CCIT Report #599

September 2006

Transrating of Coded Video Signals via Optimized Index-modified Requantization ⋆

Michael Lavrentiev [1] and David Malah

Electrical Engineering Department
Technion - Israel Institute of Technology
32000, Haifa, Israel

Abstract

Requantization is one of the tools for bit-rate reduction of pre-encoded video to adapt it to various network bandwidth constraints. Several recent works propose using Lagrangian optimization to find the optimal quantization step, for each coded macro-block, to meet a desired rate at minimum distortion. In this paper we propose to extend the Lagrangian optimization procedure by allowing modification of quantized coefficients values, including setting their values to zero, in addition to quantization step-size selection. Thus, for each selected step-size the run-level values, which serve as indices in the VLC table, may get modified so that the overall distortion for a given overall rate is reduced. Coefficient value modification and quantization step-size selection are optimally done using a low complexity trellis-based algorithm. The proposed requantization algorithm is implemented in an MPEG-2 environment. It provides higher PSNR values than the Lagrangian-based optimization method that only handles the selection of quantization steps, and still does not exceed considerably its complexity.

Key words: MPEG-video, transcoding, requantization, trellis

⋆ This work was supported by STRIMM consortium under the MAGNET program of the Ministry of Trade and Industry via the Samuel Neaman Institute.

Email addresses: `michael_l@cbyond.co.il` (Michael Lavrentiev), `malah@ee.technion.ac.il` (David Malah).

[1] Presently at Cbyond - a Gyrus-ACMI company, 19 Hamesila St., 36885, Nesher, Israel

1 Introduction

Transrating of coded video , i.e., transcoding to the same video format at a different rate, aims at reducing the bit-rate of the encoded stream. Typical situations are channel congestion, or a need to match the encoded video bit-stream rate to a low bit-rate destination, while preserving the highest possible quality of the rate-reduced video. For simultaneous transrating of several video streams that have to be transmitted over the same Constant Bit Rate (CBR) channel, the goal can be set as to increase the mean quality of every stream by dynamically transrating the input streams according to their relative activities [1]. The naive solution of simple cascading of a decoder and an encoder is put aside because of its high computational complexity, which is mainly due to the need to re-estimate motion parameters, and because of the quality degradation. The degradation is caused by imprecise Discrete Cosine Transform (DCT)/Inverse DCT (IDCT) matching, additional quantization errors, and less accurate motion estimation in the second generation encoder that operates on lossy reconstructed video. Most of the transrating schemes are operating in the compressed domain and utilize the decisions made by the initial encoder to improve the output video quality. A common approach to bit-rate reduction in the compressed domain is requantization, carried out by increasing the quantization step-size used to quantize the DCT coefficients in each block. Several works propose low complexity open-loop transcoding [10,12], while other take advantage of error compensation provided by a closed-loop scheme [4,9]. Methods for requantization error reduction that are based on estimated statistical distribution of the DCT coefficients, were also proposed [16].

Furthermore, instead of "simple" requantization, which applies the standard complexity model and rate control of TM5 [2], to set a new quantization step-size for each macro-block (MB), several recent works proposed using Lagrangian optimization to find the optimal quantization step for each MB to meet a total desired bit-rate at minimum distortion [5,4]. It is shown in [4] that the optimally transrated bit stream provides a higher peak-signal-to-nose-ratio (PSNR) than a cascade of decoder-encoder (i.e., fully decoding the coded video and re-encoding it to the desired reduced bit-rate), and can even provide a better video quality than the standard TM5 encoder applied to the *original* video sequence at the reduced rate (this is possible because TM5 is not an optimal encoder). Lagrangian optimization was also used recently for discarding certain quantized coefficients in I-frames [10]. Dynamic programming is used to decide on which coefficients to skip, or, optionally, from which coefficient index to discard all codewords used to encode those coefficients.

Since many video encoding standards, like MPEG2 and h.263, use constant variable-length-coding (VLC) tables to encode quantized coefficients, in this paper we develop an efficient algorithm that allows modification of quantized coefficient *indices*, the principles of which were briefly presented in [3], so that shorter VLC codewords can be used, combined with minimum distortion, and test its performance when applied for MPEG2 transrating. This is done together with the selection of an optimal quantization step-size for each MB. Using a novel trellis-based optimization scheme that is specially adapted to the above problem, our scheme provides improved performance over previously proposed Lagrangian optimization over requantization step-sizes only, with only a reasonable increase in computational load.

The organization of this paper is as follows: Section 2 describes the MPEG-2 encoding procedure. Section 3 presents the pertinent Lagrangian optimization method. In section 4 we introduce the idea of modifying quantized coefficients indices, and in section 5 we present its efficient implementation using a special trellis diagram. Complexity issues are discussed in section 6. In sections 7 and 8, experimental results are shown and conclusions are drawn, respectively.

2 MPEG-2 AC coefficients encoding

Following the application of the DCT to each of 4 luminance 8×8 blocks and from 2 up to 8 chrominance blocks (depending on video format), which form a MB, the DCT coefficients (except for the DC coefficient) are quantized [8]. For each MB, a value from one of two possible tables, each having 32 quantization step-size values, is selected (a different table can be chosen for each frame). The actual quantization step-size used for each coefficient is the product of the selected step-size from the table and a value defined by a suitable quantization matrix that depends on the MB type ('Intra' or 'Inter'). The 63 quantized AC coefficients are concatenated in an order defined by one of two possible zig-zag scans. The resulting 6 to 12 vectors, of 63 quantized coefficients each, constituting a MB, are entropy coded by a variable-length-coding (VLC) table. Each coefficient vector is segmented into several parts, with each part consisting of a *run* of consecutive zeros followed by a non-zero *level* value, defining a run-level pair. In case of adjacent non-zero level values, the run length is defined to be zero. The MPEG-2 standard defines for every run-level pair a variable-length codeword. There are two VLC tables that can be used. It is possible to use the same table for all types of MBs, or to use a different one for Intra MBs [8].

3 Requantization via Lagrangian optimization

The requantization problem can be formulated as an optimization problem of determining a set of quantization step-sizes that minimize the total distortion in each frame, under a given bit-rate constraint:

$$\min_{\{q_k\}} D, \text{ under the constraints } R \leq R_T, q_k \in Q \tag{1}$$

with ,

$$D = \sum_{k=1}^{N} d_k(q_k),\ R = \sum_{k=1}^{N} r_k(q_k), \tag{2}$$

where,

N - number of MBs in the frame;

q_k - quantization step-size for the k-th MB;

d_k - distortion caused to the k-th MB;

r_k - number of bits produced by the k-th requantized MB;

Q - set of allowed values of quantization step-size.

A general solution for the allocation of a given quota of bits to an arbitrary set of different quantizers was derived in [18]. An analysis for the conventional MSE distortion metric in video transrating is presented in [4]. The problem can be converted into an unconstrained one by merging rate and distortion through a Lagrange multiplier $\lambda \geq 0$ into the cost function:

$$J_{total} = D + \lambda R, \tag{3}$$

where λ defines the relative importance of rate against distortion in the optimization procedure.

The Lagrangian cost can be independently calculated for each MB [18,4]. Thus, for the k-th MB:

$$J_k(\lambda) = \min_{q_k}\{d_k(q_k) + \lambda r_k(q_k)\}. \tag{4}$$

Let $\{r_k(\lambda), d_k(\lambda)\}_{k=1}^{N}$ be the set of solutions for a particular λ that achieves the minimum Lagrangian cost for every MB in the frame. If for a particular value $\lambda = \lambda_s$ the total rate $R(\lambda_s) = \sum_{k=1}^{N} r_k(\lambda_s)$ is equal to the target rate R_T , then the set $\{q_k(\lambda_s)\}_{k=1}^{N}$ is the optimal set of quantizer step sizes to be used for transrating. λ_s has to be found for every frame - if the problem is solved on a picture level, or for every slice - if bit-rate allocation is provided on that level. If the total rate is bigger/lower than R_T, than λ is increased/decreased, and the minimization in (4) is repeated with the new value of λ.

It is shown in [4], by simulations, that by using the same set of coding decisions as produced by the initial encoder, optimal requantization can achieve higher PSNR than that achieved by direct encoding of the original video sequence to the final lower bit-rate by the usual encoder. This is possible because common encoders do not perform this kind of optimization, for finding the best possible quantization steps, to avoid an increase in the encoding complexity.

4 Modification of quantized coefficients indices

The first steps in modifying the levels of quantized DCT coefficients before applying VLC, for bit-rate reduction, were taken in [10,7]. However, [10] discusses only methods for excluding AC coefficients in I-frames, and [7] considers only discarding several last non-zero coefficients in the zig-zag scan. We propose here to extend the Lagrangian optimization presented in the previous section to allow the modification of the values of all quantized DCT coefficients in an efficient way. The suggested optimization procedure aims at choosing quantized AC coefficient vectors values, as well as optimal quantization step sizes, that will provide a bit-rate that is as close as possible to the desired rate with minimal distortion.

Direct encoding of requantized coefficients using a given fixed run-level coding table does not necessarily provide the minimum possible distortion for a given total rate. It is possible to reduce the bit-rate by changing the values of each quantized vector elements before run-level coding. To preserve the best possible quality at a given bit-rate, the selection algorithm uses a penalty for distortion caused by selecting reconstructed values away from the optimal one. An improvement, as compared to selecting optimal quantization step sizes only, is expected due to the following reason: It is possible to reduce the total bit-rate by breaking VLC pairs with long runs into several smaller ones that give a smaller total bit-rate than that of the initial pair, with minimal increase in the distortion. This way it is possible to achieve a desired bit-rate reduction with smaller total quality reduction.

Table 1 gives an example of the bit allocation for some run-level pairs in an MPEG-2 VLC table. For example, it is possible to change the number of bits needed to encode run-level pair (2,6) from 24 bits to 14 bits by substituting it by the pair (2,5), with a distortion increase by the square of just one quantization step. If one decides to split run-level pair (7,3), which takes 24 bits to encode, into (6,1) and (0,3), one will need 7 plus 6 bits to encode the two

Table 1
Number of bits needed to encode some run-level pairs according to one of MPEG-2 VLC tables.

level \ run	0	1	2	3	4	5	6	7
1	3	4	5	6	6	7	7	7
2	5	7	8	9	11	11	13	13
3	6	9	11	13	13	14	17	24
4	8	11	13	14	24	24	24	24
5	9	13	14	24	24	24	24	24
6	9	14	24	24	24	24	24	24
7	11	14	24	24	24	24	24	24

pairs, saving 11 bits for an increase in distortion by the square of just one quantization step at the location of the split.

As in the optimization problem stated in section 3, we may select a different quantization step-size for each MB, but here we also allow changing the quantized DCT coefficients values *by modifying their level values (quantization indices)* after a particular quantization step-size has been selected. The minimization problem stated in (1) remains the same, but now (2) is replaced by:

$$D = \sum_{k=1}^{N} d_k(q_k, \mathbf{v}_k),\ R = \sum_{k=1}^{N} r_k(q_k, \mathbf{v}_k), \tag{5}$$

where $\mathbf{v}_k$ denotes the index vector. All other parameters remain the same as in (2). The problem is still separable at the MB level - like in (4). But now, for every q_k, an additional minimization over all possible $\mathbf{v}_k$ values must be performed.

The same problem statement can be used for direct encoding of the original image, except for the fact that in transcoding we use an approximation of the original signal from previously encoded data.

Distortion is usually measured by Squared Error from the source video sequence in the pixel domain. In transrating applications there is no source video available, so distortion is calculated relative to the originally encoded stream, or to some estimation of original video stream from the initially encoded one, as was proposed in [16] . By the orthonormality of the DCT we calculate the distortion directly from the DCT coefficients, without performing full decoding. Again, for every MB, in addition to minimizing over q_k, a minimization over all possible $\mathbf{v}_k$ values must also be performed. Thus, for the k-th MB, (4) takes the following form:

$$J_k(\lambda) = \min_{q_k} \min_{\mathbf{v}_k} \{d_k(q_k, \mathbf{v}_k) + \lambda r_k(\mathbf{v}_k)\}, \tag{6}$$

and the set $\{\mathbf{v}_k(\lambda_s)\}_{k=1}^{N}$ is the optimal set of quantized coefficients index vectors that provides the minimum distortion for a given total rate constraint, R_T. Since the algorithm modifies $\mathbf{v}_k$ directly, and not via q_k, r_k is not an explicit function of q_k in (6). An efficient solution for the stated problem is presented in the next section.

5 Trellis-based optimization

In this section a Trellis-based implementation of the above Lagrangian optimization procedure is presented and discussed.

Suppose that particular values of λ and q_k have been chosen for the k-th MB. Let's define each position in the zig-zag scanned quantized DCT coefficients vector as a different stage in a trellis (Fig. 1). The quantized coefficient index values we may choose from, define the states of each stage. A path in the trellis is defined as a particular choice of one state for each stage. It can be presented as a chain of run-level pairs. For each run-level pair the distortion d and the number of bits r needed for encoding are known, so it is possible to define the cost of a run-level pair as $d + \lambda r$. The cost of a path is the sum of the costs of run-level pairs defined by this path. The optimal path up to a particular stage is the path that has the minimal cost value over all possible paths ending at that stage. The essence of a trellis-based algorithm is the fact that minimization of the cost value at each state of the current stage is the minimization of the sum of the current stage local-cost at each state and the minimal path cost already calculated at the previous stages of the trellis. It turns out that for the current problem, in which different run-lengths need to be considered, the conventional trellis needs to be modified, so that every decision in a given stage may depend on previous stages. Luckily the dependence is only on a single, already determined, state in each previous stage, as described below.

Fig. 1 shows how the cost function is evaluated for a particular stage in the trellis. For trellis stage i (corresponding to the i-th coefficient) we have states from zero to $v\_max(i)$. $v\_max(i)$ is determined by multiplying the original index value by the original quantization step-size (both obtained from the partially decoded stream), and dividing by the new quantization step-size, followed by rounding upwards. In general, every possible v, $0 < v \leq v\_max(i)$, should be examined to see if it minimizes the total cost function $J(v, i)$ in (7) below. This cost depends not only on the value of v, but also on the number of

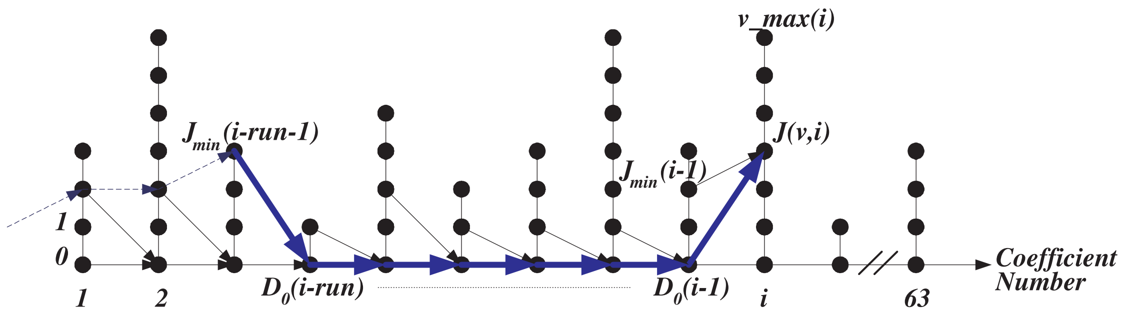


Fig. 1. Trellis diagram with search patterns for value v in i-th AC coefficient in zig-zag ordered quantized coefficients vector of DCT transformed 8x8 block.

zeros, i.e., the *run* leading to it, which defines the run-level pair for the VLC:

$$J(v,i) = \min_{run}\{J_{min}(i-run-1) + \\ + \sum_{j=i-run}^{i-1} D_0(j) + \lambda\rho(run,v) + D(v,i)\} \quad (7)$$

where,

$J_{min}(i-run-1)$	- the cost of the minimal path up to the stage $(i-run-1)$
$D_0(j)$	- the distortion caused by zeroing the j-th DCT coefficient
$D(v,i)$	- the distortion introduced by choosing v to be the index value of the i-th quantized coefficient
$\rho(run,v)$	- the number of bits needed to encode the run-level pair (run,v) using the VLC

The dotted thin line at the left side of Fig. 1 shows the minimal path till stage $i-run-1$, which has the minimal cost $J_{min}(i-run-1)$. Thin arrows connect the last values of optimal paths in the previous stages to zero; or, in the case of stage $i-1$, directly to the value v in stage i that is being examined. Different run-lengths need to be examined, but for a particular run the optimal state in the preceding stage $i-run-1$ is already known. The heavy line indicates the optimal path for the particular value of v. To determine the optimal value of v, the minimum cost over all its possible values of v has to be found:

$$J_{min}(i) = \min_{0<v\leq v\_max(i)} J(v,i) \quad (8)$$

Fig. 2 shows a schematic view of the trellis diagram to be used in the optimization algorithm on a Macro-block level. The trellis is built for every

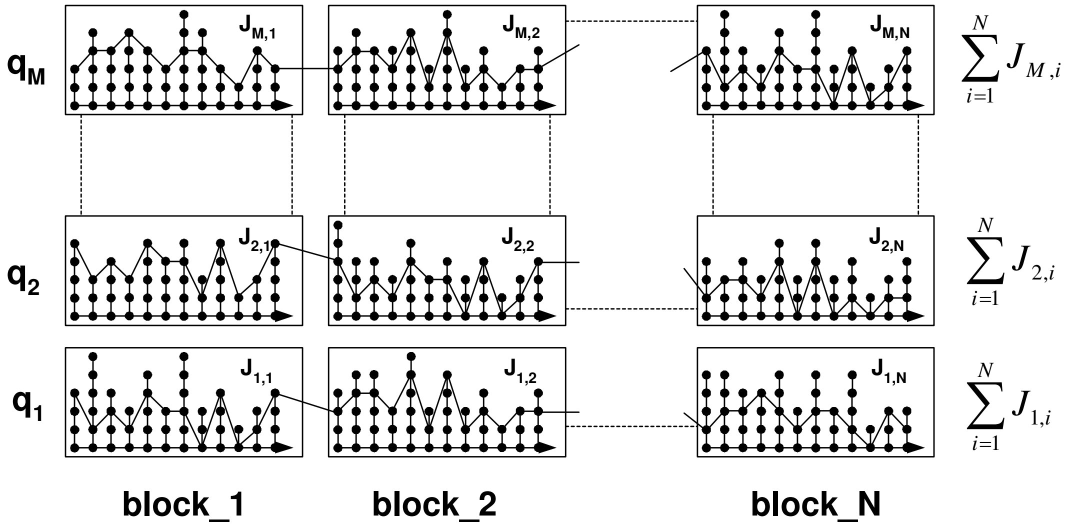


Fig. 2. Trellis algorithm on Macroblock level.

considered requantization step-size, presented as block rows in Fig. 2. Corresponding step-sizes are denoted as q_i, where $\{q_i\}_{i=1}^{M} \in Q$, where Q denotes a set of permissible quantization step-sizes. The block-row trellis is divided into sub-trellises on a block level. For a particular block and a step-size of q_i the range of quantized AC coefficient index values is built (dotted columns). The optimal path (thin lines) throughout a block is calculated for a particular λ as was described above. For a fixed λ and quantization step q_i, the cost function $J_{i,j}$ can be calculated independently for each block j in the particular MB. At the end of each block the best vector quantized is chosen in every row, which is equivalent to a trellis path with minimal cost for the particular block. For the next block, the optimal path pointer of the previous block trellis is provided, shown by the line connection to the preceding block. At the end of each macro-block the optimal quantization step size is chosen (for the given λ):

$$q_{k,opt}(\lambda) = \arg\min_{\{q_i\}_{i=1}^{M} \in Q} \sum_{j=1}^{N} J_{i,j}(\lambda, q_i) \tag{9}$$

After solving the optimization problem for all MBs, the total bit-rate is calculated and compared with the target bit-rate constraint, and λ is changed accordingly. This process is repeated until the target bit-rate is reached, or closely approximated.

Yet, even when the above trellis is used, the number of calculations needed to perform the optimization is rather high. Hence, in the next section we consider ways to speed up the algorithm.

6 Complexity Considerations

The discussed method needs, in principle, many iterations over a large number of parameter values because:

(1) The number of examined *run* values for every *level* in a particular stage increases with the index value of the DCT coefficient being processed.
(2) The computations in all stages, till the stage corresponding to the last non-zero coefficient, in every block, need to be carried out.
(3) A separate trellis has to be constructed for every requantization step-size that we wish to examine.
(4) There are several index values v at each stage that need to be examined.
(5) Several values of λ need to be tried (in a directed way) before the total rate will match the constraint.

Note, however, that while the number of index values (levels) to be examined at each stage seems to be large at first sight, it is no so in reality. This is because the mean value of the AC coefficients is typically in the range of 30-50. Hence, for example, if the initial quantization step-size is 6, then even on the finest scale there are on average only about 5÷10 values to choose from. When the quantization step is increased, we reach a single value very quickly.

As for quantization step-sizes, there were a number of works that propose to restrict the range of requantization step-sizes based on the initial MB quantization step-size [11,9,12].The most recent [12] shows that for open-loop transcoding, Lagrangian optimization [6] can be restricted to step-sizes that are even multiples of the initial quantization step-size for I frames and odd multiples for P-B frames, respectively. At those step-sizes many quantized coefficients are zeroed out after rounding. However, in a closed-loop scheme this result does not hold because of the error compensation.

As for searching over different values of λ, applying a simple bi-section search, as in [4], requires, on average, about 3 iterations only.

6.1 Complexity Reduction

As mentioned above, the number of examined *run* values for every *level* in a particular stage increases with the index value of the DCT coefficient being processed. Let's have a look at one of the MPEG-2 VLC tables presented in Fig. 3. We observe that, practically, the number of level values that should be considered for obtaining a rate reduction is actually not that large. All pairs that need 24 bits to encode are out of the destination scope (presented by white rectangles on Fig.3), as they do not allow bit-rate reduction. Moreover, if we consider choosing a run for a particular level v, the number of options to examine - $run_{max}(v)$, before getting to the maximum no. of bits in the VLC table, $R_{max} = 24$, is very small for most levels (the exceptions are levels 1 and 2, for which there are 31 and 16 possible runs, respectively).

Thus, (7) can be rewritten as follows:

$$J(v,i) = \min\{J_1(v,i), [J_2(run_{max}(v),i) + \lambda R_{max}]\} + D(v,i) \qquad (10)$$

where,

$$J_1(v,i) = \min_{run<run_{max}(v)} \{J_{min}(i-run-1) + \sum_{j=i-run}^{i-1} D_0(j) + \lambda R(run,v)\} \quad (11)$$

is the part that depends on the (*run, level*) values chosen and thus needs to be evaluated for every *run*, and

$$\begin{aligned} J_2(run_{max}(v),i) &= \min_{run\geq run_{max}(v)} \{J_{min}(i-run-1) + \sum_{j=i-run}^{i-2} D_0(j)\} + D_0(i-1) \\ &= J_2(run_{max}(v)-1, i-1) + D_0(i-1) \end{aligned} \qquad (12)$$

is the part known from previous calculations. $J_2(m,i)$ is determined by the

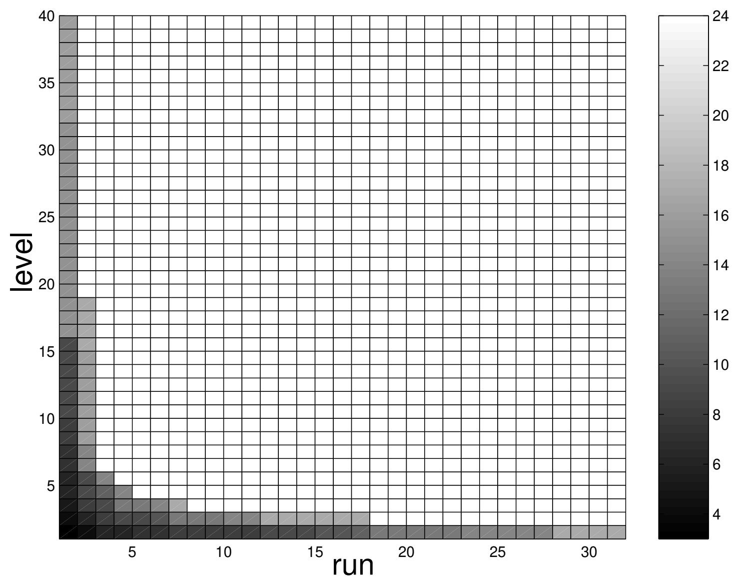


Fig. 3. One of two possible MPEG-2 VLC tables. Number of bits needed to encode each run-level pair is presented by different gray levels. The grayscale on the right relates each gray level to the number of bits and vice a versa.

following recurrent formulae:

$$
\begin{aligned}
&J_2(0,0) = 0; \\
&J_2(m,i) = \begin{cases} J_2(m-1,i-1) + D_0(i-1) & ,\, m < i \\ \min\{J_2(i-1,i), J_{min}(i)\} & ,\, m = i \end{cases}
\end{aligned}
\qquad (13)
$$

To calculate $J_1(v,i)$, $run_{max}(v)$ iterations are needed, while $run_{max}(v)$ is usually a small number. $J_2(m,i)$ is calculated by (13) for all $m \leq i$ to be used by next iterations. However, following (13) it is only adding the same value $D_0(i-1)$ to all the results from the previous stage. There is no need for $J_2(i,i)$ on i-th stage, so it can be calculated by one comparison after $J_{min}(i)$ is found. So, $J_2(run_{max}(v)-1, i-1)$ is available from the last stage, and no search is needed to find $J_2(run_{max}(v), i)$. To get $J(v,i)$, only $run_{max}(v)+1$ comparisons are needed. Using (10), (11) and (12) instead of (7) *does not affect the optimality of the solution*, but reduces the number of calculations needed by up to 40% - in our simulations.

As mentioned in section 4, it may be useful sometimes to split a run-level pair into two pairs with shorter runs in order to reduce the total bit-rate at the cost of additional distortion. It is reasonable to just substitute one the of zero values by a level value = 1 because it provides a lower distortion than with any other non-zero value. At the same time, this selection provides the possibility of a higher rate reduction because the run-level pair cost in bits is an increasing function of level. It is difficult to predict where in the zig-zag ordered vector this substitution will be the optimal one because of the weight matrix that modifies the quantization step of each AC coefficient. If one wants to use a distortion metric different from MSE, things will become even more complicated.

Thus, a useful simplification that reduces complexity is to disallow splitting. I.e., the trellis shown on Fig. 1 is reconstructed to exclude the stages with initially zeroed values. This is, of course, a sub-optimal solution, but is found to have a very small effect on performance, as reported in the sequel. In typical MPEG-2 encoded blocks about 70-90 % of the quantized coefficients are zeros, so the computational complexity reduction is very pronounced - about 60 % in our simulations. This simplification reduces the complexity at the price of being a sub-optimal solution.

7 Experimental Results

In this section we compare the performance of the proposed scheme to several other transrating schemes. We report here the results obtained for the following video sequences: FOOTBALL, MOBILE, GARDEN (in SIF 4:2:0 format, 45 frames) and FOREMAN (in CIF 4:2:0 format, 45 frames). The sequences were encoded using a standard TM5 encoder [2] at 4Mbps. This rate was reduced by transrating to rates varying from 3 to 1Mbps. In all trasrating schemes we used the simplest frame-level scheme [19], which divides the new bit-budget among frames using the same ratio that they get in the input stream:

$$\frac{R_{out}}{R_{in}} = \frac{T_{out}}{B_{in}} = const \tag{14}$$

where:

R_{out} - desired output average bit-rate

R_{in} - input sequence average bit-rate

B_{in} - bits spent on current frame in input stream

T_{out} - target bit allocation for the transrated frame

This approach is straight-forward. It does not provide any problem with virtual buffer fullness - the buffer size can be simply decreased by the same ratio that the stream rate goes down. Transrating approaches differ in how they achieve target bit allocation within each frame. The average PSNR values obtained by each of the following transrating scheme are shown in Fig. 4:

(1) "Simple" requantization that uses TM5's complexity model to transrate each frame (denoted 'Sim'). This method is explained in detail in Appendix A.
(2) Lagrangian optimization that finds the optimal set of quantization step-sizes for each frame ('Lag'), like in [4].
(3) Proposed Trellis-based optimization ('Tr').
(4) Proposed reduced complexity Trellis-based optimization ('TrNZ'), i.e., trellis diagram is modified to include initially non-zero stages only, since no splitting of runs is allowed, as described at the end of subsection 6.1.

The original video sequence and the decoded 4Mbps video were also encoded to the desired bit-rates using a standard TM5 encoder. They are denoted as 'Enc' and 'Re', respectively. In Table 2 we summarize results for all the above sequence, transrated from 4Mbps to 2Mbps:

Table 2
Results measured for different sequences, transrated from 4Mbps to 2Mbps

Method	Run-time ratio	PSNR Improvement over Re-encoding (dB)			
		FOOTBALL	MOBILE	FOREMAN	GARDEN
'Sim'	0.15	0.3	0.45	0.5	0.45
'Lag'	1	0.9	1.05	1.4	1.05
'Tr'	7.9	1.3	1.55	1.75	1.5
'TrNZ'	3.3	1.2	1.45	1.65	1.4

Measuring complexity in terms of run-time, we obtained in our implementation that: 'Lag' runs about 6 times slower than 'Sim' transrating, and 'Tr' runs about 8 times slower than 'Lag', while 'TrNZ' is just 3 times slower than 'Lag'. On FOREMAN sequence 'Tr' outperforms simple re-encoding by 1.75 dB.

It is seen from the Fig.4 (a) that, for FOOTBALL sequence, 'Tr' outperforms 'Enc' at all the rates by about 0.62 dB, in average PSNR, while the gain of 'Lag' over 'Enc' is only about 0.28 dB. The difference in PSNR between 'Tr' and 'Sim' transcoding is about 0.95 dB.'TrNZ', which is sub-optimal but of greatly reduced complexity, suffers a loss of only 0.07 dB as compared to 'Tr'. For other video sequences we got similar results with the proposed Trellis-based schemes, relative to the Lagrangian scheme. The other methods show more significant variations, but always have lower PSNR than the proposed Trellis-based schemes , as can be seen from Fig. 4 (b),(c) and (d).

The variation of the PSNR from frame to frame along the FOOTBALL sequence, when transrating from 4Mbps to 2Mbps by 'Tr', 'Lag' and 'Sim' schemes, is shown in Fig. 5 (a). The results of encoding the original sequence ('Enc') at 2Mbps is also added for comparison. Fig. 5 (b) shows the performances of 'Enc', 'Re' and 'Tr'.

8 Conclusions

This paper presents an extension of the Lagrangian optimization of quantization step-sizes only [5,4,6], by allowing the modification of quantized AC coefficients indices. This method consistently results in better performance than other known schemes, including even those obtained by encoding the original video sequence at the reduced rate with a standard coder. The cost of the proposed scheme is in its complexity. A Trellis-based Lagrangian optimization for MPEG-2 encoded video transrating has been developed. In this approach, at each trellis state we need to check one state in each previous stage. The fact that when the run-level pair encoding length reaches the maximum, the optimal decision is known from previous stages, has provided an additional complexity reduction that has not affected the optimality of the solution. To further reduce the complexity, a sub-optimal solution that does not allow run-level pair splitting, was examined and found to reduce significantly the complexity with a very small reduction in performance. The run-time of the optimal extended Lagrangian scheme is about 8 times the complexity of the algorithm proposed in [5,4,6], and it is reduced to about 3 times only by the sub-optimal solution.

In terms of average PSNR the improvement obtained over the cascading of the decoder with a standard TM5 encoder (scheme 'Re' in section 6) is about 1.5 dB. The sub-optimal solution results in a loss of gain of less then 0.1 dB in average PSNR. The new scheme even provides better results (in average PSNR) than encoding of the original video sequence at the desired lower rates with a standard TM5 encoder. This is possible because the TM5 encoder is not optimal, in rate-distortion sense.

In this work we have focused on rate reduction of MPEG-2 encoded bitstreams. However, the proposed trellis-based algorithm can be applied to other coding standards that use run-level VLC. It can also be used to build an encoder that will outperform the TM5 encoder (at the cost of encoder complexity).

Appendix A "Simple" requantization

The 'Simple' transrating method is based on a variation of the picture complexity measure, proposed in [20] for pictures in a GOP. In the present work this notion is adapted to MBs in a picture. The original definition of picture complexity from [20] is given below, followed by the description of out modification. Finally, the quantization step modification method for 'Simple' transrating is described.

Picture layer bit-budget allocation in a GOP that is based on picture complexity is proposed in [20]. Picture complexity is defined as the product of the average quantizer step-size and the number of bits generated, divided by some empirical constant:

$$X_I = S_I Q_I, X_P = \frac{S_P Q_P}{K_P}, X_B = \frac{S_B Q_B}{K_B} \tag{A.1}$$

where

X_I, X_P, X_B	- Estimated picture complexity
	of I-,P- and B-frames, respectively
S_I, S_P, S_B	- number of bits generated by encoding
Q_I, Q_P, Q_B	- average quantization step sizes used in encoding
K_P, K_B	- universal constants (1.0 and 1.4)

To estimate the output picture complexity, it is proposed in [20] to multiply the ratio of output/input complexities in the previously transrated frame by input complexity of the current transrated frame.

We adapted the above definition of picture complexity to MB level. Requantization reduces not only the bit-rate, but the MB and picture complexity as well. To further simplify the problem, we assume that the complexity reduction factor is the same for all kinds of MBs in the same frame. In this case the ratio of output complexities must remain the same as it was for input complexities, and we can estimate the output bit-rate of MBs left to encode by the bit-allocation of the last encoded MB, and update the quantization step-size accordingly:

$$q_{2,n+1}^{ind} = \begin{cases} q_{2,n}^{ind} + 1 \ if \ \hat{B}_{out} > B_{out} \\ q_{2,n}^{ind} - 1 \ if \ \hat{B}_{out} < B_{out} \\ q_{2,n}^{ind} \quad if \ \hat{B}_{out} = B_{out} \end{cases}, \ \hat{B}_{out} = \frac{\sum_{n+1}^{N} X_k}{q\_tab[q_{2,n}^{ind}] \cdot B_n} \qquad (A.2)$$

where

B_n	-	bit allocation of the last transrated n-th MB
$\hat{B}_{out}$	-	estimated bit number needed to encode MBs left
B_{out}	-	bits left to encode the rest of the frame
$q_{2,n}^{ind}$	-	index of last quantization step-size used
$q\_tab[\,]$	-	quantization step-size table used to get
		appropriate quantization step-size value
X_k	-	Complexity of k-th MB

The above scheme is called here 'Simple' transrating.

References

[1] H. Sorial, W. E. Lynch, A. Vincent "Joint transcoding of multiple MPEG video bitstreams", *Circuits and Systems, Proceedings of the 1999 IEEE International Symposium on*, vol. 4,pp. 251-254, 1999.

[2] ISO/IEC-JTCI/SC29/WG11, MPEG 93/457, Test Model 5, April 1993.

[3] M.Lavrentiev and D. Malah "Transrating of MPEG-2 coded video via requantization with optimal trellis-based DCT coefficients modification", *EUSIPCO-2005, European Signal and Image Processing Conference, Proc.*, pp.1963-1966 , Sep. 2004.

[4] A. A. Assunção and M. Ghanbari "Optimal bit rate conversion of MPEG-2 video bit streams", *Electronics Letters*, vol. 33(8), pp. 675-677, Apr. 1997.

[5] A. A. Assunção and M. Ghanbari "Optimal Transcoding of Compressed Video", *ICIP'96, IEEE International Conference on Image Processing, Proc.*, vol. 1, pp. 739-742, Apr. 1996.

[6] A. A. Assunção and M. Ghanbari "A frequency-domain video transcoder for dynamic bit-rate reduction of MPEG-2 bit stream " *Circuits and Systems for Video Technology, IEEE Transactions on* , vol. 8, pp. 953-967, Dec. 1998.

[7] A. Eleftheriadis and D. Anastassioiu, "Constrained and General Dynamic Rate Shaping of Compressed Digital Video", *ICIP'95, IEEE International Conference on Image Processing, Proc.*, vol. 3, pp. 396 -399, 1995.

[8] ISO/IEC 13818-2 "Information technology - Generic coding of moving pictures and associated audio information - Part 2: Video", 1995.

[9] S. Kadono, M. Etoh and N. Yokoya "Rationality of restricted re-quantization for efficient MPEG transcoding",*ICIP'00, International Conference on Image Processing, Proc.*, vol. 1, pp. 952-955, 2000.

[10] R. L. Lagendjik, E. D. Frimout and J. Biemond "Low-complexity rate-distortion optimal transcoding of MPEG I-frames", *Signal Processing: Image Communication*, vol. 15, pp. 531-544, 2000.

[11] K. Wang and J. W. Woods "MPEG2 Transcoding", *CDVMR PR 98-01, Center for Digital Video & Meadia Research Presentation, Rensselaer Polytechnic Institute, Troy, NY*, 1998.
http://www.ecse.rpi.edu/CNGV/publications/1998_PR/iab_talk_jw.pdf

[12] A. Leventer and M. Porat "Towards Optimal Bit-Rate Control in Video Transcoding", *ICIP'03, IEEE International Conference on Image Processing, Proc.*, vol. 3, pp. 265-268, 2003.

[13] H. Sorial, W. E. Lynch and A. Vincent "Estimating Laplacian parameters of DCT coefficients for requantization in the transcoding of MPEG-2 video", *ICIP'00, IEEE International Conference on Image Processing, Proc.*, vol. 1, pp. 956-959, 2000.

[14] B. -U. Lee "Modeling quantization error from quantised Laplacian distribution", *Electronics Letters* , vol. 36, no. 15, pp. 1270-1271, 2000.

[15] G. Zihua, O. C. Au and K. B. Letaief "Parameter Estimation for Image/Video Transcoding", *ISCAS 2000 Geneva. The 2000 IEEE International Symposium on Circuits and Systems, Proc.*, vol. 2, pp. 269-272, 2000.

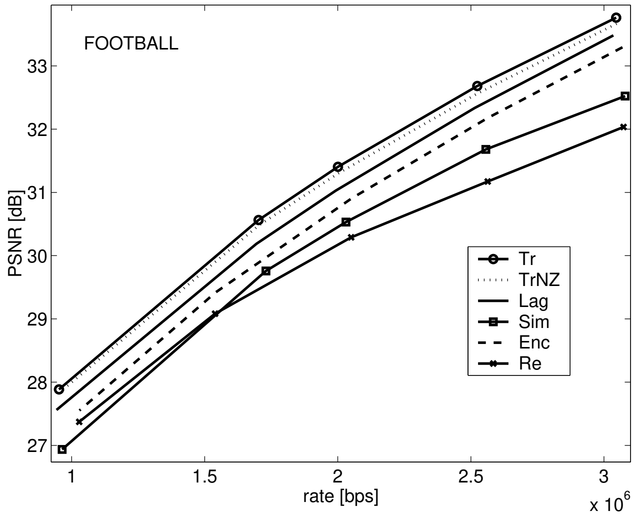
[16] O. Werner "Requantization for transcoding of MPEG-2 Intraframes", *Image Processing, IEEE Transactions on*, vol. 8(2), pp. 179-191, 1999.

[17] R. C. Reininger and J. D. Gibson "Distribution of the Two-Dimensional DCT Coefficients for Images", *IEEE trans. on Communication*, vol. COM-31, no. 6, pp. 835-839, 1983.

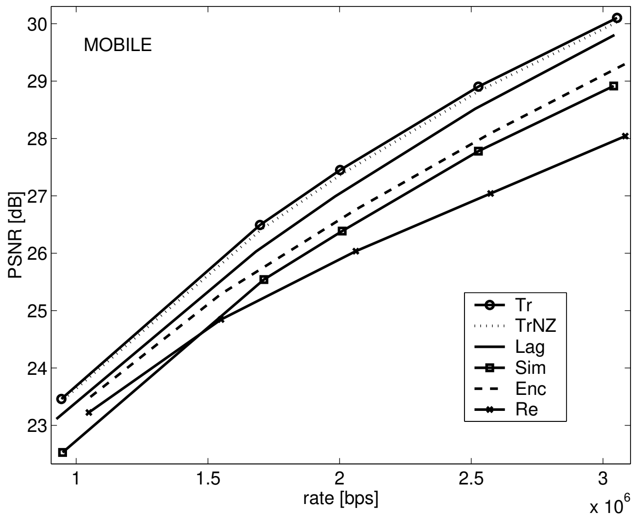
[18] Y.Shoham and A.Gersho. "Efficent bit allocation for an arbitrary set of quantizers", *IEEE Trans. Acoust. Speech, and Signal Proc.*, vol 36(9), pp.1445-1453 ,1988.

[19] J. Xin, M.-T. Sun and K. Chun "Bit-allocation for transcoding of pre-encoded video streams", *SPIE, Visual Communications and Image Processing, Proc.* vol. 4671, pp. 164-171, 2002.

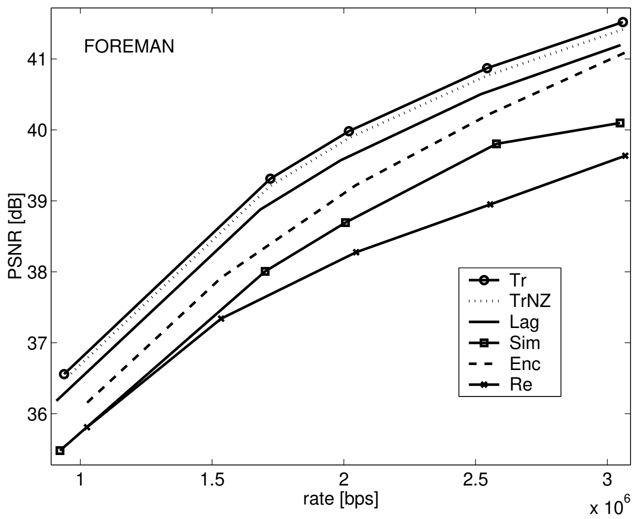
[20] W.-N. Lie and Y.-H. Chen "Dynamic rate control for MPEG-2 bit stream transcoding", *ICIP'01, IEEE International Conference on Image Processing, Proc.*, vol. 1, pp.477-480, 2001.



(a)



(b)



(c)

Fig. 4. Average PSNR vs. rate for (a)FOOTBALL, (b)MOBILE and (c)FOREMAN sequences transrated from 4Mbps into lower rates.

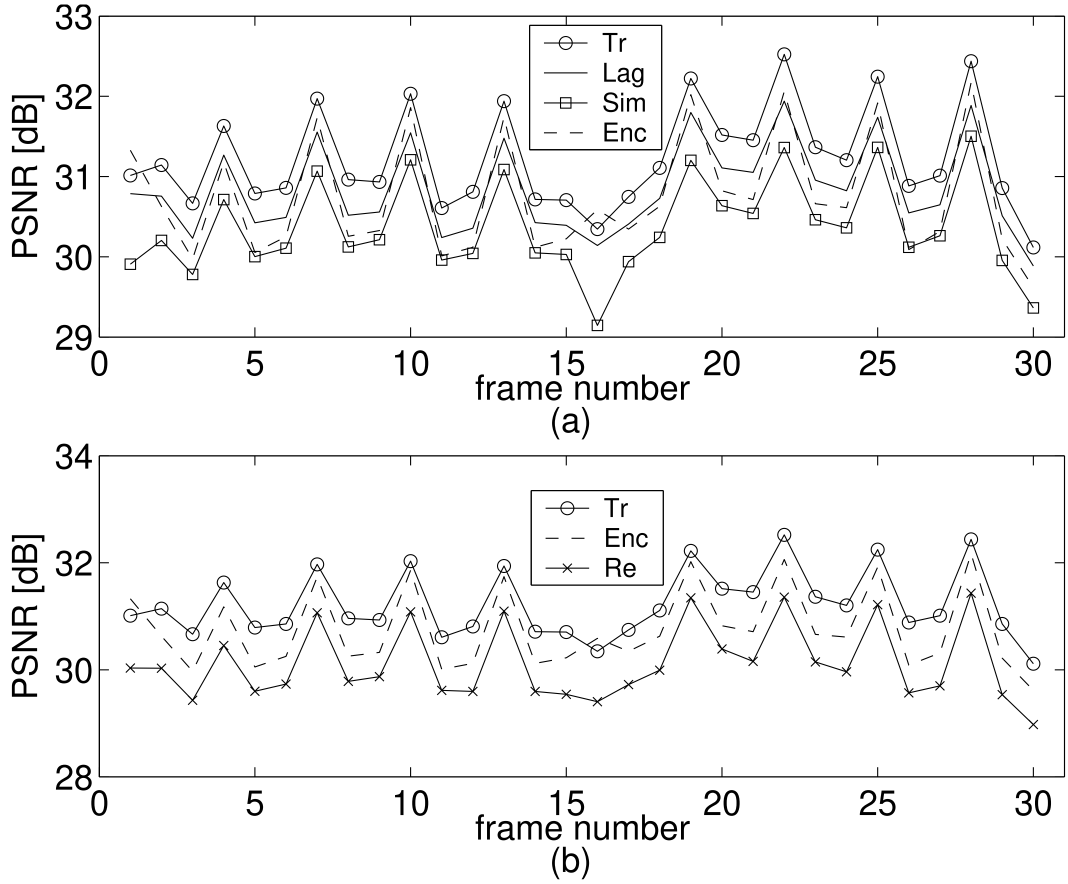


Fig. 5. PSNR as function of frame number for FOOTBALL sequence, transrated from 4Mbps into 2Mbps, for different transrating algorithms.