Optik 125 (2014) 3980–3985

Contents lists available at ScienceDirect

Optik

journal homepage: www.elsevier.de/ijleo

The rate-distortion optimized quantization algorithm in Compressive Sensing

Wei Jiang*, Junjie Yang

School of Computer and Information Engineering, Shanghai University of Electric Power, Shanghai 200090, China

ARTICLE INFO

Article history:
Received 29 July 2013
Accepted 26 January 2014

Keywords:
Rate-distortion
Quantization
Lagrange multiplier method
Compress sensing

ABSTRACT

Rate-distortion optimization greatly improves performance of compression coding system. In this paper, the rate-distortion optimized quantization algorithm is proposed for block-based Compressive Sensing. Compressive Sensing is the emerging technology which can encode a signal into a small number of incoherent linear measurements and reconstruct the entire signal from relatively few measurements. In the algorithm the sampling measurements are quantized optimally based on the rate-distortion theory. For the coefficients near dead-zone the quantization level with best rate-distortion performance is chosen. Moreover, in order to acquire the best performance, a fast Lagrange multiplier solving method is proposed to find the optimal slope λ^* of the rate-distortion curve at the given bit budget. Experimental results show that the proposed algorithm improves objective and subjective performances substantially. The average gain about 0.7 dB can be achieved with the same rate.

1. Introduction

In digital image coding, the best coding performance originates from Shannon's rate-distortion function. It shows the minimum distortion can be acquired subject to the given rate R_c without consideration of coding complexity, delay and memory. In practical coding systems the optimal performance can be achieved by selecting the proper parameters in the chosen implementations for the given set of test data. For the given coding system, rate-distortion optimization (RDO) can be used to improve the performance. The target of the RDO is to minimize the distortion for a given rate by appropriate selections of coding parameters. Generally, rate-distortion optimization can be used to optimize the quantizer [1–4]. The rate control scheme can be improved by RDO [5–8]. Moreover, discrete cosine transform is optimized to approximate the optimum transform for intraprediction residue [9,10]. To solve such a constrained problem, there are two popular approaches: Lagrangian optimization [11] and Dynamic Programming (DP) [12]. Compared to DP, Lagrangian optimization is simple and easily implemented. Therefore it is widely used in the image compression coding system to improve the performance.

In 2006 Compressive Sensing (CS) was proposed by Donoho and Candes [13,14]. By using linear measurement, it converts the signals in a high dimension space into the signals in a smaller dimension space. If the input signal is sufficiently sparse, it can be reconstructed in high precision. Compressive sensing has been successfully applicable to image compression [15–19]. Moreover, some theoretical results regarding the rate-distortion performance of CS have been published recently. Adriana Schulz carried out an empirical analysis of the rate-distortion performance of CS in image compression [20]. The issues such as the minimization algorithm used, the transform employed and quantization error are analyzed. The corresponding rate distortion function is got in case of uniform and non-uniform quantization [21]. A convex optimization technique is proposed for the compressive sensing reconstruction that uses quantized measurements. The CS reconstruction errors can be improved about 4 dB [22].

In this paper, the rate-distortion optimized quantization algorithm is proposed for compressive sensing. The proposed algorithm focuses on the wavelet coefficients near the dead-zone. The quantization indices of these coefficients are determined according to rate-distortion criterion. Moreover, a fast Lagrange multiplier solving method is proposed in order to determine best rate-distortion point for the given bit budget.

The remaining part of this paper is organized as follows. Section 2 provides introduction to the Lagrange multiplier method. Section 3 highlights the intuition and main idea of the algorithm. Section 4 describes the proposed algorithm in detail. The performance comparison is carried out in Section 5. Finally, this paper is concluded in Section 6.

* Corresponding author.
E-mail addresses: zzm7406@sina.com (W. Jiang), shiepdouma@163.com (J. Yang).

http://dx.doi.org/10.1016/j.ijleo.2014.01.147
0030-4026/

2. Rate-distortion optimization

Rate-distortion optimization (RDO) plays a critical role in the hybrid video coding. It aims at minimizing the distortion for a given rate R_c by appropriate selections of coding parameters,

$$\min_{s\in S}\{D(s)\} \quad \text{s.t.}\, R(s) \le R_c, \tag{1}$$

where S is the set of coding parameters, $D(s)$ is the distortion caused for the parameter configuration s in S and $R(s)$ be the rate for s in S. Lagrangian optimization converts the constrained problem to the following unconstrained form by introducing Lagrange multiplier λ,

$$\min_{s\in S}\{J(\lambda) = D(s) + \lambda R(s)\}, \tag{2}$$

where λ is the Lagrangian multiplier. The different values of λ corresponds the different rate-distortion characteristics. Generally speaking, a large value of λ corresponds to the higher distortion and the lower bit-rate. On Contrary, a smaller λ corresponds to the lower distortion and the higher bit-rate. The right lambda should be found in order to achieve the optimal solution at the required rate. Since for different positive values of λ, the unconstrained problem results in the tracing out of convex hull points of the rate-distortion curve, the optimal convex hull point is the one with the minimum distortion in the case of $R \le R_c$. Since the solution of Eq. (2) is the point at which the line of absolute slope is tangent to the convex hull of the R-D characteristic, λ is normally referred as the slope of the R-D curve. When the coding system is operated at the point where the line with negative slope is the tangent to the convex hull of the R-D curve, the target of RDO can be easily achieved at the same time.

3. Basic idea of the algorithm

The rate-distortion performance of CS can be improved by increasing the efficiency of CS measurement quantization. According to rate-distortion theory, the optimal quantization performance can be acquired by minimizing D subject to a bit budget constraint $R \le R_c$. Apparently, quantizing every value in such a way can produce high performance. However, it is simply impossible because of huge complexity. It is known that the coefficients of the common transform, e.g. DCT and DWT, are Laplacian distribution which is sharp in small signal zone. In addition, the uniform quantizer which is popular in many image and video compression systems performs poor in the dead-zone. Thus it is reasonable to focus on the coefficients near dead-zone. For convenience, we define the processed window W_z containing coefficients. Generally the processed window can be defined as the first and second quantization cell. Accordingly the possible quantization levels are limited into $I = \{0, 1\}$. Then the optimal quantization level of the coefficients falling into the processed window can be gotten:

$$\min_{z\in I} D(x, z) = D_Q(x, z) \quad \text{s.t.}\, R(z) \le R_c, \tag{3}$$

where x is the input signal, z is the quantization level, I represents the finite set of all admissible quantizer choices and D_Q is the quantization distortion.

3.1. Unconstrained optimization approach

The constrained optimization problem of (3) can be solved by converting it to an unconstrained problem using Lagrange multiplier λ. The optimal convex hull point can be obtained by

$$z^* = arg \min_{x\in W_z, z\in I}\{J_\lambda(x, z) = D(x, z) + \lambda \times R(z)\} \tag{4}$$

In this case the optimal quantization performance can be acquired by finding the minimum rate-distortion cost. Therefore the optimal rate-distortion slope λ^* should be found correctly.

3.2. Finding the optimal slope

In this paper, the biased Lagrange cost $J_B(\lambda)$ similar to the bisection search algorithm is proposed [23]. The biased Lagrange cost $J_B(\lambda)$ is defined as follows

$$J_B(\lambda) = J_B(\lambda, x^*(\lambda)) = \lambda R_c - J^*(\lambda) = \lambda R_c - \min_{s\in S}(D(s) + \lambda R(s)), \tag{5}$$

where S is the set of coding parameters. We have the following results:

1. $J_B(\lambda)$ is the convex function of λ.

$$\begin{aligned} J_B(\lambda_3) &= \lambda_3 R_c - J^*(\lambda) = \lambda_3 R_c - \min_{s\in S}\{D(s) + [\theta\lambda_1 + (1-\theta)\lambda_2]R(s)\} \le \\ &\quad \theta \min_{s\in S}\{[\lambda_1 R_c - D(s) - \lambda_1 R(s)] + (1-\theta)\min_{s\in S}[\lambda_2 R_c - D(s) \\ &\quad - \lambda_2 R(s)]\} = \theta J_B(\lambda_1) + (1-\theta) J_B(\lambda_2), \end{aligned} \tag{6}$$

where $\lambda_3 = \theta\lambda_1 + (1-\theta)\lambda_2$, $0 \le \theta \le 1$.

2. λ^* and $s^*(\lambda^*)$ that minimize $J_B(\lambda)$ are the optimal slope and optimal operating point for the given budget constraint.

Suppose λ' is the slope of the convex hull face which "straddles" the budget constraint line on the R-D plane. For $\lambda < \lambda'$, invoking the lower rate operating point s'

$$\begin{aligned} J_B(\lambda) \quad -J_B(\lambda') &= \min_{s\in S}\{[\lambda R_c - (D(s) + \lambda R(s))] \\ &\quad - [\lambda' R_c - (D(s) + \lambda' R(s'))]\}]\} \ge [\lambda R_c - (D(s') + \lambda R(s'))] \\ &\quad - [\lambda' R_c - (D(s') + \lambda' R(s'))] \ge (\lambda - \lambda')(R_c - R(s')) \ge 0. \end{aligned} \tag{7}$$

Similarly, for $\lambda > \lambda'$, invoking the higher rate operating point s'', it can be got

$$J_B(\lambda) - J_B(\lambda') \ge (\lambda - \lambda')(R_c - R(s'')) \ge 0. \tag{8}$$

Therefore, for all positive values of λ, $J_B(\lambda') < J_B(\lambda)$. It is proved that the optimal slope for the bit budget R_c is the point minimizing the biased cost $J_B(\lambda)$. Then the problem of searching the optimal slope is converted to solving the minimum value of $J_B(\lambda)$. The optimal slope λ^* can be found:

$$J_B(\lambda^*) = J_B(\lambda^*, s^*(\lambda^*)) = \min_{\lambda\ge 0} J_B(\lambda). \tag{9}$$

Since $J_B(\lambda)$ is convex, the optimal slope λ^* is unique. In addition, no matter how the operational point is distributed on R-D curve, the minimization can be searched definitely. Consequently the problem of searching the optimal slope is converted to solving the minimum value of $J_B(\lambda)$. Apparently it can be accomplished by the fast convex search algorithm like Newton's method or bisection method [24].

With above discussion, the optimal quantization performance can be acquired by solving

$$\overset{(b)}{\min_{\lambda\ge 0}} \left(\lambda R_c - \overset{(a)}{\min_{x\in W_z, z\in I}} [(D(s) + \lambda R(s))] \right), \tag{10}$$

where the innermost minimization (a) involves the search for the best quantization level for signals in the processed window W_z for fixed operating slope λ and the outermost optimization (b) is the convex search for λ^* subject to R_c. Accordingly, the final solution (λ^*, z^*) is obtained in two sequential optimization steps.