A Review of the Minimum Maximum Criterion for Optimal Bit Allocation among Dependent Quantizers

Guido M. Schuster

3COM, Carrier Systems Business Unit

Advanced Technologies Research Center

1800 W. Central Rd.

Mount Prospect, Illinois 60056-2293

Tel: (847) 222-2486

Fax: (847) 222-0573

E-mail: guido_schuster@3com.com

Gerry Melnikov and Aggelos K. Katsaggelos

Northwestern University

Department of Electrical and Computer Engineering

McCormick School of Engineering and Applied Science

Evanston, Illinois 60208-3118

Tel: (847) 491-7164

Fax: (847) 491-4455

E-mail: {gerrym,aggk}@ece.nwu.edu

Abstract

In this paper we review a general framework for the optimal bit allocation among dependent quantizers based on the minimum maximum (MINMAX) distortion criterion. Pros and cons of this optimization criterion are discussed and compared to the well known Lagrange multiplier method for the minimum average (MINAVE) distortion criterion. We argue that, in many applications, the MINMAX criterion is more appropriate than the more popular MINAVE criterion. We discuss the algorithms for solving the optimal bit allocation problem among dependent quantizers for both criteria and highlight the similarities and differences. We point out that any problem which can be solved with the MINAVE criterion, can also be solved with the MINMAX criterion, since both approaches are based on the same assumptions. We discuss uniqueness of the MINMAX solution and the way both criteria can be applied simultaneously within the same optimization framework. Furthermore we show how the discussed MINMAX approach can be directly extended to result in the lexicographically optimal solution. Finally we apply the discussed MINMAX solution methods to still image compression, inter-mode frame compression of H.263, and shape coding applications.

I. Introduction

A compromise between the rate and the distortion is an inherent feature of every lossy compression scheme. One common approach to mathematically formulate this tradeoff is to minimize the average (or total) distortion for a given bit rate, or vice versa, to minimize the bit rate for a given average distortion. In other words, the minimum average (MINAVE) criterion is employed. The philosophy behind this approach is that if the average distortion is minimized then, in the long run, the best quality is obtained.

The well-known Lagrangian multiplier method [3], [31], [20], [23], [19] is well suited for these kinds of constrained optimization problems. It converts the "hard" constrained problem into a set of "easy" unconstrained problems, parameterized by the Lagrange multiplier λ. These unconstrained optimization problems need to be solved optimally and efficiently so that the overall scheme is optimal and efficient. In the case where the quantizers are independent, the solution can be found by solving a set of independent problems [3], [31]. In the case of dependent quantizers, these unconstrained problems can be solved optimally and efficiently by dynamic programming (DP) [20], [23]. Quantizers are called dependent when the selection of a particular quantizer depends on the selection of a neighboring (in space and/or time) quantizer. Differential pulse code modulation (DPCM) is a prime example, since a quantizer applied to the current value changes the prediction value for the next sample. This in turn changes the prediction error which is then quantized by another quantizer. Hence since the quantizer applied to the prediction error is non-linear, the quantizer applied to the current sample has a direct influence on the selection of the quantizer used for the prediction error. We will give more examples of dependent quantizers later on. The search for the optimal solution consists of first finding the optimal λ^* using, for example, the bisection method, such that the optimal solution to the unconstrained problem also solves the constrained problem optimally. It is interesting to notice that with this popular approach, a large variability among the different source distortions is possible. Since in general the sources are consecutive in time and/or space, such as, different frames in a sequence, different blocks in a frame, or different boundary segments, this variability in quality can be very disturbing and the perceived quality is low even though the average distortion is minimized.

A different approach to formalize the relationship between the rate and the distortion is the minimum maximum (MINMAX) distortion approach, where the goal is to minimize the maximum source distortion for a given bit rate, or vice versa, to minimize the bit rate for a given maximum source distortion. The philosophy behind this approach is that by minimizing the maximum source distortion, no single source distortion will be extremely bad and hence the overall quality will be quite constant. In fact, the MINMAX criterion is an ideal choice, when the goal is to achieve an almost constant distortion which is as small as possible for the available bit rate. It may also be the criterion of choice when subjective quality is taken into account. The lexicographic approach is a generalization of the MINMAX criterion for the case where the MINMAX solution is not unique. We will discuss the extension of the MINMAX approach to the lexicographic approach later on.

There are many applications where a global distortion metric, i.e., MINAVE criterion, is not a satisfactory measure of quality or acceptability. For example, to an investor whose stock went down there is little consolation in the fact that the Dow Jones index reached a new high. Similarly, in coding applications it may be meaningless, in certain cases, to rely on the global mean squared error (MSE) to judge performance. One example is given in section V-C for a boundary encoding problem. In Figure 18 the optimal MINAVE solution is found for a rate constrained problem. In this case the optimal MINAVE solution did not include the encoding of the area between the leg of the child. Note that otherwise, the boundary approximation is very good (error pixels are in white). This is the expected result for the MINAVE criterion, since encoding this long but rather small object requires many bits (long boundary) while not encoding it at all results in a relative small error (small area). Therefore it is the best trade-off in the MINAVE sense not to encode the object and instead use the saved bits to improve the boundary approximation for the other objects in the scene. Hence the MINAVE results in a small global MSE at the expense of a large localized MSE. A MINAVE solution has an inherent problem with outliers, which, though may be statistically insignificant, can play an important role as we have shown in the above example.

With the MINMAX approach, on the other hand, this problem does not exist, since the local maximum allowable distortion is explicitly bounded. This is one motivation for the recent interest in near-lossless compression techniques [22], [7], [17], where quality is maintained at the local level. Note that the MINMAX approach discussed in this paper can include the near-lossless technique used in [22], using a proper definition of distortion. Another strong argument in favor of MINMAX comes from perceptual quality considerations. It has long been known in the image coding field that the human visual system exhibits more tolerance to distortions in regions with high spatial activity than to distortions in relatively flat areas [10], [1]. If the chosen distortion metric is inversely related to a local (mean-removed) SNR, the application of the MINMAX criterion has the effect of allowing more errors in high frequency regions and fewer errors in relatively uniform areas, which is consistent with perceptual quality considerations.

The MINAVE criterion, however, is much more commonly used in the literature than the MINMAX criterion. This is mostly due to the fact that efficient algorithms have been available for the MINAVE criterion, while such algorithms have been lacking for the MINMAX criterion. The MINMAX problem for independent quantizers has been studied in [14] where a simple algorithm has been proposed. The MINMAX algorithm for the optimal bit allocation among dependent quantizers described in this paper was first published in [27] and also in chapter 4 of [23]. In [26] the connection between the MINMAX and the MINAVE algorithm is outlined while the first application of the MINMAX algorithm to boundary coding was published in [28]. The first application of the lexicographic criterion for the optimal bit allocation among independent quantizers in a continuous framework was published in [8]. The basic idea of secondary distortion measures was introduced for boundary coding in [23]. The main contribution of this review paper is the unification and generalization of the previously published results. We establish a common framework for the MINMAX and MINAVE criterion and show how the MINAVE can be extended to include the lexicographic criterion. Furthermore we generalize the concept of a secondary distortion measure first introduced for boundary coding. In addition, we show how the boundary encoding problem is an application of the general theory discussed in this paper. Furthermore we present new results of the MINMAX criterion for video compression and compare it to previously published results using the MINAVE criterion.

In section II we introduce the notation and assumptions and formulate the problem mathematically. In section III we review the Lagrangian multiplier method for dependent quantizers and cast the solution in the DP paradigm. In section IV we review an efficient algorithm for the optimal bit allocation among dependent quantizers for the MINMAX criterion. In section III we also discuss the uniqueness of the solution and show how the MINMAX approach can be extended to result in a lexicographically optimal solution. In section V we apply both algorithms (MINMAX and MINAVE) to the Intra and Inter frame encoding schemes used in H.263 and compare the different results with respect to the mean and variance of the resulting distortion. There we also present an example of applying both criteria to the shape coding problem. Finally in section VI we compare the MINMAX and the MINAVE criterion and summarize the paper.

II. Notation, Assumptions and Problem Formulation

In this section we introduce the necessary notation, the underlying assumptions and the mathematical formulation of the optimal bit allocation problem in a dependent coding framework.

Before we proceed it is important to discuss in more detail the notion of a source signal (henceforth referred to as source), since the source distortion and the source rate are of critical importance in the development of the paper. In general, we consider as source any signal which is quantized. For example, if the quantizers are selected per block (or macro-block in H.263 notation) then each block is considered a source. On the other hand, if the quantizer is selected and fixed on a frame by frame basis (which is a popular strategy), then the frames are considered different sources. In speech coding, it is common to segment the original waveform into blocks of 10 to 30 milliseconds length. Each of these blocks is then considered a source. For shape coding, the identification of a source becomes more difficult. Basically the original boundary is considered a sequence of boundary segments and each boundary segment is approximated (quantized) by either a straight line or a higher order curve. Hence the boundary segments are considered sources. A complicating factor is that the original boundary is not only segmented into one possible sequence of segments, but into all possible sequences of segments. Therefore, a given source (boundary segment) will in general consist of several other sources (smaller boundary segments) and this relationship holds until a boundary segment consists simply of two consecutive pixels. Nevertheless, with this view of a source in mind the boundary encoding is simply an application of the more general theory.

A dependent coding framework implies that the source rate $r_i(x_{i-a}, \ldots, x_{i+b})$ and/or source distortion $d_i(x_{i-a}, \ldots, x_{i+b})$ for a given source S_i depends not only on the quantizer x_i applied to that source, but also on neighboring quantizers $x_{i-a}, \ldots, x_{i+b}$ in a neighborhood defined by two non-negative integers a and b. Examples of such dependent coding frameworks are all predictive coding schemes, such as motion compensated video compression, segmentation encoding, image coding, etc. For example, in motion compensated video compression, the quantizer selected for the previous frame has a direct influence on the rate-distortion characteristic of the current frame, since the reconstructed previous frame and the motion information is used to predict the current frame.

The total rate for encoding all sources is the sum of the source rates, and defined as follows,

$$R(x_0, \ldots, x_{N-1}) = \sum_{i=0}^{N-1} r_i(x_{i-a}, \ldots, x_{i+b}). \tag{1}$$

Depending on the employed distortion criterion, the distortion for encoding all sources $D(x_0, \ldots, x_{N-1})$ is the sum (or average) of the source distortions (MINAVE),

$$D(x_0, \ldots, x_{N-1}) = \sum_{i=0}^{N-1} d_i(x_{i-a}, \ldots, x_{i+b}), \tag{2}$$

or the maximum of the source distortions (MINMAX),

$$D(x_0, \ldots, x_{N-1}) = \max_{i \in [0, \ldots, N-1]} \{d_i(x_{i-a}, \ldots, x_{i+b})\}, \tag{3}$$

with $x_{-a}, \ldots, x_{-1}$ and $x_N, \ldots, x_{N+-1+b}$ specifying the boundary conditions.

In either case, two optimal bit allocation problems can be formulated, the minimum rate problem and the minimum distortion problem (which is also called the rate constrained problem). In the minimum rate problem we are looking for the quantizer sequence which results in the smallest bit rate for a given maximum distortion D_{max}. This can be formulated as follows,

$$\min_{x_0, \ldots, x_{N-1}} R(x_0, \ldots, x_{N-1}), \;\; \text{s.t.:} \;\; D(x_0, \ldots, x_{N-1}) \leq D_{max}. \tag{4}$$

In the minimum distortion problem we are looking for the quantizer sequence which results in the smallest distortion for a given maximum bit rate R_{max}. This can be formulated as follows,

$$\min_{x_0, \ldots, x_{N-1}} D(x_0, \ldots, x_{N-1}), \;\; \text{s.t.:} \;\; R(x_0, \ldots, x_{N-1}) \leq R_{max}. \tag{5}$$

The above formulations hold for either of the two distortion criteria. It should be noted, however, that in some applications, N, the number of sources retained for quantization, may be different from the number of sources present in the original signal. In this case, finding the optimal N is part of the optimization.

III. The MINAVE Criterion

In this section we show how the optimal bit allocation problem can be solved for the MINAVE criterion which is defined in Eq. (2). The solution is based on the Lagrange multiplier method and DP. Note that for the MINAVE criterion, the total rate and the total distortion are of exactly the same form and hence the solution approach for the minimum rate problem is equivalent to the solution approach for the minimum distortion problem. Therefore, in this section, we will concentrate on the minimum distortion problem, keeping in mind that by re-labeling the function names, i.e., $r(\cdot) \leftarrow d(\cdot)$, $d(\cdot) \leftarrow r(\cdot)$, $R(\cdot) \leftarrow D(\cdot)$, and, $D(\cdot) \leftarrow R(\cdot)$, the minimum rate problem can be solved using the same approach.

The basic idea behind the Lagrange multiplier method is to merge the rate and the distortion with a Lagrangian multiplier λ. This results in the Lagrangian cost function which is of the following form,

$$J_\lambda(x_0, \ldots, x_{N-1}) = D(x_0, \ldots, x_{N-1}) + \lambda \cdot R(x_0, \ldots, x_{N-1}). \tag{6}$$

The goal of this method is to convert the "hard" constrained problem of Eq. (5) into a set of "easy" unconstrained problems parameterized by λ.

It has been shown in [3], [31] that if there is a λ^* such that,

$$[x_0^*, \ldots, x_{N-1}^*] = \arg \min_{x_0, \ldots, x_{N-1}} J_{\lambda^*}(x_0, \ldots, x_{N-1}), \tag{7}$$

leads to $R(x_0^*, \ldots, x_{N-1}^*) = R_{max}$, then $[x_0^*, \ldots, x_{N-1}^*]$ is also an optimal solution to the minimum distortion problem of Eq. (5) for the MINAVE criterion. It is well known that when λ sweeps from zero to infinity, the solution to problem (7) traces out the convex hull of the operational rate distortion curve, which is a non-increasing function. Hence bisection [6] can be used to find λ^*. The main problem with the Lagrange multiplier method is that only solutions which belong to the convex hull can be found.

Clearly the efficiency of the Lagrange multiplier method depends on the assumption that the unconstrained problem of Eq. (7) can be solved efficiently. Based on the assumptions made in section II, the source rates and distortions only depend on the quantizers selected in a neighborhood around the current source. Therefore the Lagrangian cost function can be expressed as follows,

$$J_\lambda(x_0,\ldots,x_{N-1}) = \sum_{i=0}^{N-1}(d_i(x_{i-a},\ldots,x_{i+b}) + \lambda \cdot r_i(x_{i-a},\ldots,x_{i+b})) = \sum_{i=0}^{N-1} j_{\lambda,i}(x_{i-a},\ldots,x_{i+b}), \quad (8)$$

where $j_{\lambda,i}(\cdot) = d_i(\cdot) + \lambda \cdot r_i(\cdot)$. The Lagrangian multiplier method is based on the assumption that the above unconstrained problem can be solved optimally and efficiently. In the following section we will closely follow the development of the DP recursion formula in [23]. Using it we will then show how DP can be used to optimally solve the unconstrained problem efficiently, as long as, the size of the neighborhood $(a+b)$ is reasonably small.

A. DP recursion formula

Even though there exists only a finite number of combinations in which source quantizers $(x_0,\ldots,x_{N-1})$ in Eq. (8) can be combined, the exhaustive search is too complex. Assuming the same cardinality $|X_i|$ for all N source quantizers, the complexity of exhaustive search is $O(|X_i|^N)$, where complexity refers to the number of times the cost function $J_\lambda(x_0,\ldots,x_{N-1})$ needs to be evaluated. With the use of the proposed DP algorithm, this complexity can be significantly reduced.

Dropping the subscript λ, we denote by $j_l^*(\cdot)$ the minimum Lagrangian cost up to and including neighborhood l, that is

$$j_l^*(x_{l-a+1},\ldots,x_{l+b}) = \min_{x_0,\ldots,x_{l-a}} \sum_{i=0}^{l} j_i(x_{i-a},\ldots,x_{i+b}). \quad (9)$$

From Eq. (9) it follows that,

$$j_{l+1}^*(x_{l+1-a+1},\ldots,x_{l+1+b}) = \min_{x_0,\ldots,x_{l+1-a}} \sum_{i=0}^{l+1} j_i(x_{i-a},\ldots,x_{i+b})$$
$$= \min_{x_{l+1-a}} \left[\min_{x_0,\ldots,x_{l-a}} \left(\sum_{i=0}^{l} j_i(x_{i-a},\ldots,x_{i+b}) + j_{l+1}(x_{l+1-a},\ldots,x_{l+1+b}) \right) \right]. \quad (10)$$

Since $j_{l+1}(x_{l+1-a},\ldots,x_{l+1+b})$ does not depend on $x_0,\ldots,x_{l-a}$, it can be moved outside the inner minimization and the following DP recursion formula results,

$$j_{l+1}^*(x_{l+1-a+1},\ldots,x_{l+1+b}) = \min_{x_{l+1-a}} \left[j_l^*(x_{l+1-a},\ldots,x_{l+b}) + j_{l+1}(x_{l+1-a},\ldots,x_{l+1+b}) \right]. \quad (11)$$

Assuming that all source quantizers have the same cardinality $|X_i|$, this recursive formula requires only $|X_i|^{a+b+1}$ comparisons to find $j^*_{l+1}(x_{l+1-a+1},\ldots,x_{l+1+b}),\forall\, x_{l+1-a+1},\ldots,x_{l+1+b}$ given $j^*_l(x_{l+1-a},\ldots,x_{l+b})$. Thus the overall complexity is $O(N\cdot|X_i|^{a+b+1})$, which is a significant reduction from $O(|X_i|^N)$ required by the exhaustive search. It is important to note that the time complexity depends directly on the size of the neighborhood $(a+b)$. In other words, if the the dependency among quantizers is well localized, such as in single predictor DPCM $(a=1,b=0)$ or a second order B-spline ($a=1,b=1$, see section V-C), the above DP formulation results in a low time complexity. Clearly, if the dependency is not well localized, as for example the vector median based DPCM for the motion vector prediction in H.263 $(a=11,b=0)$ [24], [29] the time complexity increases exponentially. Furthermore, in cases where the dependency is global, for example the current quantizer depends on every previous quantizer, then the above DP algorithm degenerates to an exhaustive search. Nevertheless, most problems have a strong local dependency and hence the DP approach works well.

B. Forward DP algorithm

Having established the DP recursion formula (Eq. 11), we apply the Viterbi algorithm [5] to arrive at the optimal solution, which consists of the minimum cost $\min_{x_{N-a},\ldots,x_{N-1+b}} j^*_{N-1}(x_{N-a},\ldots,x_{N-1+b})$ and its associated sequence of source quantizers. The recursion is first initialized,

$$j^*_{a-1}(v_0,\ldots,v_{a+b-1})=\sum_{i=0}^{a-1} j_i(x_{i-a},\ldots,x_{i+b})\quad\forall\,[x_0,\ldots,x_{a+b-1}] \tag{12}$$

and the optimal selection back-pointer is introduced,

$$i_{a-1}(x_0,\ldots,x_{a+b-1})=[x_{-a},\ldots,x_{b-1}]. \tag{13}$$

Next, the recursion formula is applied for $l=a-1$ up to and including $l=N-2$, that is,

$$j^*_{l+1}(x_{l+1-a+1},\ldots,x_{l+1+b})=\min_{x_{l+1-a}}\left[j^*_l(x_{l+1-a},\ldots,x_{l+b})+j_{l+1}(x_{l+1-a},\ldots,x_{l+1+b})\right], \quad \forall\,[x_{l+1-a+1},\ldots,x_{l+1+b}]. \tag{14}$$

Again the back pointer is assigned, for which the argument which minimizes the above problem needs to be known,

$$x^*_{l+1-a}(x_{l+1-a+1},\ldots,x_{l+1+b})=\arg\min_{x_{l+1-a}}\left[j^*_l(x_{l+1-a},\ldots,x_{l+b})+j_{l+1}(x_{l+1-a},\ldots,x_{l+1+b})\right], \quad \forall\,[x_{l+1-a+1},\ldots,x_{l+1+b}], \tag{15}$$

which is

$$i_{l+1}(x_{l+1-a+1},\ldots,x_{l+1+b}) = [x^*_{l+1-a}(x_{l+1-a+1},\ldots,x_{l+1+b}), x_{l+1-a+1},\ldots,x_{l+b}], \qquad (16)$$
$$\forall\ [x_{l+1-a+1},\ldots,x_{l+1+b}].$$

The last back-pointer is the argument that minimizes the cost at the last source,

$$[x^*_{N-a},\ldots,x^*_{N-1}] = \arg\min_{[x_{N-a},\ldots,x_{N-1}]} j^*_{N-1}(x_{N-a},\ldots,x_{N-1+b}). \qquad (17)$$

Consequently, the optimal sequence of source quantizers $x^*_{N-a},\ldots,x^*_{N-1+b}$ is found by following the back-pointers during the backtracking stage, i.e.,

$$x^*_{l-a} = \left[i_l(x^*_{l-a+1},\ldots,x^*_{l+b})\right]_1, \quad l = N-1,\ldots,a \quad , \qquad (18)$$

where $[\cdot]_1$ refers to the first element in the vector.

IV. The MINMAX Criterion

In this section we propose a general algorithm for the optimal bit allocation among dependent quantizers for the MINMAX criterion [27]. The basic idea behind the proposed algorithm is to solve the minimum rate problem optimally using DP. This is possible since the maximum distortion constraint D_{max} applies to *each* source and not to the sum of the source distortions. We then prove that the operational rate distortion function is non-increasing. Therefore we can solve the minimum distortion problem, which is a min max problem, using bisection, where in each bisection iteration the minimum rate problem is solved using a different D_{max}.

A. The Minimum Rate Problem

In this section, we solve the minimum rate problem which is described in Eq. (4) for the MINMAX criterion. The key observation for the derivation of the optimal solution is that, the maximum distortion D_{max} constraint applies to each source, and not, as in the case of the MINAVE criterion, to the sum of the source distortions. We can make use of this fact by redefining the source rates as follows,

$$r_i(x_{i-a},\ldots,x_{i+b}) = \begin{cases} \infty & : \quad d_i(x_{i-a},\ldots,x_{i+b}) > D_{max} \\ r_i(x_{i-a},\ldots,x_{i+b}) & : \quad d_i(x_{i-a},\ldots,x_{i+b}) \leq D_{max} \end{cases} . \qquad (19)$$

In words, the rate for a source with a distortion which is larger then the maximum permissible distortion is set to infinity. This results in the fact that, given that a feasible solution exists, the quantizer sequence which minimizes the total rate, as defined in Eq. (1), will not result in any source distortion greater than D_{max}. If no feasible solution exists, then the resulting minimum total rate is infinite, hence this situation can easily be detected and D_{max} can be increased. In other words, the minimum rate problem, which is a constrained optimization problem, can be transformed into an unconstrained optimization problem using the above re-definition of the source rates.

The structure of the total rate formula in Eq. (1) is equivalent to the structure of the Lagrangian cost function for the MINAVE case in Eq. (8). Hence the optimal solution to the unconstrained minimum rate problem can also be solved by DP as shown in section III-B.

We can now calculate the operational rate distortion function $R^*(D_{max})$ as follows,

$$R^*(D_{max}) = \min_{x_0,\ldots,x_{N-1}} R(x_0,\ldots,x_{N-1}), \;\; \text{s.t.:} \;\; D(x_0,\ldots,x_{N-1}) \leq D_{max}, \tag{20}$$

where we assume that D_{max} is a variable.

B. The Minimum Distortion Problem

The proposed optimal bit allocation algorithm for the minimum distortion problem is based on the fact that we can optimally solve the minimum rate problem. In other words, for every given D_{max} we can find the quantizer sequence which results in $R^*(D_{max})$, the minimum rate for encoding the combined sources, where each source distortion has to be below the maximum distortion D_{max} (see Eq. (20)). We use the following Theorem to formulate an iterative procedure to find the optimal solution for the minimum distortion problem.

Theorem 1: $R^*(D_{max})$ is a non-increasing function of D_{max}.

Proof: Let $D^2_{max} \geq D^1_{max}$, $[{}^1x^*_0,\ldots,{}^1x^*_{N-1}]$ be the optimal solution of Eq. (4) for $D_{max}{=}D^1_{max}$, and $[{}^2x^*_0,\ldots,{}^2x^*_{N-1}]$ the optimal solution of Eq. (4) for $D_{max}{=}D^2_{max}$. Since $D^1_{max} \leq D^2_{max}$, $[{}^1x^*_0,\ldots,{}^1x^*_{N-1}]$ is a possible solution of Eq. (4) for $D_{max}{=}D^2_{max}$, using $R^*(D^1_{max})$ bits. Since $[{}^2x^*_0,\ldots,{}^2x^*_{N-1}]$ is the optimal solution of Eq. (4) for $D_{max}{=}D^2_{max}$, it follows that $R^*(D^2_{max}) \leq R^*(D^1_{max})$. ■

The above Theorem is intuitively clear since it simply states that if a greater maximum distortion is permissible, then we should be able to encode the sources with a smaller number of bits. Note that even though this seems obvious, this only holds true because we can solve the minimum rate case optimally.

Having shown that $R^*(D_{max})$ is a non-increasing function, we can use bisection to find the optimal D^*_{max} such that $R^*(D^*_{max}) = R_{max}$, which solves the minimum distortion problem of Eq. (5) for the MINMAX criterion. The bisection method starts with two points $(D^l_{max}, R^*(D^l_{max}))$ and $(D^u_{max}, R^*(D^u_{max}))$ which bracket the optimal solution (see Fig. 1). Then a middle point $(D^m_{max}, R^*(D^m_{max}))$ is found by invoking the minimum rate algorithm for,

$$D_{max} = D^m_{max} = (D^l_{max} + D^u_{max})/2. \tag{21}$$

The new bracketing points of the optimal solution are then the middle point and one of the original points which results in a bracket which includes the optimal solution. This procedure is then iterated until the optimal solution is found or the bracket is small enough for the purpose at hand.

Since this is a discrete optimization problem, the function $R^*(D_{max})$ is not continuous and exhibits a staircase characteristic (see Fig. 1). This implies that there might not exist a D^*_{max} such that $R^*(D^*_{max}) = R_{max}$. In that case, the proposed algorithm will still find the optimal solution, which is of the form $R^*(D^*_{max}) < R_{max}$, but only after an infinite number of iterations. Hence we stop the algorithm after a fixed number of iterations.

C. Breaking the tie

Sometimes the solution to the MINMAX distortion problem is not unique. That is, two or more source quantizer sequences result in the same minimum rate $R^*(D_{max})$ for the same distortion D_{max}. One way out of this problem is to arbitrarily select one of the solutions, for example the one with the smallest index. This is in fact the way most encoders resolve this problem.

Another way to break this tie is with the use of secondary distortion measures first introduced for boundary encoding in [23]. As shown above, DP is used to solve the minimum rate MINMAX problem directly by redefining the source rates. Secondary objectives can easily be included in the DP framework. We propose to use the MINAVE criterion as a secondary objective to break the tie between two or more optimal MINMAX solutions. The source rates are changed again, and in fact something very similar to a Lagrangian is defined. Recall the source rates were redefined before in Eq. (19) for the MINMAX case. The following definition changes Eq. (19) to include a MINAVE based secondary distortion measure, weighted by β,

$$j_{i,\beta}(x_{i-a}, \ldots, x_{i+b}) = \begin{cases} \infty & : \quad d(x_{i-a}, \ldots, x_{i+b}) > D_{max} \\ r(\cdot) + \beta * \tilde{d}(\cdot) & : \quad d(x_{i-a}, \ldots, x_{i+b}) \le D_{max} \end{cases}, \tag{22}$$

where $r(\cdot)$ and $\tilde{d}(\cdot)$ stand for $r(x_{i-a},\ldots,x_{i+b})$ and $\tilde{d}(x_{i-a},\ldots,x_{i+b})$, respectively, with the latter one being the MINAVE criterion-based secondary distortion measure and β is a positive real number. As shown, the DP algorithm finds the sequence of dependent quantizers which minimizes $\sum_{i=0}^{N-1} j_{i,\beta}(x_{i-a},\ldots,x_{i+b})$.

We would like to select β such that it will identify a winner among the optimal solutions, but it should be impossible for any other solution to outperform the optimal ones. Let us denote by $J_1 = R_1 + \beta \cdot \tilde{D}_1$ and $J_2 = R_2 + \beta \cdot \tilde{D}_2$ the costs of two solutions to the original MINMAX problem and let us assume without loss of generality that solution J_1 is optimal. Clearly, if J_2 is also optimal, i.e., $R_1 = R_2$, any positive real β will cause the solution with the smaller secondary distortion $\tilde{D}$ to be selected. If, on the other hand, J_2 is not optimal, i.e., $R_1 < R_2$, then we would like solution J_1 to be selected regardless of the magnitudes of the secondary distortions. That is, $\beta > 0$ must be such that, in this case,

$$R_1 + \beta \cdot \tilde{D}_1 < R_2 + \beta \cdot \tilde{D}_2, \qquad \forall \tilde{D}_1 > 0,\ \tilde{D}_2 > 0. \tag{23}$$

If $\tilde{D}_1 < \tilde{D}_2$, the J_1 solution is correctly selected. Hence we assume $\tilde{D}_1 > \tilde{D}_2$. Rearranging the variables, the following upper bound results,

$$\beta < \frac{\delta R}{\tilde{D}_{max}} < \frac{R_2 - R_1}{\tilde{D}_2 - \tilde{D}_1}, \tag{24}$$

where δR is the smallest possible difference between an optimal and a suboptimal solution, which is usually 1 bit, and $\tilde{D}_{max}$ is the largest possible value of the secondary distortion measure. Consequently, any β satisfying $0 < \beta < \frac{\delta R}{D_{max}}$ is capable of discriminating between two solutions based on a secondary distortion measure. By selecting β as proposed above a suboptimal MINMAX solution will never be selected as the optimal solution, and the secondary distortion measure will only be used to find a winner among the optimal solutions.

It is also possible to incorporate secondary distortion measures, specifically the MINMAX measure, into a DP algorithm operating on the MINAVE criterion. We cannot include a MINMAX measure into the Lagrangian cost function directly. Note, however, that the maximum source distortion exhibits the same order of dependency as the sum of the source distortions. In other words, knowing the current maximum source distortion makes the future of the maximum source distortion independent from the past. We now apply this knowledge to the the DP recursion formula of Eq. 11. If two (or more) x_{l+1-a} result in the same $j^*_{l+1}(\cdot)$ the one resulting in the smallest maximum source distortion is selected and the new minimum maximum source distortion is stored together with the backpointer identifying the optimal x_{l+1-a}.

Yet another way to break the tie between two or more optimal MINMAX solutions is through application of the lexicographic optimality principle [8], which can be considered an extension of the MINMAX criterion. If the MINMAX solution is unique, then it is also the lexicographically optimal solution. This is the reason why we discuss this approach in this section. The lexicographical optimality criterion can be explained as follows. If two or more MINMAX solutions exist, a sorted list is created for all candidate optimal solutions. It contains the source distortions introduced by individual source quantizers. These distortions are sorted in the decreasing order and the lists are then scanned sequentially. When there is a tie between two optimal MINMAX solutions the first elements of the respective lexicographically ordered lists must be equal. The lexicographically ordered list which has the smallest second largest source distortion is then considered to represent the smaller total distortion and hence it identifies the optimal solution. If the second largest source distortion still results in a tie, then the third largest source distortion is compared, and so on.

Clearly we do not want to keep all possibly optimal solutions until the end of the DP. Ideally, we would like to eliminate lexicographically suboptimal solutions as early as possible. This can be achieved easily in the proposed DP framework, since a lexicographically ordered list exhibits the same order of dependency as the maximum of the source distortions. In other words, knowing the current lexicographically optimal list of source distortions makes the future of the lexicographically optimal list independent from the past. We now apply this knowledge to the DP recursion formula of Eq. (11). If two (or more) x_{l+1-a} result in the same $j^*_{l+1}(\cdot)$ the one resulting in the smallest lexicographically ordered source distortion list is selected and this new list is stored together with the backpointer identifying the optimal x_{l+1-a}.

The main difference between the idea of a secondary distortion measure and a lexicographically optimal solution is that the secondary distortion measure allows for the influence of a completely different distortion measure. In contrast to this, the lexicographically optimal solution is all based on one type of distortion measure. For example, if there are two MINMAX optimal solutions for a boundary approximation, it makes intuitive sense to select the one which results in a smaller MSE. This is the nature of a secondary distortion measure. Clearly the tie can also be broken using a lexicographical criterion, but in this case the braking of the tie is still based on a smallest maximum distortion and not on something else, like an intuitive MSE. In other words, the secondary distortion measure is appealing to us since it allows to combine MINMAX and MINAVE criteria which are both well known concepts for which we inherently have an intuitive understanding.

V. Applications

In this section we present several examples from different areas of data compression to compare the MINAVE and the MINMAX approaches. They all have in common the theme of optimal bit allocation among dependent quantizers, for which solution techniques derived in sections III and IV are applied.

In section V-A we discuss the optimal block quantizer selection for a a still-frame compression scheme [23], [27]. In section V-B we discuss the optimal quantizer, mode and motion vector selection for a motion compensated video compression scheme. Note that the mode and quantizer selection scheme for the MINAVE criterion has been reported in [24] and is very similar to the scheme reported in [32]. The combined selection of motion vector, mode and quantizer for the MINAVE criterion has been reported in [29] while the selection of optimal motion vectors in a lossless video coder are discussed in [21]. Finally, in section V-C we discuss the optimal boundary approximation for a shape encoding scheme [11], [30].

It is important to notice that the presented theory can also be used for completely different coding schemes, as long as the assumptions stated in section II are satisfied.

A. Still frame compression

The dependent image coding scheme we use for this example is the Intra frame scheme employed in TMN4 [4], which is the test model four of the H.263 standard.

For the example at hand, we encode the first frame of the QCIF color sequence “Mother and Daughter”. We use the TMN4 mechanism for transmitting the quantizer step sizes which is based on a modified delta modulation scheme. In TMN4, the quantizer step size of the current macro block must be within ± 2 of the quantizer step size employed for the previous macro block. Then the difference between the quantizer step sizes is entropy coded. This DPCM scheme results in a first order dependency between two consecutive blocks, since the operational rate distortion curve of the current block depends on the quantizer selected for the previous block.

First we fix the quantizer step size for all macro blocks to 10. The resulting rate ($R_{Q=10} =$

18297 bits) and distortion are listed in Table I. Note that the mean squared error (MSE) of the luminance (Y) channel is used as the distortion measure. For both, the MINAVE and the MINMAX criterion, we solve the minimum distortion problem, where we set the maximum rate equal to the the rate TMN4 uses for a fixed quantizer of 10 ($R_{max} = R_{Q=10}$). Again, the resulting rate and distortion are listed in Table I.

In Fig. 3 the MSE per macro block for the three implementations is shown and in Fig. 2 the corresponding quantizer selections are displayed. It is interesting to notice in Fig. 2 that there are quite a few blocks where the quantizers are the same for both optimal schemes. These blocks tend to coincide with the blocks where the MSE (see Fig. 3) is very small, i.e., blocks with no high frequency components. In Fig. 4 all three approaches are compared to the original image. The compression ratio for the three compressed images is approximately 11, and, even though in each case the block MSEs are quite different, the visual quality of the three compressed images is similar, which is another argument in favor of the fact that the global MSE does not tell the whole story. It is clear from Fig. 3 that the minimum maximum distortion scheme results in a more even quality for the entire frame than the minimum total distortion approach. In fact, discounting the blocks with very low MSE, the distortion profile is quite flat and very close to the minimum average distortion achieved by the MINAVE approach. In other words, the result shows that if the goal is to have almost constant distortion, which is almost as low as the smallest possible average distortion, for a given bit budget, the MINMAX criterion is an excellent choice.

B. Inter-mode frame compression

An example of applying the MINMAX approach to a dependent quantizer framework is presented here for the Inter frame coding scheme employed in TMN4 [4]. This example is an optimal block-based (16×16) motion estimation and residual error quantization scheme.

As in the previous section, we use the QCIF color sequence "Mother and Daughter" to compare performance of MINAVE [29], MINMAX, as well as a fixed-quantizer scheme.

The displaced vector field (DVF) encoding scheme used here differs from TMN4 in two aspects. First, instead of a raster scan, the modified Hilbert scan [23] is employed to achieve a higher correlation between consecutive motion vectors. Second, only a first order DPCM encoding is used and not a vector median-based one, which, because of the neighborhood's small size ($a = 1,\ b = 0$ in section III-A), results in a faster optimization procedure. The motion is estimated with a half-pixel accuracy with the maximum range of $\pm$ 15.5 pixels.

We use the TMN4 modified delta modulation scheme for transmitting the quantizer step sizes between the current and the previous blocks, with the difference limited to ± 2. This difference is then entropy coded and sent to the decoder together with the quantization index.

Conceptually, for a given block, we may think of selecting a particular motion vector followed by selecting the residual error quantizer as a single quantization operation. Since the number of these generalized quantizers is finite and they exhibit a first order (rate) dependency, the techniques of sections III and IV are applicable.

In the following experiments we apply the scheme mentioned above to encode the "Mother and Daughter" sequence in the Inter mode at 7.5 frames/second, i.e., every 4^{th} frame. As with the Intra mode, the MSE of the luminance (Y) channel is used as the distortion measure.

First we fix the quantizer step size for all macro blocks in all frames to 10. In this case, the encoder can only decide in which mode (Intra, Inter, Skip, or Prediction) a given block is encoded, and, in case of Inter, which motion vector to use. These decisions, however, are made using a heuristic algorithm. The resulting rates, for each frame, are stored in the rate profile to be matched by the MINAVE and MINMAX optimal encoders, which, in addition to the mode, can also optimally choose both a motion vector and a quantizer for each block. That is, both MINAVE and MINMAX solve the minimum distortion problem, subject to the rate being equal (within 25 bits) to that used by TMN4 with the fixed quantizer step size encoding. The resulting average block MSE evolution for the three encoders is shown in Fig. 5 for all coded frames. Similarly, Figures 6 and 7 show the evolutions of the maximum and the standard deviation of the block MSE. Clearly, the MINMAX approach leads to a greater uniformity in the maximum MSE across all frames. It may seem strange that, within each frame, the MINAVE and MINMAX approaches exhibit similar degree of variability of the block MSEs, as evidenced by Fig. 7. However, a global measure like the standard deviation is inherently incapable of penalizing the few outliers possible under MINAVE, and, hence, obscures the benefits of MINMAX. As expected, the MINMAX approach is outperformed by both MINAVE and fixed quantizer schemes in terms of the average block MSE. The optimal quantizer selections and the optimal motion vector fields are displayed in Figures 8, 9, 10, 11, 12, and 13 for the fixed quantizer, MINAVE, and MINMAX encodings, respectively. It is interesting to note that in all cases no motion vectors are found in the hand area, since this newly appeared object was not present in the previous reference frame. In general, with the MINMAX approach an almost uniform level of quality is achieved, which, arguably, is more consistent with the subjective assessment of quality than the global MSE.

C. Shape coding

In this example we demonstrate the application of MINMAX and MINAVE approaches to the problem of lossy boundary encoding. Recently this problem has attracted considerable attention as a result of emerging multimedia applications and the MPEG-4 standardization effort.

A number of algorithms have been recently reported, like for example, the context-based CAE coder [2], the Modified Modified Read (MMR) coder [33], the baseline coder [13], the vertex-based polynomial coders [9], [18], as well as, the recently proposed optimal B-spline coders [15], [16], [30]. A review and comparison of shape coding algorithms can be found in [11].

The problem at hand is the lossy approximation of a given closed discretized contour by connected segments of a given order (lines or higher order curves). In the following discussion we refer to 2^{nd}-order B-spline segments, with each segment being defined by three consecutive control points. The precise mathematical definition of this parametric curve is given in [11]. A continuous spline segment is quantized to fit the discreet support grid of the image. Optimal placement of the control points, under a chosen differential encoding scheme of their position and a segment distortion measure, constitutes the solution to the contour approximation problem. Although on the surface this problem seems quite different from the other examples presented in this paper, it can be formulated and solved using the same methods.

The definition of a distortion between the original and the approximating boundaries is not unique. When the MINMAX criterion is employed, the problem at hand is to minimize the rate while guaranteeing that none of the pixels of the approximating contour is located farther than D_{max} (Eucledian distance) away from the original contour. To aid in the implementation, as well as, understanding of the algorithm, we define a distortion band, centered around the original boundary, to which a MINMAX approximation must belong. (In [12], D_{max} was allowed to vary). Figure 14, in which also the original and the approximating contours are shown, illustrates this concept. We note, however, that D_{max} is not a true metric, since there may be pixels on the original contour located a distance $D > D_{max}$ from the approximating contour.

For the MINAVE criterion, we choose to adapt the distortion metric used by MPEG-4 to evaluate efficiency of competing shape coders. A contour distortion is defined as the number of incorrectly labeled pixels, i.e., all pixels in the interior of the original object and in the exterior of the approximating object, or vice versa, normalized by the total number of interior pixels in the original frame. A frame distortion d_n, which is the sum of its contour distortions, is then defined as

$$d_n = \frac{\text{number of pixels in error}}{\text{number of all interior pixels}}. \tag{25}$$

Let p_{u-1}, p_u and p_{u+1} be three consecutive control points defining a B-spline segment of the approximating boundary. This segment originates at the midpoint (p_{u-1}, p_u) and terminates at the midpoint (p_{u+1}, p_u). Let us also associate the beginning and the end of the spline segment with l' and m', the two closest points on the original boundary, respectively. Then the segment distortion can be evaluated, as in Fig. 15, by counting the number of pixels in error. Clearly, the contour distortion under the MINAVE criterion, in this case, is additive, since it can be defined on a segment-by-segment basis.

In order to de-correlate consecutive control point locations a 2^{nd}-order prediction model is used [11]. Every control point is encoded in terms of the relative angle α and the length β (in pixels), as depicted in Figs. 16A and 16B. Control point locations are not restricted to belong to the original boundary, since the problem at hand is that of approximation and not interpolation. For reasons of computational complexity, however, a fixed width admissible control point band is defined around the original contour, similar to the distortion band of Fig. 14, to which control points must belong. We sequentially number all original boundary pixels and associate every control point band pixel with the boundary pixel closest to it. Implementation details, including specification of the VLC tables for α and β, can be found in [11], [16].

Having defined the segment rate $r(p_{u-1}, p_u, p_{u+1})$ and the segment distortion $d(p_{u-1}, p_u, p_{u+1})$, the problem is the selection of control points under either MINMAX or MINAVE criterion in the framework of resource allocation among dependent quantizers or states, so that the tools developed in Sec. III and IV can be applied to find the optimal solution. We note that when the MINMAX criterion is employed, $d(p_{u-1}, p_u, p_{u+1})$ is two-valued: 0 when the segment falls within the distortion band, and ∞ otherwise. Let us define a state as a grouping of two consecutive control points. A transition between two states (p_{u-1}, p_u) and (p_u, p_{u+1}) is then labeled with a Lagrangian cost $w(p_{u-1}, p_u, p_{u+1}) = r(p_{u-1}, p_u, p_{u+1}) + \lambda \cdot d(p_{u-1}, p_u, p_{u+1})$, representing one B-spline segment. Thus we can associate a "quantizer" with every pair of consecutive states, since it has both the rate and the distortion characteristics. Thus, an ordered set of control points corresponding to the sequence of dependent quantizers with the least total cost constitutes the optimal solution to the contour approximation problem. The optimal solution is obtained by applying the DP recursion formula, described in Sec. III-A.

We conduct experiments on the 100 frames of the "Kids" sequence in the intra mode, i.e., without taking into account the temporal correlation between frames. We compare the MINAVE and the MINMAX algorithms in terms of their rate vs distortion characteristics in Fig. 17,where both the rate and the distortion were averaged over 100 frames. As expected, MINAVE outperforms MINMAX with respect to the global distortion measure used for its optimization. When we consider the visual quality, however, the MINMAX result may be preferred. Figures 18 and 19 show the same object encoded using the MINAVE and MINMAX criteria with 343 and 347 bits, respectively. With the areas in error shown in white, it can be seen in Fig. 18 that the object representing the area between the legs of the kid was deemed unimportant based on the global rate distortion tradeoff. With the MINMAX approach the encoder must approximate every object with a given maximum distortion D_{max} which prevents such objects from being skipped.

VI. Conclusions

We conclude this paper by comparing the two optimal bit allocation algorithms for the MINAVE and the MINMAX criterion. The MINAVE approach is based on the Lagrange multiplier method. This method is used to transform the constrained optimization problem into a set of unconstrained optimization problems parameterized by the Lagrangian multiplier λ. These unconstrained problems are then solved optimally using DP. The optimal λ^*, which results in the solution of the original constrained problem, is then found using an iterative approach, such as bisection, where for each iteration the unconstrained problem needs to be solved. For the MINAVE approach, the minimum rate and the minimum distortion problem are both solved by the same algorithm. This is one of the main differences between the MINAVE and the MINMAX approach.

For the MINMAX approach, the minimum rate problem, which is a constrained optimization problem, can be transformed into an unconstrained problem using the re-definition of the source rates. Then this unconstrained problem can be solved directly using DP. In other words, *no* iteration is necessary to solve the minimum rate problem. The minimum distortion problem is then solved using the fact that we can find the optimal solution to the minimum rate problem, which results in a non-increasing operational rate distortion function. The solution to the minimum rate problem is also found by an iterative search for the optimal D^*_{max} using bisection. For each iteration, the minimum rate problem (i.e., the unconstrained problem) is solved using DP.

While these algorithms have many similarities, they are quite different with respect to finding all optimal solutions. This cannot be guaranteed for the Lagrangian approach, since only solutions belonging to the convex hull can be found. Furthermore, while the Lagrangian multiplier method needs an iterative search, for both, the minimum rate and the minimum distortion problem, the MINMAX approach only needs an iteration for the minimum distortion problem. Hence the minimum rate problem can be solved much faster for the MINMAX approach. Ultimately, showing that the MINMAX approach has several algorithmic advantages over the MINAVE approach does not help somebody who needs to solve a MINAVE problem. We do believe however that many real world problems are better served by using a MINMAX approach than a MINAVE approach and that the MINAVE approach has only been so popular since its optimal solution has been known. We have argued that the MINMAX approach based on maintaining a minimum level of local SNR is highly correlated with perceptual quality measures. Now that the optimal solution to the MINMAX criterion is known, we believe many problems should be solved using this approach. This is especially true, since both approaches are based on the same underlying assumptions. Hence every DP algorithm which can solve the MINAVE problem can easily be changed to solve the MINMAX problem.

Both algorithms discussed in this paper (MINAVE and MINMAX) were applied to the intra-frame encoding scheme used in H.263 [23], [27], inter-frame mode of the same standard, as well as, shape coding [11], [30]. The results obtained were compared in terms of the global MSE-like measures and visually. It was clear from these experiments that the MINMAX approach resulted in a more even quality for the entire source than the MINAVE approach. In the shape coding example, the MINMAX approach avoided the problem of skipping perceptually important features and objects. The sacrifices in the average distortion in the MINMAX approach in the conducted experiments were not significant. In conclusion, if the goal is to have almost constant distortion, which is almost as low as the smallest possible average distortion, for a given bit budget, the MINMAX criterion is an excellent choice.

References

[1] G. L. Anderson and A. N. Netravali, "Image restoration based on a subjective criterion," *IEEE Trans. Syst. Man. Cybern.*, SMC-6, pp. 845-853, 1976.

[2] N. Brady, F. Bossen, and N. Murphy, "Context-based arithmetic encoding of 2D shape sequences," *Proc. Intl. Conf. Image Processing*, (Santa Barbara), pp. I-29-32, 1997.

[3] H. Everett, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Operations Research*, vol. 11, pp. 399–417, 1963.

[4] Expert's Group on Very Low Bitrate Visual Telephony, *Video Codec Test Model, TMN4 Rev1.* ITU Telecommunication Standardization Sector, Oct. 1994.

[5] G. D. Forney, "The Vitterbi algorithm," *Proceedings of the IEEE*, vol. 61, no. 3, pp. 268-278, March 1973.

[6] C. F. Gerald and P. O. Wheatley, *Applied numerical analysis.* Addison Wesley, fourth ed., 1990.

[7] D. Haugland, J. G. Heber, and J. H. Husoy, "Optimization algorithms for ECG data compression," *Medical and Biomedical Engineering and Computing.* vol. 35, pp. 420-424, 1997.

[8] D. T. Hoang, E. L. Linzer, and J. S. Vitter, "Lexicographic bit allocation for MPEG video," *Journal of visual communication and image representation*, vol. 8, no. 4, December 1997.

[9] M. Hötter, "Object-oriented analysis-synthesis coding based on moving two-dimensional objects," *Signal Processing: Image Communications*, vol. 2, pp. 409-428, Dec. 1990.

[10] N. Jayant, J. Johnston, and R. Safranek, "Signal compression based on models of human perception," *Proceeding of the IEEE*, vol. 81, no. 10, pp.1385-1422, 1993.

[11] A. K. Katsaggelos, L. Kondi, F. W. Meier, J. Ostermann, G.M. Schuster, "MPEG-4 and Rate Distortion Based Shape Coding Techniques". *Proc. IEEE*, pp. 1126-1154, June 1998.

[12] L. Kondi, F. W. Meier, G. M. Schuster, A. K. Katsaggelos, "Joint optimal object shape estimation and encoding", *Proc. of SPIE*, vol. 3309, pp. 14-25, Jan. 1998.

[13] S. Lee, D. Cho, Y. Cho, S. Son, E. Jang, and J. Shin, "Binary shape coding using 1-D distance values from baseline", *Proc. Intl. Conf. on Image Processing*, (Santa Barbara), pp. I-508-511, 1997.

[14] D. W. Lin, M.-H. Wang, and J.-J. Chen, "Optimal delayed-coding of video sequences subject to a buffer size constraint," in *Proc. Conf. on Visual Communications and Image Processing*, vol. 2094, pp. 223–234, SPIE, 1993.

[15] G. Melnikov, P. V. Karunaratne, G. M. Schuster, A. K. Katsaggelos "Rate-Distortion optimal boundary encoding using an area distortion measure", *Proc. Intl. Symposium on Circuits and Systems*, June 1998.

[16] G. Melnikov, G. M. Schuster, and A. K. Katsaggelos "Simultaneous Optimal Boundary Encoding and Variable-Length Code Selection", *Intl. Conf. on Image Processing*, pp. I-256-260, October 1998.

[17] R. Nygaard and D. Haugland, "Compressing ECG Signals by Piecewise Polynomial Approximation", *Proc. Intl. Conf. on Acoustics, Speech and Signal Processing*, pp. 1809-1812, May 1998.

[18] K. J. O'Connell, "Object-adaptive vertex-based shape coding method", *IEEE Trans. Circuits and Systems for Video Technology*, vol. 7, pp. 251-255, Feb. 1997.

[19] A. Ortega, K. Ramchandran, editors, "Rate-Distortion Techniques for Image/Video Compression", *IEEE Signal Processing Magazine*, Nov. 1998.

[20] K. Ramchandran, A. Ortega, and M. Vetterli, "Bit allocation for dependent quantization with applications to multiresolution and MPEG video coders," *IEEE Transactions on Image Processing*, vol. 3, pp. 533–545, Sept. 1994.

[21] J. Ribas-Corbera and D. L. Neuhoff, "Optimal bit allocations for lossless video coders: motion vectors vs. difference frames," *Intl. Conf. on Image Processing*, pp. II-747-760, 1995.

[22] D. Saupe "Optimal piecewise linear image coding," *Proc. SPIE Conf. on Visual Communications and Image Processing*, vol. 3309, pp. 747-760, 1997.

[23] G. M. Schuster and A. K. Katsaggelos, *Rate-Distortion Based Video Compression, Optimal Video Frame Compression and Object Boundary Encoding*. Kluwer Academic Publishers, 1997.

[24] G. M. Schuster and A. K. Katsaggelos, "Fast and efficient mode and quantizer selection in the rate distortion sense for H.263," in *Proc. Conf. on Visual Communications and Image Processing*, pp. 784–795, SPIE, Mar. 1996.

[25] G. M. Schuster and A. K. Katsaggelos, "A video compression scheme with optimal bit allocation among segmentation, motion and residual error," *IEEE Transactions on Image Processing*, 1997.

[26] G. M. Schuster and A. K. Katsaggelos, "The Minimum-Average and Minimum-Maximum Criteria in Lossy Compression," *Vistas in Astronomy*, vol. 41, no. 3, pp. 427-437, 1997.

[27] G. M. Schuster and A. K. Katsaggelos, "Optimal bit allocation among dependent quantizers for the minimum maximum distortion criterion," in *Proc. Intl. Conf. on Acoustics, Speech and Signal Processing*, vol. 4, pp. 3097–3100, May 1997.

[28] G. M. Schuster and A. K. Katsaggelos, "An optimal polygonal boundary encoding scheme in the rate distortion sense," *IEEE Transactions on Image Processing*, vol. 7, no. 1, pp. 13-26, Jan. 1998.

[29] G. M. Schuster, A. K. Katsaggelos, "A theory of the optimal bit allocation between displacement vector field and displaced frame difference", *Selected Areas in Communications*, vol. 15, no. 9, pp. 1739-1751, December 1997.

[30] G. M. Schuster, G. Melnikov, A. K. Katsaggelos, "Operationally Optimal Vertex-Based Shape Coding", *Signal Processing Magazine*, pp.91-108, Nov. 1998.

[31] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 36, pp. 1445–1453, Sept. 1988.

[32] T. Wiegand, M. Lightstone, D. Mukherjee, T. G. Campbell, and S. K. Mitra, "Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard", *Circuits and Systems for Video Technology*, vol. 6, no. 2, pp. 182-190, April 1996.

[33] N. Yamaguchi, T. Ida, and T. Watanabe, "A binary shape coding method using modified MMR", *Proc. Intl. Conf. Image Processing*, (Santa Barbara), pp. I-504-508, 1997.

	Rate	Distortion (MSE)			
		mean	min	max	std
Q=10	18297	27.8	0.6	66.8	17.5
min total	18431	27.1	0.6	65.0	17.3
min max	18293	29.9	0.6	46.2	15.9

TABLE I

EACH OF THE 99 MACRO BLOCKS (16×16) RESULTS IN A PARTICULAR MSE, AND THE MEAN MSE IS THE MEAN OF THESE 99 MSES. THE SAME HOLDS FOR THE MINIMUM, MAXIMUM AND STANDARD DEVIATION COLUMN.

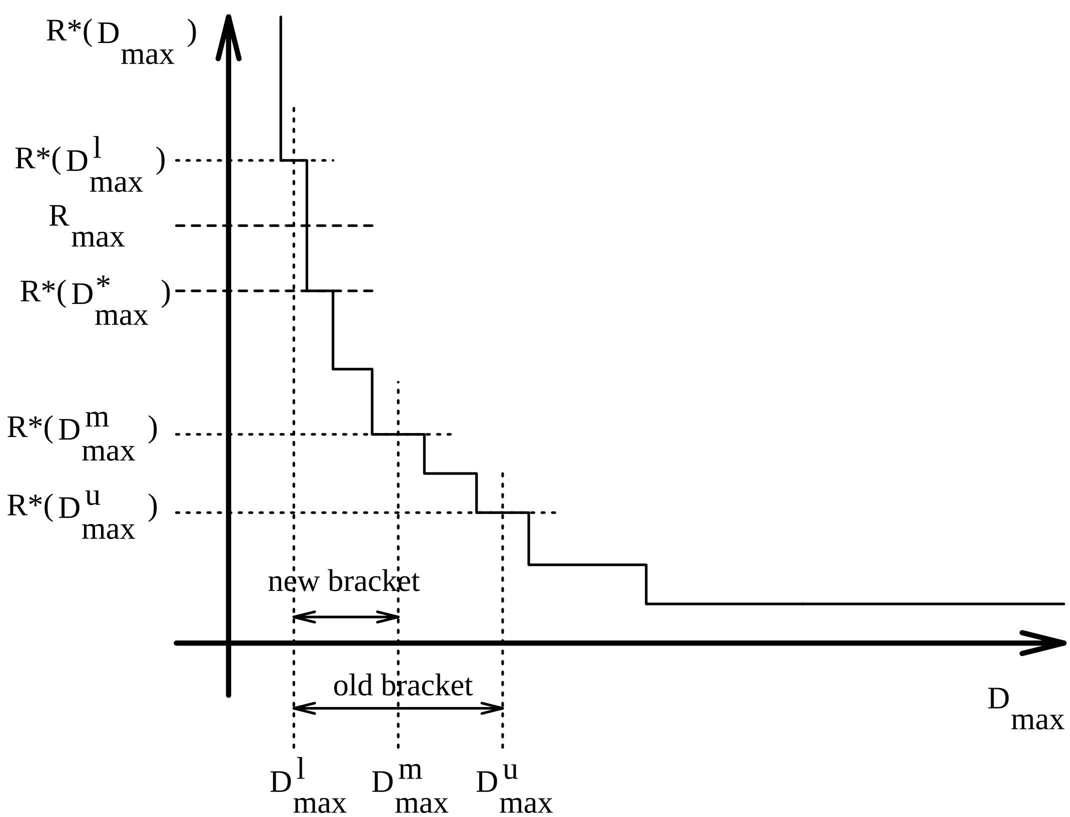

Fig. 1. The $R^*(D_{max})$ function, which is a non-increasing function exhibiting a staircase characteristic. The selected R_{max} falls onto a discontinuity and therefore the optimal solution is of the form $R^*(D^*_{max}) < R_{max}$, instead of $R^*(D^*_{max}) = R_{max}$.

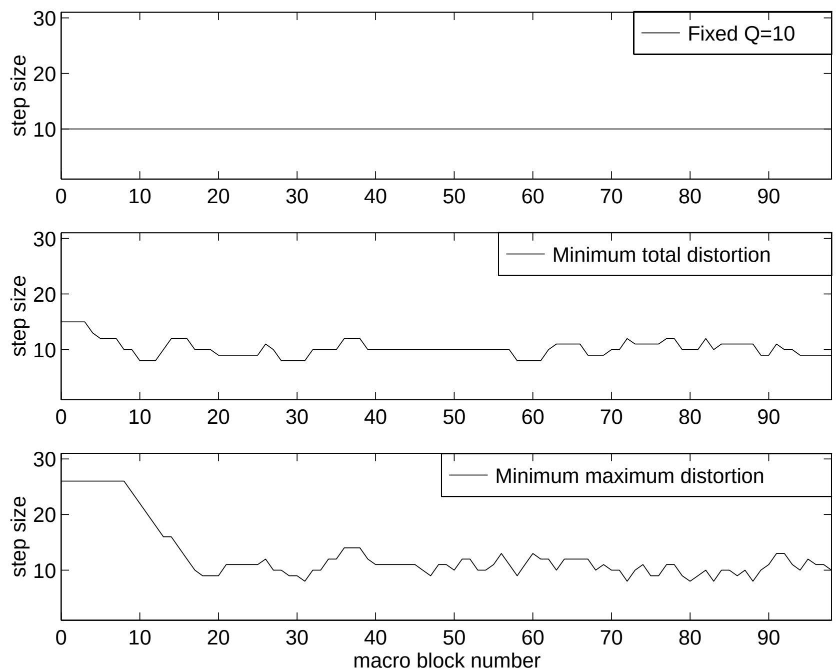

Fig. 2. Macro block quantizer step sizes; First row: fixed quantizer step size of Q=10. Second row: step sizes for the minimum total distortion approach. Third row: step sizes for the minimum maximum distortion approach.

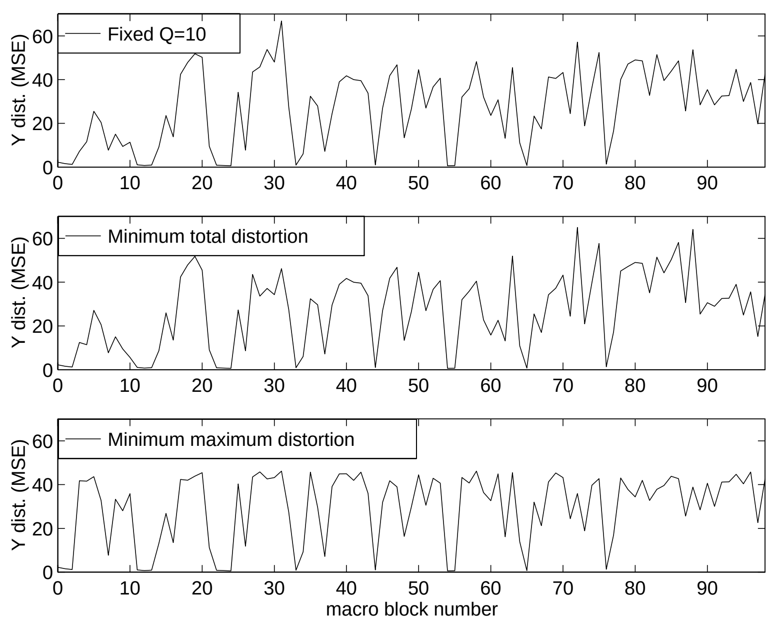

Fig. 3. MSE of each macro block of the luminance channel; First row: MSE for a fixed quantizer step size of Q=10. Second row: MSE for the minimum total distortion approach. Third row: MSE for the minimum maximum distortion approach.

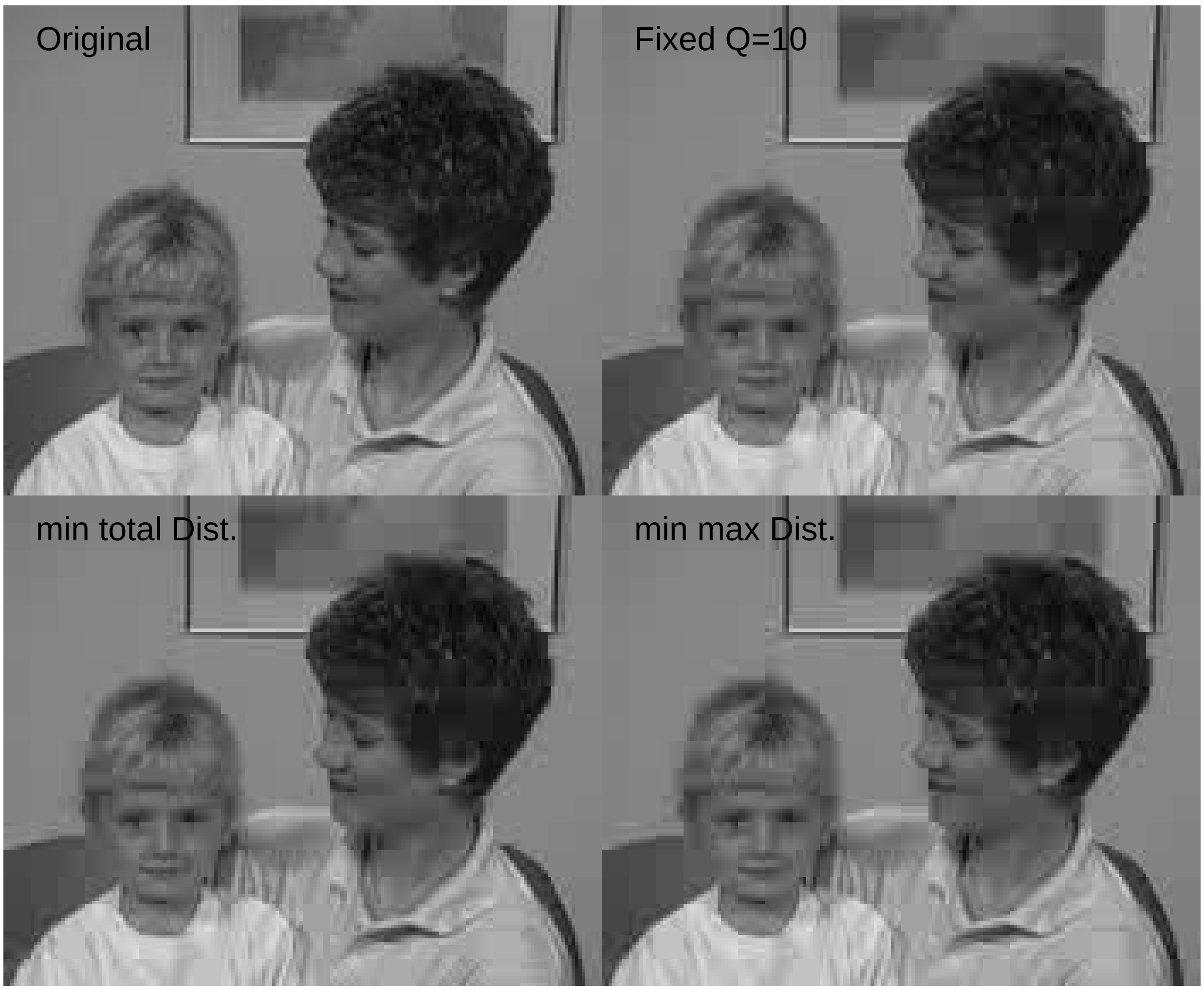

Fig. 4. Comparison between MINMAX, MINAVE, and fixed quantizer approaches with the same compression ratio.

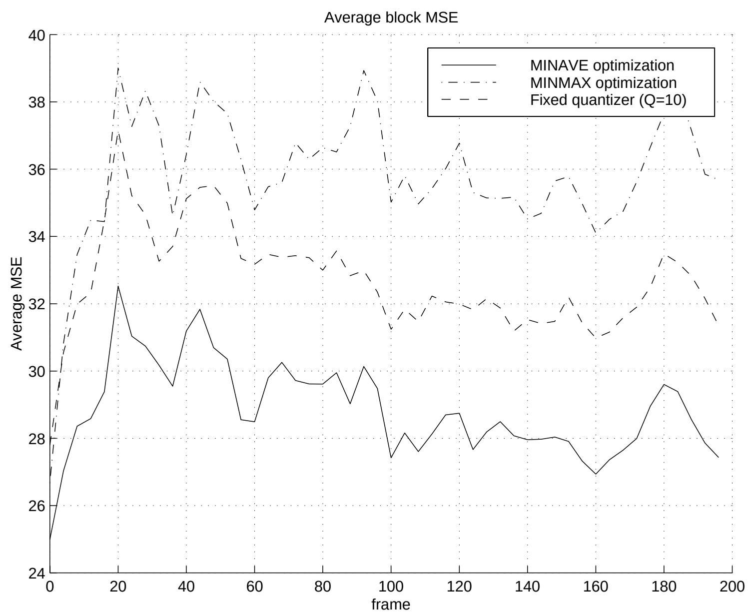

Fig. 5. Comparison between the average block PSNR ($PSNR = 10 * log_{10}(255^2/MSE)$) for the MINMAX, MINAVE, and fixed quantizer approaches with the same compression ratio.

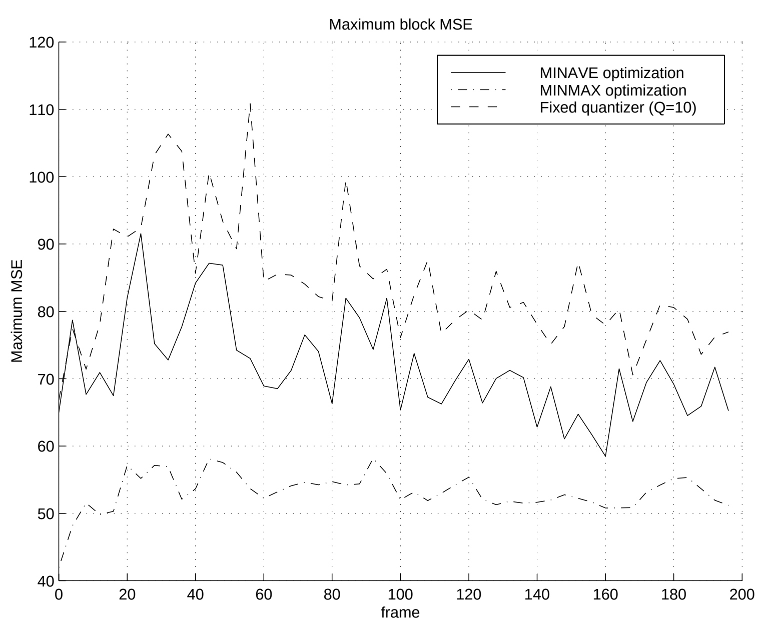

Fig. 6. Comparison between the maximum block PSNR ($PSNR = 10 * log_{10}(255^2/MSE)$) for the MINMAX, MINAVE, and fixed quantizer approaches with the same compression ratio.

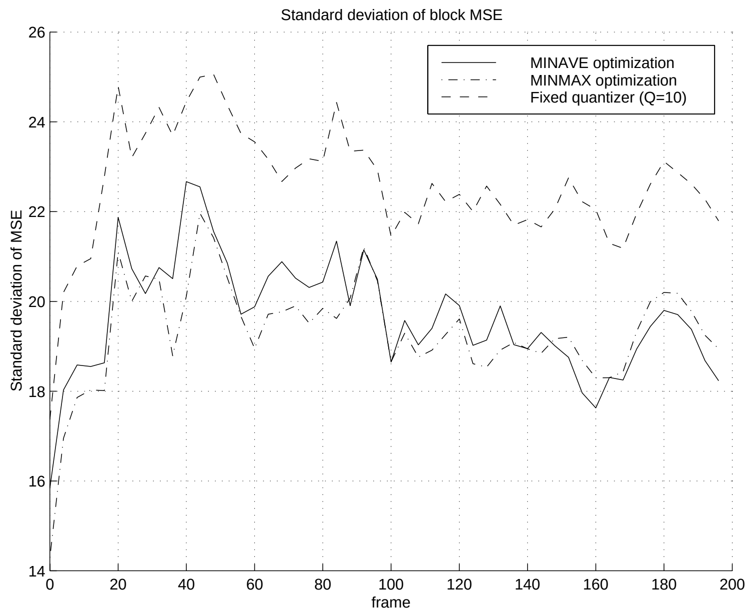

Fig. 7. Comparison between the standard deviation of block MSE for the MINMAX, MINAVE, and fixed quantizer approaches with the same compression ratio.

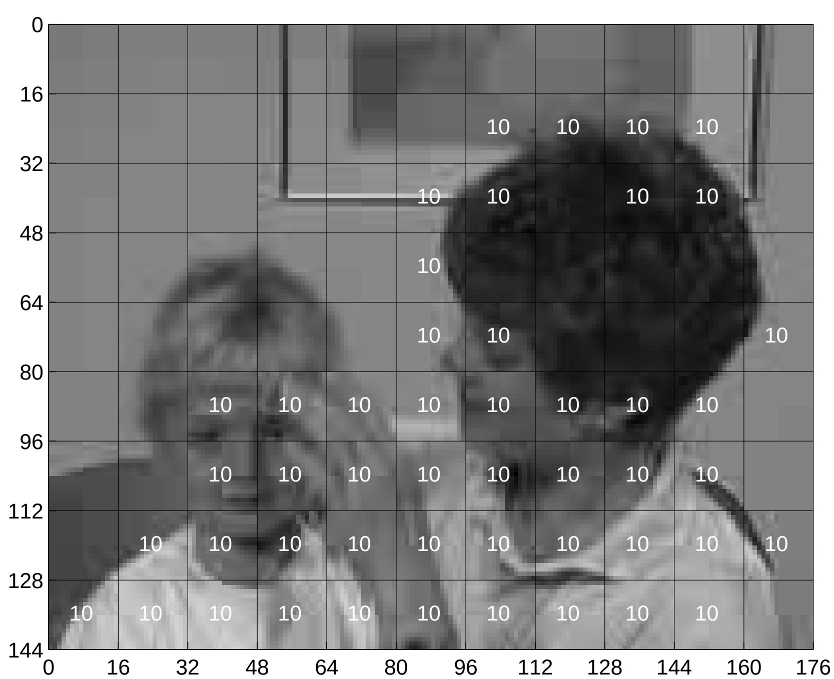

Fig. 8. Quantizers in the fixed quantizer scheme.

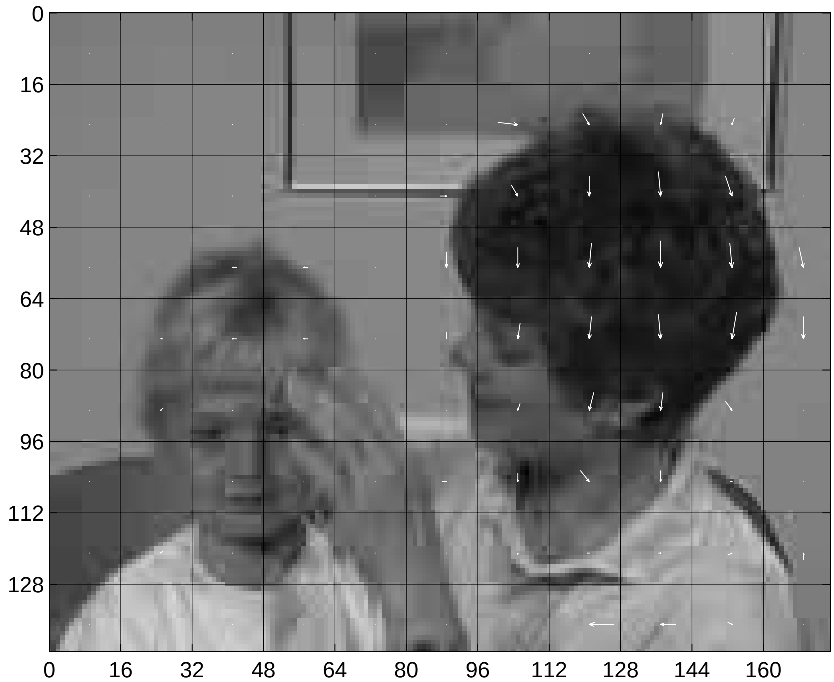

Fig. 9. Motion vectors in the fixed quantizer scheme.

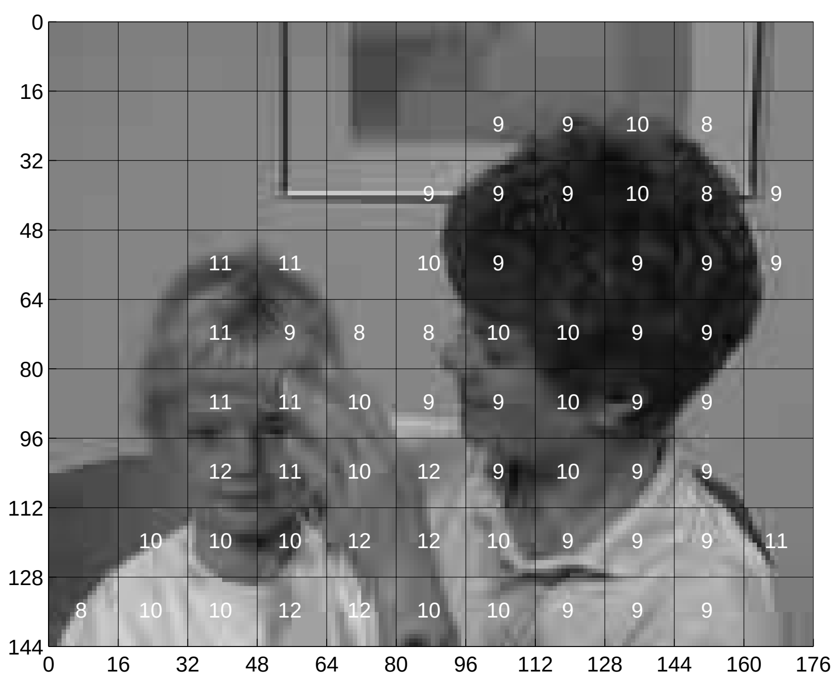

Fig. 10. Optimal quantizers in the MINAVE scheme.

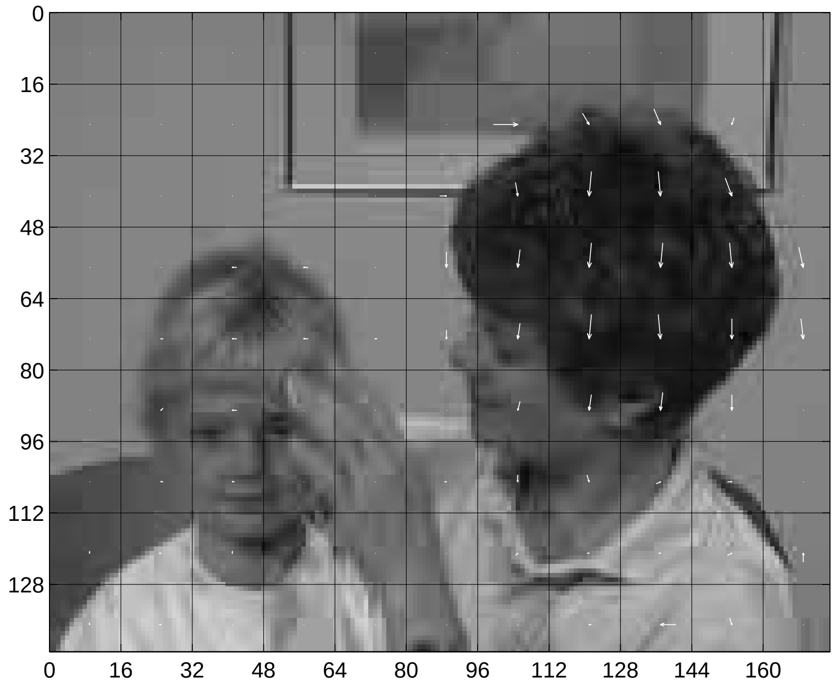

Fig. 11. Optimal motion vectors in the MINAVE scheme.

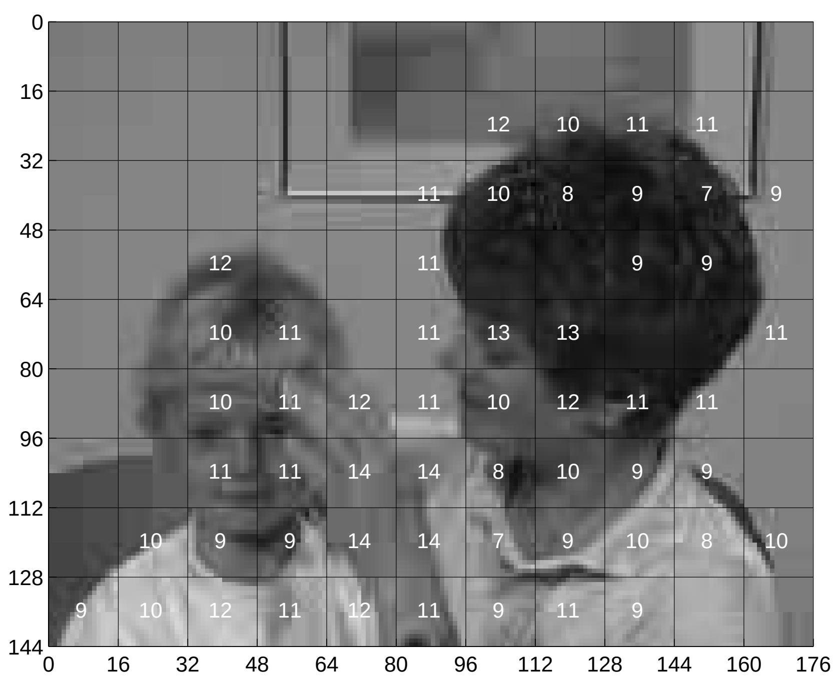

Fig. 12. Optimal quantizers in the MINMAX scheme.

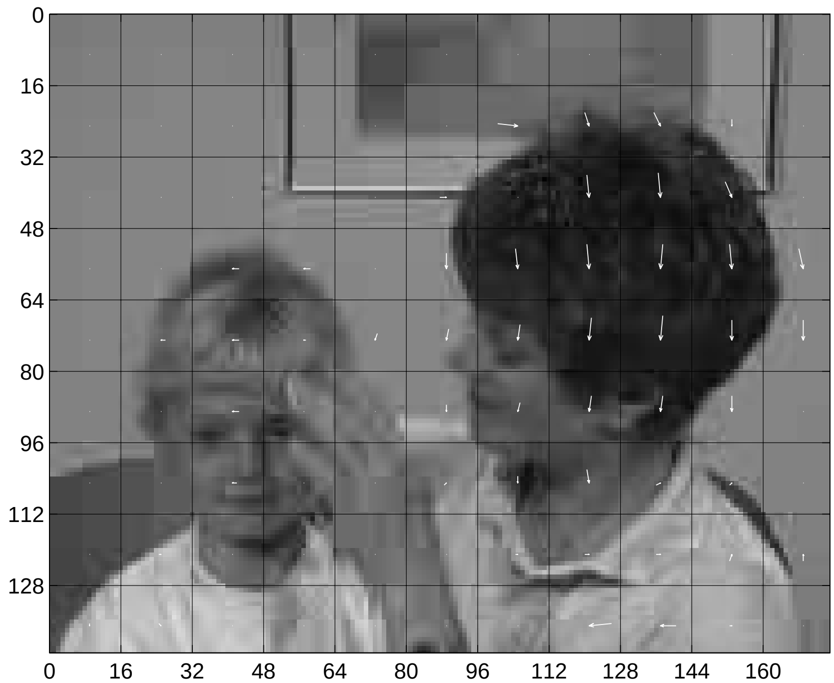

Fig. 13. Optimal motion vectors in the MINMAX scheme.

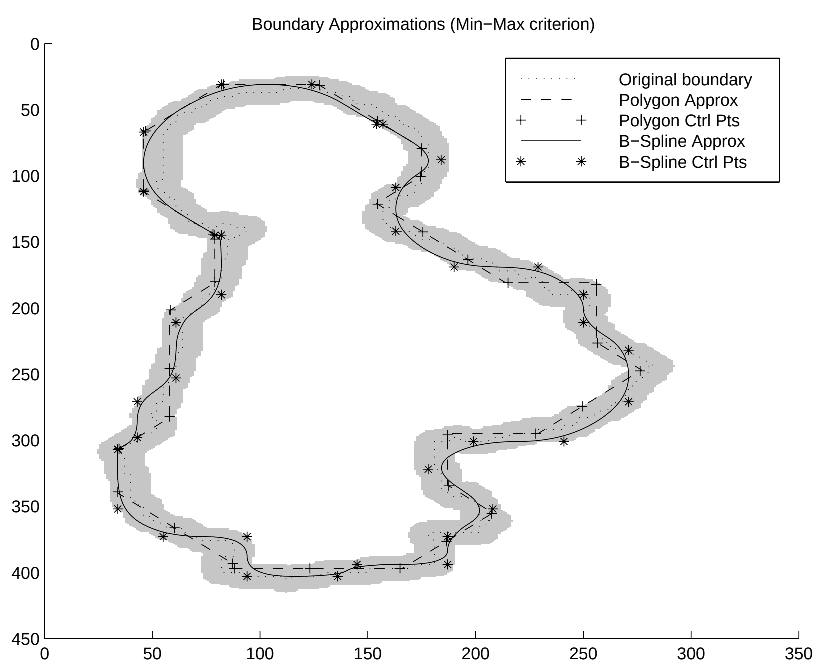

Fig. 14. Distortion band and boundary approximation using the MINMAX criterion.

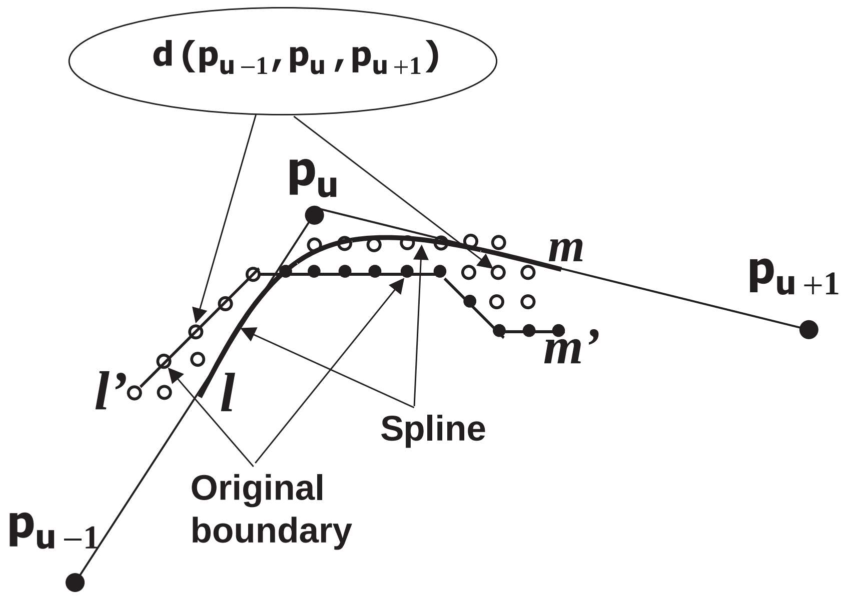

Fig. 15. Area between the original boundary segment and its spline approximation (circles).

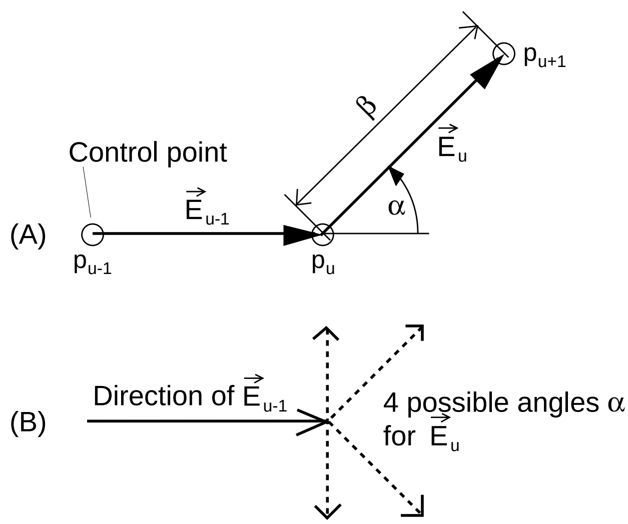

Fig. 16. Encoding of a spline control point.

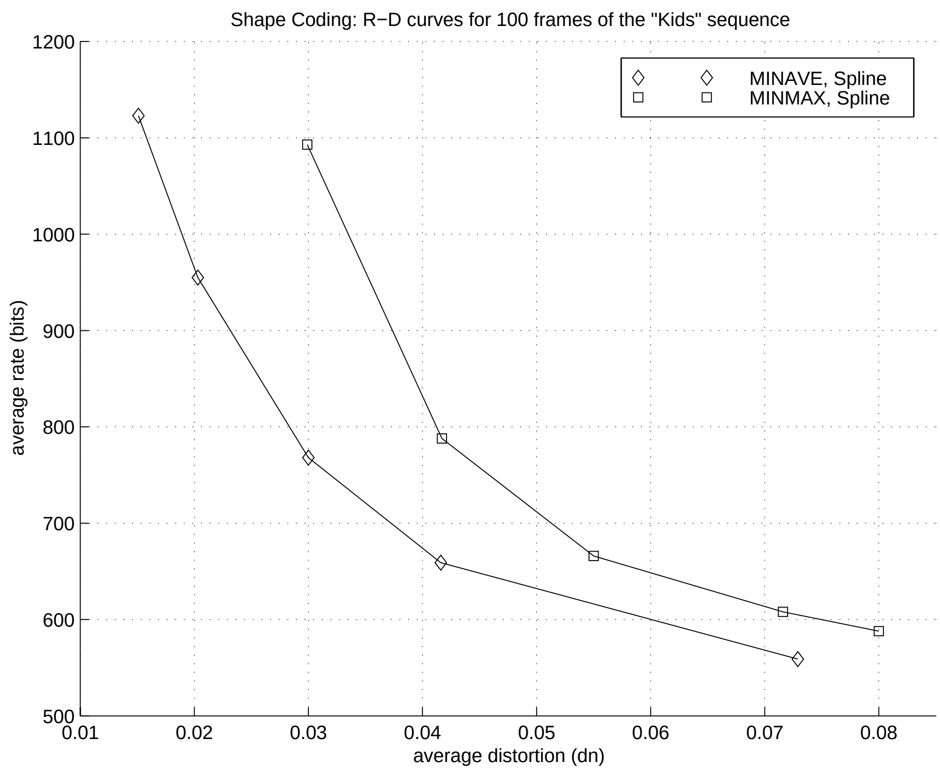

Fig. 17. Rate-distortion performance of MINAVE and MINMAX algorithms.

Fig. 18. Frame 5, kid 1 encoded with 347 bits using the MINAVE approach (white: pixels in error; grey: error-free pixels; black: background).

Fig. 19. Frame 5, kid 1 encoded with 343 bits using the MINMAX approach.