OPTIMAL SHAPE CODING UNDER BUFFER CONSTRAINTS

Jae-Beom Lee, Jin-Soo Cho and Alexandros Eleftheriadis

Department of Electrical Engineering
Columbia University, New York, NY 10027
{jbl,eleft,cjs}@ee.columbia.edu

ABSTRACT

In this paper, we consider an optimal rate control algorithm for shape coding of emerging MPEG-4 compression system. Current MPEG-4 shape coding consists of two sub-steps; First distortion is introduced by scaling down and up way. Then, second, context-based arithmetic coding is applied. Since the arithmetic coding is "lossless", scaling down and up step is considered as a virtual quantizer. We first formulate the buffer-constrained adaptive quantization problem for the shape coding. And then we propose an algorithm for the optimal solution under the buffer constraints. Experimental results are given using MPEG-4 shape coder.

1. OPTIMAL BUFFERED COMPRESSION

Recently "optimal Trellis-based buffered compression" was proposed in [2] to provide optimal buffer control strategies for signal sequences in a finite buffered environment. Authors, first, formalized the description of the buffer-constrained adaptive quantization problem. And, second, they formulated the optimal solution for a given set of admissible quantizers used to code a discrete nonstationary signal sequence in a finite buffer.

Optimal buffered compression is defined as followings.

Problem Definition: Given a set of quantizers, a sequence of blocks to be quantized, and a finite buffer, select the optimal quantizer for each block so that the total cost measure is minimized and the finite buffer is never in overflow.

Consider the allocation for N blocks, and suppose there are M quantizers available to code each block. Let d_{ij} and r_{ij} be, respectively, the distortion and the number of bits produced by the coding of block i with quantizer j, and let r be the channel rate per block. Define an admissible solution x as a selection of one quantizer for each block, i.e., a mapping from $1, 2, ..., N$ to $1, 2, ..., M$, $x = \{x(1), x(2), ..., x(N)\}$, where each $x(i)$ is the index of one of the M quantizers for block i. Therefore, $(r_{1x(1)}, ..., r_{Mx(N)})$ and $(d_{1x(1)}, ..., d_{Mx(N)})$ are, respectively, the rate and distortion for each block and a given choice of quantizers x.

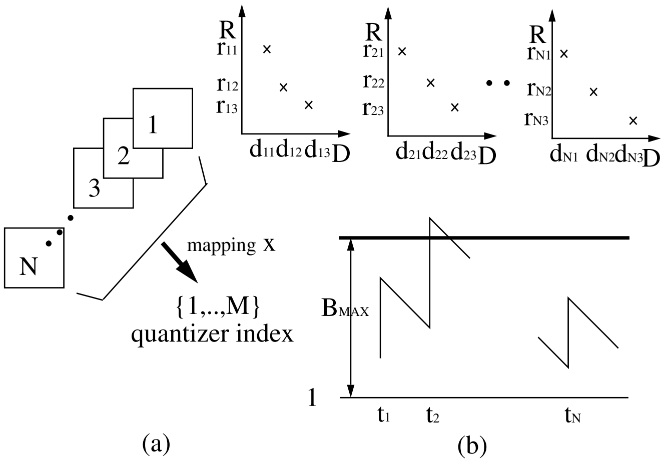

Figure 1: (a) Each block in the sequence has a different R-D characteristic (For a given choice of quantizers for the blocks in the sequence, we can obtain R-D points to form the composite characteristic), (b) R^* is not a feasible solution with the chosen buffer size (Buffer is limited).

Now, define the buffer occupancy at stage i, $B(i)$ for a given admissible solution x. To account for the fact that the buffer occupancy cannot be negative at any stage (i.e., underflow means the buffer occupancy is 0), we use a recursive definition. Let $B(1) = r_{1x(1)} + B(0), B(2) = max(B(1) + r_{2x(2)} - r, 0)$ and, in general

$$B(i) = max(B(i-1) + r_{ix(i)} - r, 0) \tag{1}$$

where the buffer occupancy at each block instant is increased by the coding rate of the current block and decreased by the channel rate. $B(0)$ is the initial buffer state.

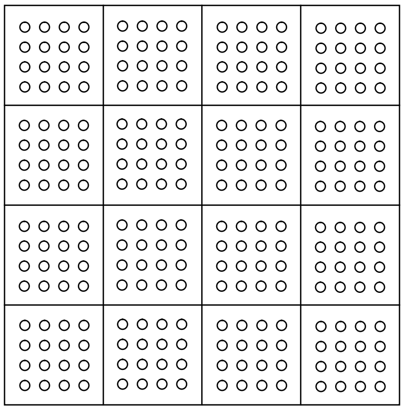

Figure 2: A BAB consists of 16 PBs.

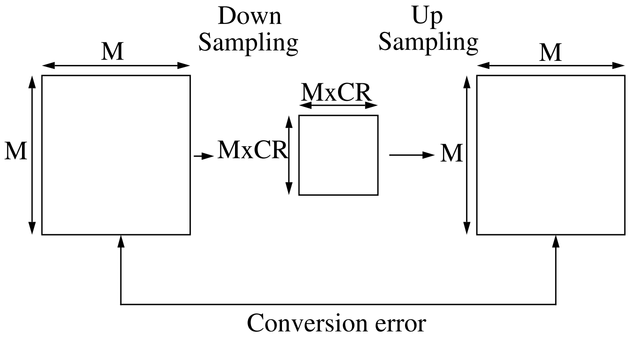

Figure 3: Size-conversion.

General Formulation : (Integer Programming)
The problem is to find the mapping x that solves

$$min(\sum_{i=1}^{N} d_{ix(i)}), \quad (2)$$

subject to

$$B(i) \leq B_{max}, \forall i = 1, ..., N$$

where B_{max} is the buffer size.

In the MPEG-4 video context, previous optimal buffered compression can be thought of the optimal solution for the "texture coding". The importance of MPEG-4 as an industry standard with extensive future use of interactive multimedia systems suggests further investigation on the optimal buffer control issue on shape information as well. This paper addresses the shape counter part of optimal rate control under buffer constraints.

Current MPEG-4 binary shape coding consists of two sub-steps; First distortion is introduced by the way shape data are scaled down and up, based on thresholding binary image. Then, second, context-based arithmetic coding is applied. Since arithmetic coding itself is lossless, we will deal scaling down and up as a "virtual quantizer". We will, above all, formulate the buffer-constrained adaptive quantization problem for a shape coding. And then we propose an algorithm for the optimal solution.

Now the problem remains in the question of what the quantier index in "shape" coding context is like.

(b) R^* is not a feasible solution with the chosen buffer size (Buffer is limited).

2. OPTIMAL BUFFERED COMPRESSION FOR MPEG-4 SHAPE CODING

In current MPEG-4, a binary alpha plane can be encoded in INTRA mode for I-VOPs and in INTER mode for P-VOPs and B-VOPs. The methods used are based on binary alpha blocks, and the principal method is block-based context-based arithmetic encoding (CAE) with block-based motion compensation. For detail explanation of MPEG-4 shape coding, we refer to MPEG-4 VM documment [1].

Current rate control and rate reduction in MPEG-4 is realized through size-conversion of the binary alpha information as in Figure 3. In the size conversion process, the determination of CR is done based on a given distortion parameter for acceptance (i.e., $alpha\_th$). That is, it is necessary to ascertain whether a BAB has an accepted quality under the size conversion process with a specific CR. The criterian is based on a 4×4 pixel block (PB) data structure. Each BAB is composed of 16 PBs as illustrated in Figure 2.

The measure of acceptance is following: Given the current original binary alpha block i.e. BAB and some approximation of it (i.e. BAB'), it is possible to define a function

$$ACQ(BAB') = MIN(acq_1, acq_2, ..., acq_{16}) \quad (3)$$

where,

$$acq_i = \begin{cases} 0, & \text{if } SAD\_PB_i > 16 \times alpha\_th \\ 1, & \text{otherwise} \end{cases}$$

and $SAD\_PB_i(BAB, BAB')$ is defined as the sum of absolute differences for PB_i, where an opaque pixel

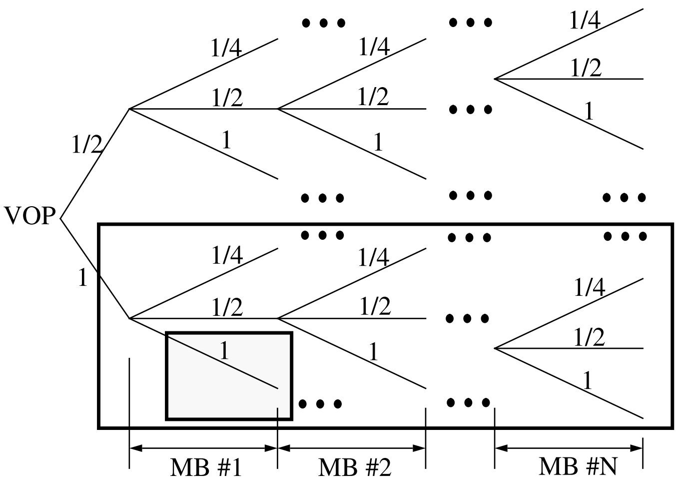

Figure 4: The algorithm: branching procedure (Branch value means VOP_CR and MB_CR).

has value of 255 and a transparent pixel has value of 0. The parameter *alpha_th* has values {0,16,32,64,...,256}. If $alpha\_th = 0$, then encoding will be lossless. A value of $alpha\_th = 256$ means that the accepted distortion is maximal i.e. in theory, all alpha pixels could be encoded with an incorrect value.

Note that in current VM rate control, only quality measure is taken into account; There is no buffer constrants. Therefore, a problem comes up when the encoding buffer is concerned (i.e., in CBR applications). To introduce a rate control concept, we consider the "optimal buffered compression" for the MPEG-4 shape coding.

In the shape coding, distortion is introduced in two levels; In VOP level, the first of all, size conversion is performed. And then, in Macroblock level, extra size conversion is done. If this is tought of "virtual quantizer", we can formulate it for shape coding. Note that in this formulation the *alpha_th* value is replaced by $(VOP(B(i))\_CR, B(i)\_CR)$ conceptually, thus making *alpha_th* erased. Therefore, problem formulation can be given as followings.

Shape Coding Formulation : (Integer Programming)
Let x be $\{x(1), x(2), ..., x(N)\}$ where
$x(i) = (VOP(B(i))\_CR, B(i)\_CR)$.
The problem is to find the mapping x that solves

$$min(\sum_{i=1}^{N} d_{ix(i)}), \tag{4}$$

subject to

$$B(i) \leq B_{max}, \forall i = 1, ..., N$$

where B_{max} is the buffer size.

In the above expressions, $B(i)$ means i-th Macroblock, and $VOP(B(i))$ means VOP in which $B(i)$ is.

3. ALGORITHM AND RESULTS

Recent texture part for "optimal buffered compression", which is proposed in [2], was more or less infeasible since the number of quantizers is usually not small; this makes the algorithm extremly computational intensive. However, the shape part in this paper is quite feasible because the number of quantizers is just 3 (i.e., 1, 1/2, and 1/4). The algorithm used here is to choose CR set in following order: (1,1), (1,1/2), (1,1/4) based on the content of the buffer. If any of these values cannot make the buffer under maximum size, one step previous MB is chosen in the branch for the same computation. Note that we choose $VOP(B(i))\_CR = 1$ for simplicity, which is illustrated in Figure 4. Figure 5 is the result of the algorithm.

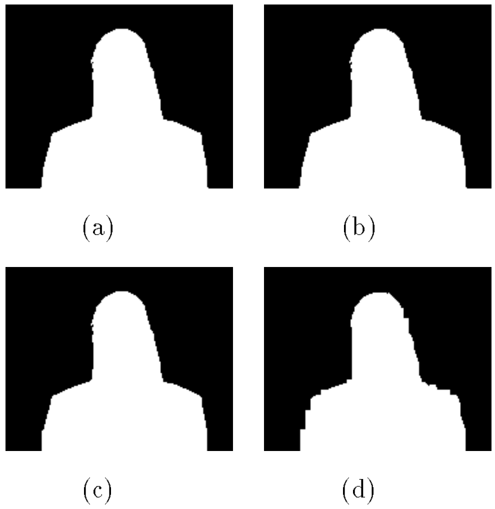

Figure 5: Decoded outputs (MPEG-4 shape decoder $B_{max} = 2000$, $r = 5$) : (a) 1st frame, (b) 2nd frame, (c) 3rd frame, (d) 4th frame.

4. REFERENCES

[1] ISO. Draft MPEG-4 : Video Verification Model Version 8.0. *ISO/IEC JTC1/SC29/WG11*, July 1997.

[2] A. Ortega, K. Ramchandran, and Martin Vetterli. Optimal Trellis-based buffered compression and fast approximations. *IEEE Iransactions on Image Processing*, 3(1):26–40, January 1994.