nodes at the same stage i have the same accumulated rate from unit 1 to i, then the one with the larger accumulated distortion would be pruned. As we will show in Section 3, the complexity of this type of DP algorithms is *pseudo-polynomial* or exponential time. If predictive coding is assumed, complexity is even higher.

To jointly select predictor frames and QoS levels for encoding and protection of different video frames during network streaming, [19] proposed an integer rounding approach to reduce complexity of a DP algorithm, where DP tables used to store computed local solutions were scaled down to reduce the number of table entries. The authors derived a performance bound for the proposed reduced-complexity DP algorithm; however, this bound gets progressively worse as the scale factor increases.

For multiview, assuming independently coded units, [20] considered a uniform rate allocation among views in a multiview video system, and proposed a DP-based algorithm to select the views for encoding and transmission such that the expected distortion among encoded and synthesized views is minimized given a rate budget. This work is extended in [21] where both the views and the coding rates for each selected view are selected to minimize distortion in a rate-constrained scenario. Due to high complexity, a greedy DP algorithm was proposed. There is no performance guarantee for the greedy DP algorithm, however, and thus the obtained solution can be arbitrarily far from the optimal solution.

2.3. Discrete Lagrangian formulation

Instead of the originally posed rate-constrained bit allocation problems, the Lagrangian approach is often used. [5] proposed a trellis-based algorithm to find the optimal Lagrangian solution for predictively coded

[1] [4] was a seminal bit allocation work and proposed the first singular value search strategy but only for independently coded data units.