On Using Game Theory to Optimize the Rate Control in Video Coding

Ishfaq Ahmad, *Senior Member, IEEE,* and Jiancong Luo, *Member, IEEE*

***Abstract*—This paper presents a game theory based technique for optimizing the bit rate control in video coding. Game theory, by virtue of its enormous potential for solving constrained optimization problems, has been effectively utilized in several branches of natural and social sciences. But this paper is the first attempt in using game theory for video compression. The objective is to optimize the perceptual quality while guaranteeing "fairness" in bit allocation among macroblocks (MBs). The proposed technique is a dual-level rate control algorithm: At the first level, the algorithm allocates the target bits to frames based on their coding complexity; a method to estimate the coding complexity of the remaining frames is proposed. At the second level, MBs of a frame play cooperative games such that each MB competes for a fair share of resources (bits) to optimize its quantization scale while considering the human visual system (HVS) perceptual property. We formulate the rate control problem by defining players, strategies and objective function. Since the whole frame is an entity perceived by viewers, MBs compete cooperatively under a global objective of achieving the best quality with the given bit constraint. The major advantage of the proposed approach is that the cooperative game leads to an optimal and fair bit allocation strategy based on the *Nash bargaining solution*. Another advantage is that it allows multi-objective optimization with multiple decision makers (e.g., MBs) in order to achieve accurate bit rate with good perceptual quality while maintaining a stable buffer level. Several extensions of the work are possible.**

***Index Terms*—Bit allocation, discrete cosine transform (DCT), game theory, human visual system (HVS), rate control (RC), video compression.**

I. Introduction

THE VISUAL quality of an encoded frame is related to the bits it consumes. To maintain stable quality throughout the video sequence, one needs to consider the distribution of the bit budget for each frame. Therefore, bits are precious resources that must be utilized effectively. Most rate-distortion optimization algorithms optimize a unique objective function, which is typically the perceptual quality or distortion of an entire frame. While such an approach is simple (but still computationally intensive), the same objective function may not yield the best results for the whole video frame or a sequence. This paper presents a game theory based rate allocation strategy that allows multi-objective optimization with multiple decision makers (e.g., blocks), while working under an overall constraint (e.g., a given bit budget for a frame). The proposed approach is based on cooperative game theory, under which each decision maker has its own objective function, which is its own perceptual quality measure. The solution for the cooperative game yields the optimal bit allocation that is fair to each macroblock (MB) under the give constraints. The proposed rate control algorithm has two stages. In the first stage, target bits are allocated at the frame level. In the second stage, the quantization scale for each MB is determined by using a game theoretical approach.

At the frame level, the algorithm allocates target bits to the current frame based on the coding complexity of the frame, which is the mean absolute prediction error. Since the remaining frames are unavailable, we estimate the coding complexity of the remaining frames from the encoded frames. The target number of bits of the current frame is optimized by using the current frame coding complexity as well as the estimated coding complexity of the remaining frames.

At the MB level, the algorithm uses game theory for bit allocation. We formulate the bit allocation as a bargaining problem [25]. Each MB competes for a share of resources, which are the target bits for a frame. Based on *Nash bargaining solution* (NBS) [27], we derive a cooperative optimal quantization scale for each MB. Furthermore, we incorporate the human visual system (HVS) perceptual property in the game theory framework. Initial visual quality of the game setting, which is guaranteed for each MB, is determined proportional to the perceptible distortion, which is the distortion that exceeds the just-noticeable distortion (JND) threshold [14]. Fig. 1 shows the block diagram of the proposed rate control algorithm that can be embedded in a discrete cosine transform (DCT)-based video encoder. The gray box in the diagram shows the modules of the proposed algorithm.

Game Theory is widely thought to have originated in the early 20th century, when von Neumann proved the min-max theorem [28]. Some of the most pioneering results were reported within a year, when Nobel Laureate John Nash made seminal contributions to both *cooperative* and *noncooperative* games. In [26], Nash proved the existence of a strategic equilibrium for noncooperative games (Nash Equilibrium). He also proposed that cooperative games were reducible to noncooperative games. He accomplished that by pioneering the axiomatic bargaining theory and proved the existence of the NBS for cooperative games (a notion similar to the Nash Equilibrium) [26]. One remarkable property of Game Theory is its abstractly defined mathematics and notions of optimality. In no other branch of Sciences do we find so many understandable definitions and levels of optimality [30]. Game Theory has been

Manuscript received June 8, 2004; revised October 1, 2004. This paper was recommended by Associate Editor C. Guillemot.

I. Ahmad is with the Computer Science and Engineering Department, University of Texas, Arlington, TX 76019 USA (e-mail: iahmad@cse.uta.edu).

J. Luo was with the Computer Science and Engineering Department, University of Texas, Arlington, TX 76019 USA. He is now with Thomson Corporate Research, Princeton, NJ 08540 USA.

Digital Object Identifier 10.1109/TCSVT.2005.856899

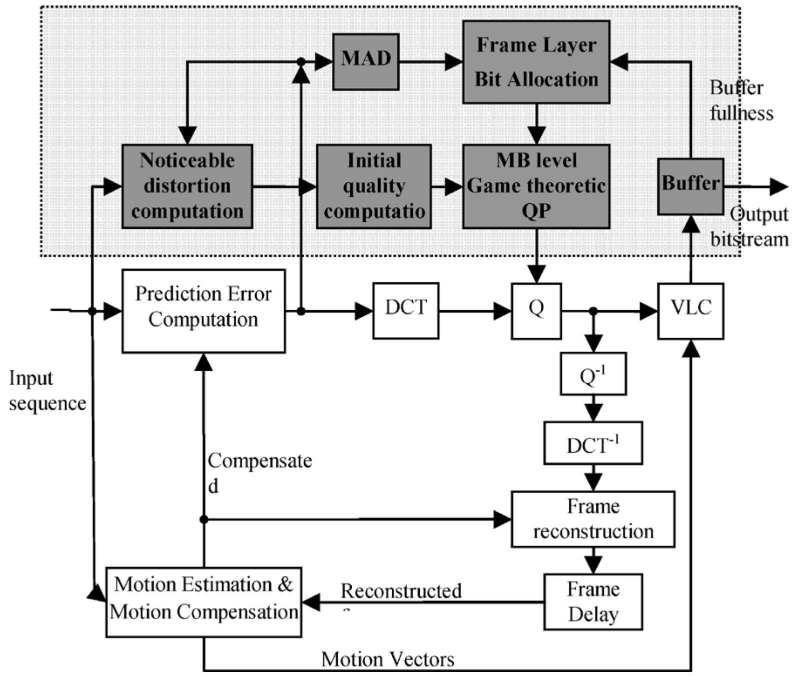


Fig. 1. Proposed algorithm embedded in a DCT-based video encoder.

used as a powerful method to analyze and solve problems that contain natural competition in several areas of social sciences [17], Biology [10], Political Science [23], Economics [22], etc. In computer science, a few applications of Game Theory applied to job scheduling and networking problems have been documented [32]. Recently, auction theory has started to be recognized as the emerging solution for problems in microeconomics [22], and agent-based systems [35]. To our knowledge, this paper is the first attempt to apply game theory in video compression.

The rest of the paper is organized as follows. Section II presents the related works on rate control reported in the literature. Section III describes the proposed frame level bit allocation scheme and the game theoretical problem formulation for MB layer quantizer optimization. Section IV presents simulation results of the proposed scheme, including a comparison to the quadratic rate control proposed in [3] and suggested by MPEG-4 VM8 [11]. Section V concludes the paper with some concluding remarks and future work.

II. Related Work

Several rate control algorithms have been proposed and utilized in video compression standards like MPEG-1/2/4 and H.261/263. In the early rate control algorithms, such as the H.263 Test Model Near-term version 5 (TMN5) [13], a fixed bit allocation for each frame is employed. The target bit budget for each frame is obtained by dividing the target bit rate by the frame rate. The MPEG Test Model 5 (TM5) [12] rate control employs a hierarchical bit allocation scheme, in which the target number of bits is decided for the GOP layer first. Constrained by the group of pictures (GOP) bit budget, target number of bits for a frame is then calculated. Target bits for I-, P-, B-frame are allocated according to the complexity of the previous frame. With respect to the model-based quantization scale decision, early rate control algorithms, such as TM5, adopt a linear rate-distortion model. Thereafter, various rate-distortion models have been proposed to improve the accuracy of the quantization scale estimation. Chiang and Zhang proposed a quadratic rate-distortion model that can be applied to both DCT and wavelet-based coders [3]. An improvement in bit allocation has been proposed by Pan *et al.* [5] by exploiting the position of a frame in a GOP. Ngan, Meier, and Chen have introduced a new constraint for the least-mean-square estimation of the model parameter of the rate-distortion function [29]. Cheng and Huang have proposed an adaptive piecewise linear model [1]. Chiang *et al.* have proposed a scheme for MB level bit allocation and two distinct models for high bit rate and low bit rate situations respectively [2]. The model for high bit rate adapts the quantization scale with the energy of the block by using finer quantization for MB of flatter image regions. The model for low bit rate maintains a near-constant quantization scales, in order to minimize the overhead bits for DQUANT, which is to define the change of quantization scale. Ribas-Corbera and Lei have proposed a rate-distortion model, which was adopted in the H.263+ testing model TMN8 [33]. Based on [33], Tsai and Hsieh have proposed to modify the encoding order of MBs to favor the more complex MBs [40]. For object-based video coding, Lee, Chiang and Zhang have proposed an algorithm that is scalable for various bit rates, spatial and temporal resolution [18], [19]. Vetro and Sun also developed a scheme for multiple video objects in [40], [42]. Lee *et al.* has proposed models for coded frames and objects as well as skipped frame and objects [20]. Other rate control algorithms have been proposed in [16], [21], [39], [44], etc. The above rate control algorithms reportedly achieve efficient bit rate regulation. What distinguishes our approach from these works is that we define a new criterion of optimizing the rate-distortion efficiency under the bit rate and "fairness" constraints, and propose an efficient solution.

A few algorithms have considered the rate-distortion optimization. Since the quantization scale can refer the distortion of the decoded frame, some rate control algorithms seek optimal quantization to maximize the aggregate perceptual quality or minimize the aggregate distortion subject to the bit rate constraints, using Lagrange multiplier or dynamic programming [31], [33], [36], [43]. He and Mitra have modeled the rate and distortion as the functions of ρ, which is the percentage of zeros among the quantized DCT coefficients, and an optimized bit allocation has been proposed based on this model [7]–[9]. Although these algorithms address the rate-distortion optimization problem, the perceptual redundancy of human vision system has not been efficiently exploited. Moreover, the fairness of bit allocation among the MBs has not been discussed.

This paper proposes an algorithm that masks the imperceptible distortion and optimizes the perceptual quality using a game theory approach ensureing optimality and fairness of bit allocation.

III. Proposed Algorithm

A. Frame-Level Bit Allocation

Since it takes more bits to encode a complex frame than a simple frame to obtain the same visual quality, we tune the frame level target bit budget according to the estimated coding complexity of the frame. We measure the coding complexity of a frame using the mean absolute prediction error, denoted by mad:

$$\mathrm{mad} = \frac{1}{W \cdot H} \sum_{x=0}^{W} \sum_{y=0}^{H} |p(x,y) - \hat{p}(x,y)| \qquad (1)$$

where (x,y) is the pixel coordinate, and W and H are the width and height of the frame in pixel. The target bit budget is allocated proportional to the coding complexity of the current coding frame. Therefore, we need to know the coding complexity of the remaining frames. Since the encoder only contain the current frame and the reference frame due to the memory and delay constraints, we estimate the coding complexity of the remaining frame by computing the weighted mean of the coding complexity of the previous coded frames. Suppose we are encoding a frame at time t. The estimated coding complexity of the remaining frames is denoted by $\mathrm{mad\_}r_t$. and given by

$$\begin{aligned}\mathrm{mad\_}r_t &= \frac{\sum_{i=1}^{t}(i \cdot \mathrm{mad}_i)}{\sum_{i=1}^{t} i} = \frac{2 \cdot \sum_{i=1}^{t}(i \cdot \mathrm{mad}_i)}{(t+1)\cdot t} \\ &= \frac{2 \cdot \sum_{i=1}^{t-1}(i \cdot \mathrm{mad}_i) + 2t \cdot \mathrm{mad}_t}{(t+1)\cdot t} \\ &= \frac{(t-1)}{(t+1)} \cdot \frac{2 \cdot \sum_{i=1}^{t-1}(i \cdot \mathrm{mad}_i)}{t \cdot (t-1)} + \frac{2}{t+1}\mathrm{mad}_t \\ &= \frac{(t-1)}{(t+1)} \cdot \mathrm{mad\_}r_{t-1} + \frac{2}{t+1}\mathrm{mad}_t \end{aligned} \qquad (2)$$

where mad_i is the coding complexity of frame at time i. The number of bits allocated to a frame at time t is given by

$$T'_t = \frac{R_{t-1}}{N_{t-1}} \cdot \times \frac{\mathrm{mad}_t}{\mathrm{mad\_}r_t} \qquad (3)$$

where R_{t-1} and N_{t-1} are the numbers of remaining bits and remaining P-frames after encoding the frame at time $t-1$.

Initially, I-frame at time 0 is encoded using an external input quantization scale. The number of the remaining bits after encoding the first I-frame is

$$R_0 = Br \cdot L - R_f \qquad (4)$$

where Br is the target bit rate of the sequence, L is the total length of the sequence in unit of second. R_f is the number of bits used to encode the first I-frame.

We need to further adjust the target bits T'_t to the current buffer fullness with the following equation:

$$T''_t = \frac{B_t + 2 \times (B - B_t)}{2 \times B_t + (B - B_t)} \times T'_t \qquad (5)$$

where B is the buffer size and B_t is the buffer fullness at time t. Equation (5) intends to maintain the buffer fullness in the middle of the buffer. If the buffer fullness is lower than the middle level, more bits will be allocated to the current frame. Otherwise, fewer bits will be allocated.

Finally, the target number of bits is bounded by

$$T_t = \min\left(\frac{2Br}{F}, \max\left(\frac{Br}{4F}, T''_t\right)\right) \qquad (6)$$

where F is the frame rate. Br/F is the average bits for a frame. Equation (6) intends to avoid allocating extremely large or extremely small number of target bits to a frame, in order to prevent the buffer overflow or under flow. The upper bound of target number of bits for a frame is two times of average bits for a frame and the lower bound is a quarter of the average bits for a frame.

B. MB-Level Bit Allocation

The problem to identify the optimal quantization scale is equivalent to find an optimal allocation of the frame target bits to maximize the perceptual quality, which is a resource optimization problem. Each MB competes for a share of resources to optimize its own performance. Since the whole frame is an entity perceived by the viewers, MBs need to work cooperatively. We solve the problem by playing a multiple-player cooperative game. An optimal and fair bit allocation strategy is derived based on the NBS.

C. Quadratic Rate Control Model

We use a quadratic rate-distortion model to formulate the relation of the consumed bits and the quantization scale used to encode a MB. The proposed model is plugged in to the game theoretical framework presented in Section III-D, to determine the optimal quantization scale. The model is given by

$$\frac{r_i}{m_i^{\alpha}} = \frac{K}{{Q_i}^2}$$

where r_i is the number of bits spend to encode the ith MB, m_i is the standard deviation of prediction errors of the ith MB,

Q_i is the quantization step size for the ith MB, K are model parameters. α is a constant ($\alpha = 0.8$).

The rate-distortion model is updated after the encoding of each MB. The rate-distortion model parameter K for the $(i+1)$th MB will be updated as the following:

$$K_{i+1} = \begin{cases} \frac{(i-n_s-1)}{i-n_s}K'_i + \frac{A_i Q_i^2}{m_i^\alpha (i-n_s)}, & A_i > 0 \\ K_i \times 0.9, & A_i = 0 \end{cases}$$
$$K'_{i+1} = \begin{cases} K_i, & A_i > 0 \\ K'_i, & A_i = 0 \end{cases} \quad (7)$$

where A_i is the actual bits used to encode the ith MB, n_s is the number of skipped MBs. The initial value K_0 is set to 3000.

D. Solving Quantizer Optimization With Game Theory

In the proposed MB level bit allocation, the bargaining game is configured as follows.

Players: Each frame contains N uncoded MBs. Each MB is regarded as a player in the game. N players compete for the use of a fixed resource, which is the target bit budget for the frame.

Strategies: The strategy of a player is the number of bits it requests for, denoted by r_i. Since the target number of bits for a frame is constrained, the sum of the bits requested by the N remaining MBs should be no more than the remaining bits for a frame, i.e.,

$$\sum_{i=1}^{N} r_i \le R_c \quad (8)$$

where R_c is the remaining bits for the N MBs.

Preference: A utility function u_i for each player i reflects its preference. We use the visual quality as a measure of utility. Higher visual quality is more preferable. Given a combination of strategies carried out by all the MBs $r = (r_1, r_2, \ldots, r_N)$, $u = (u_1(r), u_2(r), \ldots, u_N(r))$ is the utility of the game. Since in the DCT video coding, the visual quality of a MB is related only to the number of bits it obtained, the utility can be represented by $u = (u_1(r_1), u_2(r_2), \ldots, u_N(r_N))$.

Initial Utility: The initial utility of the ith MB, denoted by d_i, is the initial perceptual quality that required to be guaranteed. The initial quality is determined according to the perceptible distortion of each MB, which is defined in Section III-E. Denote the initial quality of the game $d = (d_1, d_2, \ldots, d_N)$, we have $u > d$. Define the number of bits achieving d_i as r_i^0. Since $u > d$, we have $r > r^0$, where $r^0 = ({r_1}^0, {r_2}^0, \ldots, {r_N}^0)$, which means

$$r_i > r_i^0. \quad (9)$$

Define U the set of achievable utilities. Tuple $\langle U,\ d\rangle$ represents the game setting.

NBS is a unique solution that satisfies a set of axioms for fair bargain [27]. Let strategy r^* be the NBS of the game $\langle U,\ d\rangle$, the corresponding utility $u^*(\langle U,\ d\rangle)$ is called the Nash bargaining point. The following axioms define a unique NBS:

Efficiency: The NBS is Pareto optimal, i.e., there is no other solution produces better utility for one player without hurting another player.

Linearity: Given a monotonous increasing linear transform function F, Nash bargaining point u^* satisfies that: $u^*(\langle F(U), F(d)\rangle) = F(u^*(\langle U, d\rangle)$.

Independence of Irrelevant Alternatives: Let X, Y are sets of attainable utilities, and $X \subseteq Y$. If $u^*(\langle Y, d\rangle) \in X$, then $u^*(\langle Y, d\rangle) = u^*(\langle X, d\rangle)$.

Symmetry: If U is symmetric with respect to any two players in the game, and if their initial utility is equally preferable, then exchange the two players will not affect the solution.

The efficiency axiom states that the NBS is cooperatively optimal. The linearity axiom implies that the NBS will not change if the player's objective is linearly transformed. The irrelevant alternatives axiom expressed that if we remove the irrelevant subset (the subset does not contain the Nash bargaining point), the NBS of the game will not change. The symmetry axiom says that the solution is only depended on the players' initial utilities and their utility functions. The last three axioms are the axioms of fairness.

NBS for a multiple players bargaining game is characterized by the following property [37].

The solution r^* is a NBS if and only if $\prod_i(u_i(r^*) - d_i) \ge \prod_i(u_i(r) - d_i)$ for all $r \in H$, where H is the set of all feasible combination of strategies.

Therefore, to find the NBS, we need to solve the following maximization problem:

$$\max \prod_{i=1}^{N}(u_i(r_i) - d_i) \qquad \text{s.t. } r_i > r_i^0,\ \sum_{i=1}^{N} r_i \le R_c. \quad (10)$$

The conditions in (10) are constraints from (8) and (9).

The approximate mean square error (MSE) distortion of the ith MB is $D_i = Q^2/12$ [6]. Therefore, we defined the visual quality of the ith MB by

$$u_i = \frac{1}{D_i} = \frac{12}{Q_i^2} = \frac{12 r_i}{K m_i^\alpha}. \quad (11)$$

Since u_i is concave and injective, $\ln(u_i)$ is strictly concave. The above problem is equivalent with the following problem:

$$\max \sum_{i=1}^{N} \ln(u_i(r_i) - d_i) \qquad \text{s.t. } r_i > r_i^0,\ \sum_{i=1}^{N} r_i \le R_c. \quad (12)$$

The above inequality constrained optimization problem can be solved by maximizing the following Lagrangean, using the theorem of Kuhn and Tucker [38]

$$J = \sum_{i=1}^{N} \ln(u_i - d_i) + \lambda\left(R_c - \sum_{i=1}^{N} r_i\right) + \sum_{i=1}^{N} \theta_i (r_i - r_i^0) \quad (13)$$

where λ and θ_i are the Lagrange multiplier. The optimized solution can be obtained by solving (14)

$$\begin{cases} \frac{\partial J}{\partial r_i} = & \frac{\partial \ln(u_i - d_i)}{\partial r_i} - \lambda + \theta_i = 0 \\ \frac{\partial J}{\partial \lambda} = & R_c - \sum_{i=1}^{N} r_i \geq 0 \\ \lambda \frac{\partial J}{\partial \lambda} = & \lambda \left(R_c - \sum_{i=1}^{N} r_i \right) = 0 \\ \frac{\partial J}{\partial \theta_i} = & r_i - r_i^0 \geq 0 \\ \theta_i \frac{\partial J}{\partial \theta_i} = & \theta_i (r_i - r_i^0) = 0, \end{cases}$$

where

$$i \in (1 \ldots N). \tag{14}$$

Since $\sum_{i=1}^{N} r_i^0 < R_c$, there must be a solution r that strictly superior to r^0, so that $r_i - r_i^0 > 0$. Therefore, $\theta_i = 0$. From (14), we have

$$\frac{\partial \ln(u_i - d_i)}{\partial r_i} = \lambda \Rightarrow r_i = \frac{1}{\lambda} + \frac{k}{12} m_i^{\alpha} d_i \tag{15}$$

and

$$R_c - \sum_{i=1}^{N} r_i = 0 \Rightarrow R_c - \frac{N}{\lambda} - \frac{K}{12} \sum_{i=1}^{N} m_i^{\alpha} d_i = 0. \tag{16}$$

Therefore, the NBS for the game is given by

$$r_i = \frac{R_c}{N} - \frac{K}{12N} \sum_{j=1}^{N} m_j^{\alpha} d_j + \frac{K}{12} m_i^{\alpha} d_i. \tag{17}$$

And the optimal quantization step size is given by

$$Q_i = \sqrt{\frac{K m_i^{\alpha}}{\left(\frac{R_c}{N} - \frac{K}{12N} \sum_{j=1}^{N} m_j^{\alpha} d_j + \frac{K m_i^{\alpha} d_i}{12} \right)}}. \tag{18}$$

Proof of Fairness: An assignment of resource $s = (s_1, s_2, \ldots, s_N)$ is said to be proportionally fair with respect to a utility function f, if for any other feasible allocation s', the aggregate proportional changes is zero or negative [15]

$$\sum_{i=1}^{N} \frac{f(s') - f(s_i)}{f(s_i)} \leq 0. \tag{19}$$

Denote r^* the allocation from NBS. Consider a small change of r^* to r', where $r_i' = r_i^* + \Delta r_i^*$. In the context of our game setting, the proportional fairness criteria can be rewritten as

$$\sum_{i=1}^{N} \frac{[u_i(r_i') - d_i] - [u_i(r_i^*) - d_i]}{u_i(r_i^*) - d_i} \leq 0.$$

Replace u with (11), we have

$$\sum_{i=1}^{N} \frac{[u_i(r_i') - d_i] - [u_i(r_i^*) - d_i]}{u_i(r_i^*) - d_i} \leq 0$$
$$\Leftrightarrow \sum_{i=1}^{N} \frac{\left[\frac{12 r_i}{K m_i^{\alpha}} - d_i \right] - \left[\frac{12 r_i^*}{K m_i^{\alpha}} - d_i \right]}{\frac{12 r_i^*}{K m_i^{\alpha}} - d_i} \leq 0$$
$$\Leftrightarrow \sum_{i=1}^{N} \frac{\frac{12 r_i}{K m_i^{\alpha}} - \frac{12 r_i^*}{K m_i^{\alpha}}}{\frac{12 r_i^*}{K m_i^{\alpha}} - \frac{12 r_i^0}{K m_i^{\alpha}}} \leq 0 \Leftrightarrow \sum_{i=1}^{N} \frac{r_i' - r_i^*}{r_i^* - r_i^0} \leq 0. \tag{20}$$

Let $v_i(r_i) = \ln(u_i(r_i) - d_i)$. Since r^* is the optimal solution for (12), any change in r^* will make a zero or negative change in $\sum_{i=1}^{N} v_i(r_i)$. Therefore, we have

$$\sum_{i=1}^{N} [v_i(r_i') - v_i(r_i^*)] \leq 0 \tag{21}$$

$$\sum_{i=1}^{N} \left[\frac{\partial v_i}{\partial r_i} \bigg|_{r_i = r_i^*} \Delta r_i^* \right] \leq 0 \tag{22}$$

$$\sum_{i=1}^{N} \frac{u_i'(r_i) \big|_{r_i = r_i^*}}{u_i(r_i^*) - d_i} \Delta r_i^* \leq 0. \tag{23}$$

Plug (11) into (23), we have

$$\sum_{i=1}^{N} \frac{\Delta r_i^*}{r_i^* - r_i^0} \leq 0 \Rightarrow \sum_{i=1}^{N} \frac{r_i' - r_i^*}{r_i^* - r_i^0} \leq 0. \tag{24}$$

E. Perceptually Tuned Quantizer

The perceptual redundancy is inherent in video signals. It is found that the HVS is insensitive to the signals in some spatial frequencies. Moreover, the human vision is much easier to detect the luminance difference rather than the absolute intensity. And the sensitivity to the luminance contrast is depended on the average background intensity. Due to the above observation, a metric of JND is proposed in [14] to measure the perceptual lower bound of the signal distortion. JND is a threshold below which the distortion is imperceptible. The computation of JND of a pixel at (x, y) is given by [4]

$$\mathrm{JND}(x,y) = \max\{f_1(bg(x,y), mg(x,y)), f_2(bg(x,y))\} \tag{25}$$

$$f_1(bg(x,y), mg(x,y)) = mg(x,y) \cdot \alpha(bg(x,y)) + \beta(bg(x,y)) \tag{26}$$

$$f_2(bg(x,y)) = \begin{cases} T_0 \cdot \left(1 - \left(\frac{bg(x,y)}{127} \right)^{1/2} \right) + 3 & bg(x,y) \leq 127 \\ \gamma \cdot (bg(x,y) - 127) + 3 & bg(x,y) > 127 \end{cases} \tag{27}$$

$$\alpha(bg(x,y)) = bg(x,y) \times 0.0001 + 0.115 \tag{28}$$

$$\beta(bg(x,y)) = \lambda - bg(x,y) \times 0.01 \tag{29}$$

where T_0, γ and λ are 17, 3/128, and 1/2, respectively.

1	1	1	1	1
1	2	2	2	1
1	2	0	2	1
1	2	2	2	1
1	1	1	1	1

Fig. 2. Matrix B.

G_1

0	0	0	0	0
1	3	8	3	1
0	0	0	0	0
-1	-3	-8	-3	-1
0	0	0	0	0

G_2

0	0	1	0	0
0	8	3	0	0
1	3	0	-3	-1
0	0	-3	-8	0
0	0	-1	0	0

G_3

0	0	1	0	0
0	0	3	8	0
-1	-3	0	3	1
0	-8	-3	0	0
0	0	-1	0	0

G_4

0	1	0	-1	0
0	3	0	-3	0
0	8	0	-8	0
0	3	0	-3	0
0	1	0	-1	0

Fig. 3. Matrix G_k.

$bg(x,y)$ and $mg(x,y)$ represent the average background luminance and the maximum weighted average of luminance differences around the pixel at (x,y). The calculations of $bg(x,y)$ and $mg(x,y)$ are as follows:

$$bg(x,y) = \frac{1}{32}\sum_{i=1}^{5}\sum_{j=1}^{5} p(x-3+i,\ y-3+j)\cdot B(i,j) \quad \text{for } 0 \le x < H, 0 \le y < W \tag{30}$$

where $p(x,y)$ is the luminance of pixel at (x,y). $B(i,j)$, $(i,j = 1,\dots,5)$, is a matrix of weights, which is shown in Fig. 2. mg is the maximum gradient among different directions. grad_k represents the gradient in direction k, where $k = \{1,2,3,4\}$ is one of the following four directions: 1) vertical; 2) diagonal (upper-left to lower-right); 3) horizontal; 4) diagonal (upper-right to lower-left). grad_k is computed with (32)

$$mg(x,y) = \max\left\{\left|\text{grad}_k(x,y)\right|\right\} \tag{31}$$

$$\text{grad}_k(x,y) = \frac{1}{16}\sum_{i=1}^{5}\sum_{j=1}^{5} p(x-3+i, y-3+j)\cdot G_k(i,j) \tag{32}$$

where G_k is the matrices of weights in four directions, which is given in Fig. 3.

In (18), a factor that affects the optimal quantization step size Q_i is the initial quality d_i. In the proposed algorithm, the initial quality of a MB is proportional to its noticeable distortion e. The noticeable distortion in the ith MB is given by

$$e_i = \frac{1}{256}\sum_{m=1}^{16}\sum_{n=1}^{16}\{[h(m+u,n+v) - \text{JND}(m+u,n+v)]\cdot\delta(m+u,n+v)\} \tag{33}$$

where (u,v) is the top-left corner of the ith MB, and

$$h(x,y) = \left|p(x,y) - \hat{P}(x,y)\right| \tag{34}$$

$$\delta(x,y) = \begin{cases} 1, & h(x,y) > \text{JND}(x,y) \\ 0, & h(x,y) \le \text{JND}(x,y). \end{cases} \tag{35}$$

Due to the bit rate constraint, the total bits corresponding to the initial quality of the N uncoded MBs in a frame cannot exceed the remaining bits, i.e., $\sum_{i=1}^{N} r_i^0 < R_c$. Therefore, we bound the initial quality by a scale factor C, such that the total bits corresponding to the initial quality are bounded by a half of the remaining bits for the N uncoded MBs ($0.5R_c$). C is computed as follows:

$$C = \begin{cases} 1 & R_c \ge \frac{K}{6}\sum_{j=1}^{N} m_j^\alpha e_j \\ \frac{6R_c}{K\cdot\sum_{j=1}^{N} m_j^\alpha e_j} & R_c < \frac{K}{6}\sum_{j=1}^{N} m_j^\alpha e_j. \end{cases} \tag{36}$$

The initial quality for the ith MB is given

$$d_i = C\cdot e_i. \tag{37}$$

F. Algorithm Summary

The proposed algorithm, name Game Theoretic (GT) algorithm, is embedded in a DCT-based video encoder (e.g., H.263 or MPEG-4), as shown in Fig. 1. The functionality of each block in the diagram is introduced in the previous sections. We summarize the GT algorithm in the following steps.

- Step 1) Frame layer bit allocation will be performed prior to the encoding of a frame. The mean absolute prediction error of the current frame is computed with **(2)** and **(3)**, in order to determine the initial estimation of the target number of bits T_t'. Buffer fullness is fed back to further tune T_t' with **(5)**. Then T_t'' is further adjusted with **(6)** to get T_t.
- Step 2) Before encoding a frame, R_c is initially set to T_t. The noticeable distortion for each MB in the frame is calculated with **(33)**.
- Step 3) Based on the noticeable distortion and remaining bits R_c, the initial qualities for the remaining MBs are computed with **(36)** and **(37)**.
- Step 4) The quantization step size of current MB is computed based on the derived game theoretical formulation, given by **(18)**.
- Step 5) After encoding each MB, the quadratic rate-distortion model is updated with **(7)**. R_c is updated

TABLE I

BIT RATE ACCURACY, PSNR, PSPNR, AND NUMBER OF SKIPPED FRAME COMPARISON BETWEEN VM8 AND THE GT ALGORITHM

a) QCIF sequences

video sequence	Algorithm	Bit rate Target	Bit rate Actual	bit rate accuracy	Bits saved	Average PSNR_Y	Average PSPNR_Y	Frame skipped	PSNR gain (dB)	PSPNR gain (dB)
table	VM8	64	66.64	95.88%		28.88	32.38	3		
	GT	64	64.1	99.84%	3.96%	29.04	32.6	0	0.15	0.22
	VM8	96	99.24	96.63%		30.26	34.36	3		
	GT	96	96	100.00%	3.37%	30.34	34.48	0	0.08	0.12
	VM8	128	131.54	97.23%		31.2	35.76	2		
	GT	128	128	100.00%	2.77%	31.34	35.98	0	0.15	0.22
stefan	VM8	96	101.54	94.23%		24.45	26.87	5		
	GT	96	96.12	99.88%	5.64%	24.51	26.94	0	0.06	0.07
	VM8	112	119.81	93.03%		24.86	27.39	6		
	GT	112	111.98	99.98%	6.99%	25.03	27.61	0	0.17	0.22
	VM8	128	135.9	93.83%		25.34	28.03	5		
	GT	128	128.05	99.96%	6.13%	25.46	28.18	0	0.11	0.14
foreman	VM8	64	69.7	91.09%		29.73	33.56	8		
	GT	64	64.04	99.94%	8.84%	29.84	33.7	0	0.11	0.14
	VM8	128	131.99	96.88%		32.53	37.93	2		
	GT	128	127.94	99.95%	3.17%	32.62	38.23	0	0.09	0.31
	VM8	192	195.27	98.30%		34.27	40.98	1		
	GT	192	191.79	99.89%	1.81%	34.32	41.29	0	0.05	0.31
container	VM8	192	199.06	96.32%		38.15	47.46	3		
	GT	192	192.09	99.95%	3.63%	38.48	47.81	0	0.33	0.36
	VM8	384	386.87	99.25%		41.15	54.62	0		
	GT	384	384.12	99.97%	0.72%	41.95	56.3	0	0.79	1.68
	VM8	512	512.01	100.00%		41.96	56.84	0		
	GT	512	512.2	99.96%	-0.04%	42.99	58.78	0	1.03	1.93
akiyo	VM8	192	198.22	96.76%		42.31	55.97	2		
	GT	192	191.88	99.94%	3.30%	42.42	58.05	0	0.11	2.07
	VM8	384	384.91	99.76%		44.16	60.62	0		
	GT	384	383.41	99.85%	0.39%	44.72	62.85	0	0.56	2.23
	VM8	512	511.99	100.00%		45.39	64.18	0		
	GT	512	511.32	99.87%	0.13%	46.25	67.71	0	0.86	3.53

b) CIF sequences

video sequence	Algorithm	Bit rate Target	Bit rate Actual	bit rate accuracy	bits saved	Average PSNR_Y	Average PSPNR_Y	Frame skipped	PSNR gain (dB)	PSPNR gain (dB)
container	VM8	192	193.32	99.31%		32.54	37.33	0		
	GT	192	191.82	99.91%	0.78%	32.57	37.42	0	0.03	0.09
	VM8	256	263.35	97.13%		33.37	38.64	2		
	GT	256	256.12	99.95%	2.82%	33.44	38.80	0	0.06	0.16
	VM8	384	400.84	95.61%		34.86	41.26	4		
	GT	384	384.32	99.92%	4.30%	34.90	41.32	0	0.04	0.05
Coastguard	VM8	256	270.36	94.39%		27.40	30.86	5		
	GT	256	256.04	99.98%	5.59%	27.43	30.88	0	0.03	0.02
	VM8	384	395.19	97.09%		28.69	32.67	2		
	GT	384	383.95	99.99%	2.93%	28.79	32.77	0	0.10	0.10
	VM8	512	521.55	98.13%		29.73	34.10	1		
	GT	512	511.88	99.98%	1.89%	29.78	34.15	0	0.05	0.05
foreman	VM8	256	271.10	94.10%		32.05	36.42	5		
	GT	256	255.82	99.93%	5.97%	32.35	36.97	0	0.30	0.55
	VM8	384	398.17	96.31%		33.57	38.98	3		
	GT	384	383.48	99.86%	3.83%	33.69	39.18	0	0.12	0.21
	VM8	512	522.32	97.98%		34.59	40.79	1		
	GT	512	511.41	99.89%	2.13%	34.64	40.81	0	0.05	0.02
table	VM8	384	397.00	96.61%		30.80	34.89	3		
	GT	384	383.77	99.94%	3.45%	31.08	35.29	0	0.27	0.40
	VM8	512	522.09	98.03%		31.93	36.56	1		
	GT	512	511.99	100.00%	1.97%	32.07	36.80	0	0.14	0.24
	VM8	640	648.14	98.73%		32.75	37.85	1		
	GT	640	640.03	100.00%	1.27%	32.88	38.04	0	0.12	0.20
Mobile	VM8	512	548.51	92.87%		24.10	27.01	6		
	GT	512	512.33	99.94%	7.06%	24.17	27.10	0	0.07	0.09
	VM8	640	674.20	94.66%		24.88	28.07	4		
	GT	640	640.19	99.97%	5.31%	24.95	28.16	0	0.07	0.09
	VM8	768	800.36	95.79%		25.56	29.01	3		
	GT	768	768.16	99.98%	4.19%	25.62	29.09	0	0.06	0.08

as $R_c = R_c - A_i$. Go to step 6 if all the MBs in the frame is finished, otherwise go to step 3.

Step 6) Update the buffer status after encoding each frame. Stop if all the frames are encoded, otherwise go to step 1.

IV. Experiments and Results

First, we compare the GT algorithm with the quadratic rate control algorithm suggested by MPEG-4 VM8 [11]. Both comparing algorithms are implemented in Momusys encoder for MPEG-4 Verification Model. To be consistent with the VM8 rate control algorithm, we set the buffer size to $Br/8$, which means the maximum delay is 125 ms. The initial buffer fullness is $Br/16$.

The performance of a rate control algorithm is evaluated by the following metrics:

1) bit rate accuracy;
2) percentage of bits saved;
3) number of skipped frames;
4) peak-signal-to-noise ratio (PSNR);
5) peak-signal-to-perceptible-noise ratio (PSPNR).

A good rate control algorithm should be able to control the actual bit rate as close as possible to the target bit rate. We measure the bit rate accuracy with the following equation:

$$\text{bit\_rate\_accuracy} = 1 - \frac{|R_{\text{actual}} - R_{\text{target}}|}{R_{\text{target}}} \tag{38}$$

where R_{actual} and R_{target} are the actual bit rate and the target bit rate respectively. Besides the bit rate accuracy, we also measure the percentage of bits saved of the GT algorithm compared to the VM8 algorithm. For the same visual quality, lower bit consumption means the higher rate-distortion efficiency.

Frame skip technique is employed to avoid the buffer overflow. Once the buffer fullness is over a threshold, the encoding of the next frame will be skipped and will not be buffer in order to cut down the buffer fullness level. In the decoder side, the skipped frame will be replaced by a duplication of the previous frame in order to maintain the continuity of the video decoding. However, a frame skip will degrade the signal quality. A good rate control algorithm should be able to avoid the buffer overflow and minimize the number of skipped frames.

PSNR is a widely adopted metric to measure visual quality. PSNR averages the noise of all pixels in a frame, regardless if it is perceptible to HVS or not. PSPNR proposed by Chou and Li [4] measures the perceptible visual quality incorporating the human perceptual property. PSPNR only take account of the perceptible noise and is defined as

$$\text{PSPNR} = 20\log_{10}\frac{255}{\sqrt{\frac{1}{HW}\sum_{x=0}^{H}\sum_{y=0}^{W}(h(x,y) - \text{JND}(x,y))^2 \cdot \delta(x,y)}} \tag{39}$$

where $h(x,y)$ and $\delta(x,y)$ and is defined in (34) and (35), respectively. According to the MPEG-4 core experiment, the PSNR (and PSPNR) of the skipped frame is computed by considering the skipped frame as the duplication of the previous decoded frame in the decoded sequence [33], [34].

We encode QCIF and CIF format test sequences at various target bit rates. The frame rate in the experiments is 30 frames per second. For each sequence, 100 frames are encoded. The temporal prediction structure used in the experiment is IPPP…, i.e., only the first frame is encoded as I-frame and the remaining frames are encoded as P-frame (IPP…). It is to be noted that the GT algorithm is not limited to P-frames, and easily extended to the frame level bit allocation for I-frame and B-frame. The reason of using this temporal prediction structure in the experiment is that the comparative algorithm (VM8) supports only this structure. To be consistent with VM8 in the comparison, we adopt IPPP… structure in the experiment. The detailed simulation results and the comparisons are shown in Table I. Table I(a) shows the results of QCIF format and Table I(b) shows the results of CIF format.

A. Bit Rate Accuracy and Percentage of Bits Saved

From Table I, we observe that the GT rate control algorithm produces fewer bits and achieves more accurate bit rate than the VM8 algorithm in most test cases. The average bit rate accuracy of VM8 is 96.45% for CIF format and 96.61 for QCIF format while the GT algorithm achieves 99.95% for CIF format and 99.93% for QCIF format. Compare to VM8, the GT algorithm saves 3.57% and 3.39% bits for CIF format and QCIF format, respectively.

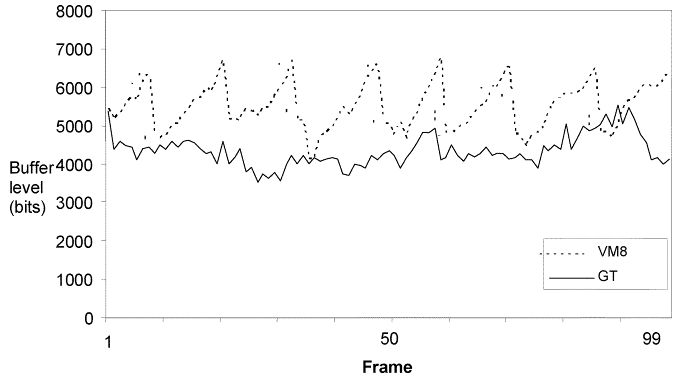
One can note that the GT algorithm uses fewer bits in the situation that without frame skipped. That means, comparing with VM8 that skips a certain number of frames to achieve the target bit rate, the proposed algorithm averagely use even less bits to encode a frame.

B. Frame Skipping

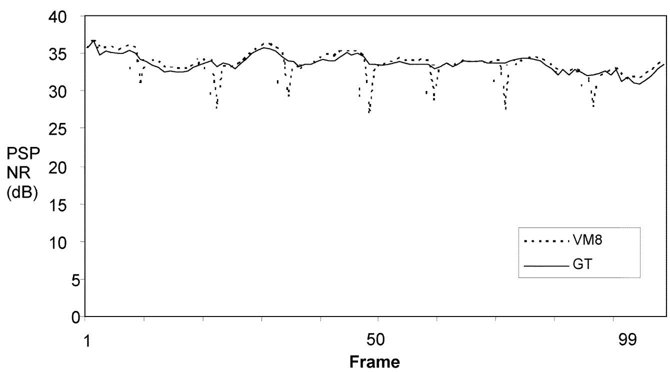
On the number of skipped frames comparison, one can notice that various numbers of frames are skipped when using VM8 algorithm. On the contrast, there is no frame skips in the experiments when using the GT rate control algorithm. The advantage of the GT algorithm is especially prominent in the low bit rate tests. For example, in the "Table" QCIF 64 kbps test and "Stefan" QCIF 96 kbps test, VM8 skips 3 frames and 5 frames, respectively, while the GT algorithm does not skip any frame.

The reason of improvement of frame skipping is because the GT can maintain a stable buffer level, which thanks to the accurate estimation of frame level target bit budge and the high accuracy of MB level rate control.

Figs. 4(a), 5(a), and 6(a) illustrate the frame-to-frame comparison on buffer fullness. One can observe that the proposed algorithm has less fluctuation on buffer level and is able to control the buffer fullness around the middle of the buffer size. The buffer fullness level is maintained within a safe margin to avoid frame skipping.

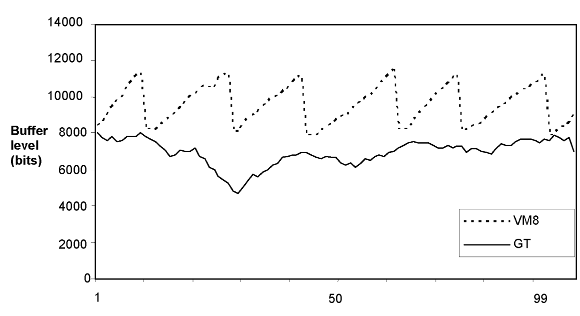


(a) Buffer fullness (buffer size = 8kbits)

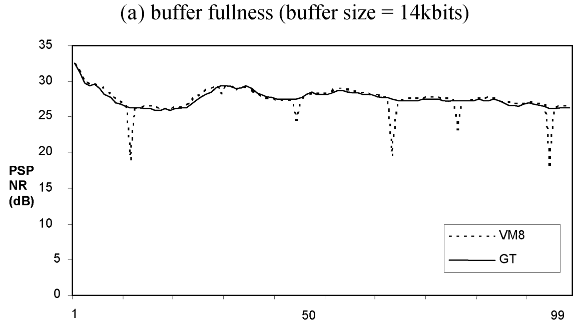


(b) PSPNR

Fig. 4. Frame-to-frame comparison of buffer occupancy and PSPNR. ("Foreman" QCIF, 64 kbps, 30 fps).

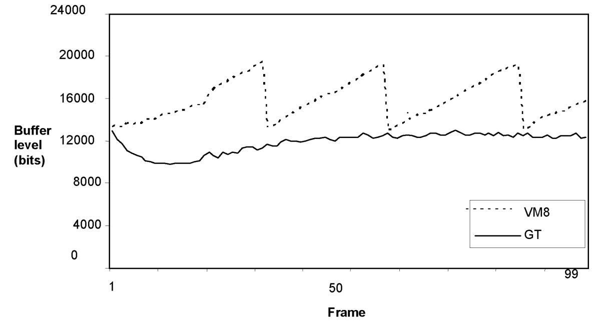


(a) buffer fullness (buffer size = 14kbits)

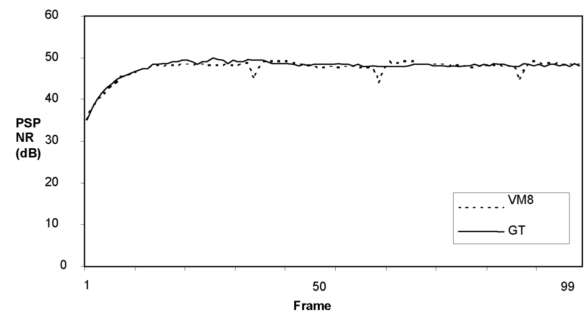


(b) PSPNR

Fig. 5. Frame-to-frame comparison of buffer occupancy and PSPNR. ("Stefan" QCIF, 112 kbps, 30 fps).



(a) buffer fullness (buffer size = 24kbits)



(b) PSPNR

Fig. 6. Frame-to-frame comparison of buffer occupancy and PSPNR. ("Container," 192 kbps, 30 fps).

C. PSNR and PSPNR

The overall visual quality is measure by the average PSNR and PSPNR. Different rate control algorithm may lead to different number of frame to be skipped. If a rate control algorithm skips more frames than the other one, more bits are used to encode the nonskipped frame, which will lead to a higher PSNR and PSPNR value on the nonskipped frame. Therefore, it is unfair to compare the average PSNR and PSPNR only taking into account the encoded frame. The distortion of the skipped frame should be considered to for a fair comparison. In the MPEG-4 decoder, a skipped frame will be repeated by the previous encoded frame. Hence, in the MPEG-4 rate control test the previous frame is used in the PSNR and PSPNR computation when a frame is skipped. [33]

Table I shows the PSNR and PSPNR comparison between VM8 and the proposed algorithm. It can be observed that the GT algorithm achieves higher PSNR than its counterpart in the PSNR comparison, although the GT algorithm is optimized for the PSPNR.

With regard to the perceptible distortion comparison, the GT algorithm successfully masks the imperceptible distortion. This is due to the bit allocation based on the noticeable distortion. Therefore, the algorithm produces substantial improvement in perceptual quality, compared to the VM8 algorithm. For example, the PSPNR improves by 1.68 dB in the "Container" QCIF 384 kbps test, and by 2.23 dB in the "Akiyo" QCIF 384 kbps test. The GT algorithm outperforms the VM8 algorithm in terms of both the bit rate and the visual quality.

Figs. 4–6 show detailed results of buffer fullness level and PSPNR values of each frame achieved by the two algorithms. One can observe that the GT algorithm achieves higher PSPNR and maintains the buffer fullness level within a safe margin to avoid frame skipping.

V. Conclusion

This paper proposed a rate control algorithm using a game theoretical approach. The algorithm models the bit allocation problem on the MB level as a bargaining problem. Bit allocation and quantization scale of each MB are decided based on the NBS. The proposed algorithm masks the imperceptible distortions by adjusting the initial quality of each MB based on noticeable distortion. The algorithm includes an efficient frame level bit allocation according to the frame coding complexity. The proposed algorithm outperforms the VM8 rate control algorithm in terms of several aspects, including bit rate accuracy, PSNR, PSPNR, and the buffer stability.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable suggestions and comments to improve the paper.

References

[1] J. Cheng and H. Hang, "Adaptive piecewise linear bits estimation model for MPEG based video coding," in *Proc. 1995 Int. Conf. Image Process.*, vol. 2, 1995, pp. 551–554.
[2] T. Chiang, H. J. Lee, and Y. Q. Zhang, "Macroblock layer rate control,", ISO/IEC JTC/SC29/WG11 MPEG97/M2555, 1997.
[3] T. Chiang and Y. Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 246–250, Feb. 1997.
[4] C. H. Chou and Y. C. Li, "A perceptually tuned subband image codec based on the measure of just-noticeable-distortion profile," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, no. 6, pp. 467–476, Dec. 1995.
[5] F. Pan, Z. Li, K. Lim, and G. Feng, "A study of MPEG-4 rate control scheme and its improvements," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 5, pp. 440–446, May 2003.
[6] A. Gersho and R. Gray, *Vector Quantization and Signal Compression*. Norwell, MA: Kluwer, 1992.
[7] Z. He, Y. K. Kin, and S. K. Mitra, "Low-delay rate control for DCT video coding via ρ-domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 8, pp. 928–940, Aug. 2001.
[8] Z. He and S. K. Mitra, "Optimum bit allocation and accurate rate control for video coding via ρ-domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 10, pp. 840–849, Oct. 2002.
[9] ——, "A linear source model and a unified rate control algorithm for DCT video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 11, pp. 970–982, Nov. 2002.
[10] J. Hofbauer and K. Sigmund, *Evolutionary Games and Population Dynamics*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[11] *Coding of Moving Picture and Associated Audio*, ISO/IEC JTC1/SC29/WG11, MPEG97/W1796, ISO/IEC 14 496–2, MPEG-4 Video VM Version 8.0, 1997.
[12] "Test Model 5 (TM5)," ISO/IEC JTC1/SC29/WG11, Document N400, 1993.
[13] *Video Codec Test Model for the Near-Term 5 (TMN5)*, ITU-T SG 15 Experts Group for Very Low Bitrate Visual Telephony, 1995.
[14] N. Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of human perception," *Proc. IEEE*, vol. 81, no. 10, pp. 1385–1422, Oct. 1993.
[15] F. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecomm. Focus Elastic Service over ATM Netw.*, vol. 8, no. 1, pp. 33–37, 1997.
[16] H. M. Kim, "Adaptive rate control using nonlinear regression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 1, pp. 432–439, Jan. 2003.
[17] V. Krishna, *Auction Theory*. San Diego, CA: Academic, 2002.
[18] H. J. Lee, T. Chiang, and Y. Q. Zhang, "Scalable rate control for very low bit rate (VLBR) video," in *Proc. 1997 Int. Conf. Image Process.*, vol. 2, Oct. 1997, pp. 768–771.
[19] ——, "Scalable rate control for MPEG-4 video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, no. 7, pp. 878–894, Sep. 2000.
[20] J. W. Lee, A. Vetro, Y. Wang, and Y. S. Ho, "Bit allocation for MPEG-4 video coding with spatio-temporal tradeoffs," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 6, pp. 488–502, Jun. 2003.
[21] L. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 4, pp. 446–459, Aug. 1998.
[22] C. Montet and D. Serra, *Game Theory and Economics, Houndmills, Basingstoke, Hampshire*. New York: Palgrave, 2003.
[23] J. Morris, M. Satyanarayanan, M. Conner, J. Howard, D. Rosenthal, and F. Smith, "Andrew: A distributed personal computing environment," *Commun. ACM*, vol. 29, no. 3, pp. 184–201, Mar. 1986.
[24] J. Nash, "Equilibrium points in N-person games," in *Proc. Nat. Acad. Sci.*, vol. 36, Jan. 1950, pp. 48–49.
[25] ——, "The bargaining problem," *Econometrica*, vol. 18, pp. 155–162, Apr. 1950.
[26] ——, "Non-cooperative games," *Annal. Math.*, vol. 54, pp. 286–295, Sep. 1951.
[27] ——, "Two-person cooperative games," *Econometrica*, vol. 21, pp. 128–140, Jan. 1953.
[28] J. von Neumann, "Zur theorie der gesellschaftsspiele," *Math. Annal.*, vol. 100, pp. 295–320, 1928.
[29] K. N. Ngan, T. Meier, and Z. Chen, "Improved single video object rate control MPEG-4," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 5, pp. 385–393, May 2003.
[30] M. J. Osborne and A. Rubistein, *A Course in Game Theory*. Cambridge, MA: MIT Press, 1994.
[31] K. Ramchandran, A. Ortega, and M. Vetterli, "Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders," *IEEE Trans. Image Process.*, vol. 3, no. 5, pp. 533–545, Sep. 1994.
[32] T. Roughgarden, "Stackelberg scheduling strategies," in *Proc. ACM STOC*, 1991, pp. 104–113.
[33] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communication," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 172–185, Feb. 1999.
[34] J. I. Ronda, M. Eckert, F. Jaureguizar, and N. Garcia, "Rate control and bit allocation for MPEG-4," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 6, pp. 1243–1258, Dec. 1999.
[35] T. Sandholm, "Limitations of vickery auction in computational multiagent systems," in *Proc. 2nd Int. Conf. Multiagent Syst.*, 1996, pp. 299–306.
[36] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. Acoust. Speech, Signal Process.*, vol. 36, no. 9, pp. 1445–1453, Sep. 1988.
[37] A. Stefanescu and M. W. Stefanescu, "The arbitrated solution for multiobjective convex programming," *Rev. Roum. Math. Pure Applicat.*, vol. 29, pp. 593–598, 1984.
[38] R. K. Sundaram, *A First Course in Optimization Theory*. Cambridge, U.K.: Cambridge Univ. Press, 1996.
[39] B. Tao, H. A. Perterson, and B. W. Dickinson, "A rate-quantization model for MPEG encoders," in *Proc. 1997 Int. Conf. Image Process.*, Oct. 1997, pp. 338–341.
[40] J. C. Tsai and C. H. Hsieh, "Modified TMN8 rate control for low-delay video communication," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 6, pp. 864–868, Jun. 2004.
[41] A. Vetro and H. Sun, "Joint rate control for multiple video objects based on quadratic rate-distortion model," in *ISO/IEC JTC1/SC29/WG11 Coding of Moving Pictures and Associated Audio MPEG 97/M1631*, Sevilla, Spain, Feb. 1997.
[42] A. Vetro, H. Sun, and Y. Wang, "MPEG-4 rate control for multiple video objects," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 186–199, Feb. 1999.
[43] S. W. Wu and A. Gersho, "Rate-constrained optimal block-adaptive coding for digital tape recording of HDTV," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 1, no. 1, pp. 100–112, Mar. 1991.
[44] K. H. Yang, A. Jacquin, and N. S. Jayant, "A normalized rate-distortion model for H.263-compatibal codecs and its application to quantizer selection," in *Proc. 1997 Int. Conf. Image Process.*, Oct. 1997, pp. 41–44.

Ishfaq Ahmad (S'88–M'92–SM'03) received the B.Sc. degree in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, in 1985, the M.S. degree in computer engineering, and the Ph.D. degree in computer science, both from Syracuse University, Syracuse, NY, in 1987 and 1992, respectively.

He is currently a Full Professor of Computer Science and engineering in the Computer Science and Engineering Department, University of Texas (UT) at Arlington. Prior to joining UT Arlington, he was an associate professor in the Computer Science Department at Hong Kong University of Science and Technology (HKUST), Hong Kong. At HKUST, he was also the Director of the Multimedia Technology Research Center, an officially recognized research center that he conceived and built from scratch. The center was funded by various agencies of the Government of the Hong Kong Special Administrative Region as well as local and international industries. With more than 40 personnel including faculty members, postdoctoral fellows, full-time staff, and graduate students, the center engaged in numerous research and development projects with academia and industry from Hong Kong, China, and the U.S. Particular areas of focus in the center are video (and related audio) compression technologies, videotelephone and conferencing systems. The center commercialized several of its technologies to its industrial partners world wide. His recent research focus has been on developing parallel programming tools, scheduling and mapping algorithms for scalable architectures, heterogeneous computing systems, distributed multimedia systems, video compression techniques, and web management. His research work in these areas is published in over 125 technical papers in refereed journals and conferences.

Dr. Ahmad has received Best Paper Awards at Supercomputing'90 (New York), Supercomputing'91 (Albuquerque), and the 2001 International Conference Parallel Processing (Spain). He has participated in the organization of several international conferences and is an Associate Editor of *Cluster Computing, Journal of Parallel and Distributed Computing,* IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, *IEEE Concurrency*, and *IEEE Distributed Systems Online*.

Jiancong Luo (S'04-M'05) received the B.S. and M.S. degrees in electrical engineering from Zhongshan (Sun Yat-sen) University, Guangzhou, China, in 1997 and 2000, respectively. From 2000 to 2001, he was a Ph.D. student in computer science at Hong Kong University of Science and Technology, Hong Kong. He received the Ph.D. degree in computer science and engineering from the University of Texas at Arlington in August 2005.

He is currently a Member of Technical Staff at Thomson Corporate Research, Princeton, NJ. His research interests include video compression, wireless multimedia communication, and 2-D/3-D image and video processing.