Video Coding with 3-D Dynamic Resolution Conversion and Rate-Distortion

Lung-Jen Wang

Department of Computer Science and Information Engineering, National Pingtung Institute of Commerce
51 Min Sheng E. Road, Pingtung 900, Taiwan, R.O.C.
ljwang@npic.edu.tw

Abstract- At low bit rates, the quality of coded video can be achieved by the dynamic resolution conversion (DRC) method for the MPEG-4 standard. In this paper, it is shown that a new three-dimensional (3-D) DRC method can be used with the MPEG-4 algorithm to improve the MPEG-4 DRC method for low bit-rate video coding. In addition, this 3-D DRC method uses a 3-D down-sampling in the encoder and a 3-D up-sampling in the decoder, respectively, for the MPEG-4 algorithm. Furthermore, a novel optimal algorithm is utilized to solve the rate-distortion (R-D) problem that joints the 3-D DRC down/up sampling and adaptive quantization for video coding. Experimental result shows that the proposed MPEG-4 with 3-D DRC and adaptive quantization obtains a better subjective quality and objective PSNR of the reconstructed frame than the standard MPEG-4 and the MPEG-4 DRC algorithm.

Keywords- MPEG-4; Dynamic Resolution Conversion (DRC); Rate-Distortion (R-D); Video Coding

I. INTRODUCTION

Video coding reduces the requirement for video data communication. The MPEG-4 is developed in response to the growing need on a video coding standard [1], [2]. However, it still causes visually blocking effects when a high quantization is used to obtain a low bit-rate video [4], [5]. In order to keep a good frame quality in real-time network traffic, the MPEG-4 in the Advanced Real-Time Simple (ARTS) profile proposes a dynamic resolution conversion (DRC) method [3], [6], which encodes video frame in a reduced resolution. In this method, the resolution of each frame is examined and when judged as a low bit-rate condition, the frame resolution is down-sampled to half of the original resolution in both vertical and horizontal directions, and encoded with a quarter number of original resolution with less bits compared to the normal encoding with very preferable subjective quality.

In this paper, a novel 3-D DRC method with adaptive quantization in a rate-distortion sense is combined with the MPEG-4 standard to improve the 2-D DRC of MPEG-4 for low bit-rate video coding. The proposed method uses a 3-D down-sampling as a pre-processing stage of the MPEG-4 encoder and a 3-D up-sampling as a post-processing stage of the MPEG-4 decoder. In this 3-D DRC down-sampling, it can be implemented by the use of a cubic down-sampling procedure to the half size among vertically, horizontally, and temporally. Furthermore, the proposed 3-D DRC up-sampling is calculated by a 3-D linear interpolation. After the 3-D DRC down-sampling, an adaptive quantization based on the rate-distortion (R-D) optimization is used to keep the compressed frame quality. Finally, computer simulation shows that the proposed method obtains a better quality of the reconstructed frame than both the 2-D DRC MPEG-4 and the standard MPEG-4 for low bit-rate video.

The rest of this paper is organized as follows. In Section II, the background of this work that includes dynamic resolution conversion (DRC) and quantization in MPEG-4 algorithm is briefly introduced. In Section III, the 3-D DRC MPEG-4 algorithm is proposed. The 3-D DRC down/up sampling is also defined. Furthermore, an adaptive quantization selection based on rate-distortion (R-D) optimization is described. In Section IV, the experimental results are discussed. Finally, Section V concludes this paper.

II. BACKGROUND OF THIS WORK

A. Dynamic Resolution Conversion Method

![Fig. 1 Block diagram of the 2-D DRC MPEG-4 encoder](image-url)
transmitted in the same way as MPEG-4. At the receiver side, the transmitted encoded frame is up-sampled to the original resolution. As a result, the 2-D DRC MPEG-4 method can be used to improve the coding frame rate without degradation of subjective quality in a low bit-rate condition. For more details on this method, see [3].

B. Quantization in MPEG-4 Algorithm

In the MPEG-4 algorithm, the quantization is crucial for determining the encoded bit rate and quality of reconstructed frame. At the encoder, suppose that the input frame is partitioned into non-overlapping 8×8 pixel blocks that are processed in a raster scan order, each block with 64 pixels. Next, each block is converted by the 2-D DCT transform into a block of 64 DCT coefficients \( C_n \) for \( 0 \leq n \leq 63 \). Then the quantization \( C'_n \) for each DCT coefficients is defined as

\[
C'_n = \text{Round}(C_n / Q_T)
\]  

where \( \text{Round}(x) \) is the closest integer to \( x \), and \( Q_T \) is a predetermined quantization table. At the decoder, the de-quantization is similarly carried out to obtain the de-quantized coefficients \( C''_n \) for \( 0 \leq n \leq 63 \) is defined by

\[
C''_n = C'_n \times Q_T
\]  

The quantization table \( Q_T \) is specified by two default quantization matrices, one for the luminance component (Y) and the other for the chrominance components (U, V), which are given in [2]. The de-quantized coefficients \( C''_n \) for \( 0 \leq n \leq 63 \) will go through the inverse DCT transform to obtain the reconstructed frame.

In a practical application of MPEG-4 video coding, a quantization parameter is involved which scales the quantization table. The bit rate and quality of a coded frame can be controlled by the quantization parameter. However, when a high quantization parameter is used to obtain a high compression ratio or low bit rate, the MPEG-4 will cause seriously blocking effects, because most of the coefficients of a block will often be quantized to be zero. To demonstrate this problem, Fig. 2 shows the frame 30 of the Foreman sequence (352×288 resolution, 24 frames/s) and its MPEG-4 coded at a high compression ratio of 300:1, it is easy to find that the reconstructed frame appears many serious blocking artifacts.

![Fig. 2 Frame 30 (P frame) of the Foreman sequence (352×288 resolution, 24 frames/s). (a) Original Y frame. (b) Reconstructed Y frame with serious blocking artifacts that used MPEG-4 method at a compression ratio of 300:1.](image)

III. PROPOSED 3-D DRC MPEG-4 ALGORITHM

In order to reduce the blocking effects of MPEG-4 decompressed frame illustrated in Section II-B at a high compression ratio or a low bit rate, a new 3-D DRC MPEG-4 algorithm, using an extension of the ideas of the 2-D DRC MPEG-4 method described in Section II-A, is proposed for video coding. Furthermore, this 3-D DRC MPEG-4 algorithm, shown in Fig.3, joints 3-D down/up sampling and adaptive quantization with an optimal rate-distortion (R-D) selection described in Section III-B to improve the low bit-rate video. That is, in a high bit-rate condition, the normal resolution mode is selected in the same way as the standard MPEG-4. However, in a low bit rate, the reduced resolution mode is selected, and an input frame is encoded by the 3-D DRC method mentioned in Section III-A.

A. Framework of Proposed 3-D DRC MPEG-4

In this section, an adaptive 3-D DRC method based on dynamic spatial and temporal down-sampling is along with the MPEG-4 algorithm. That is, the proposed algorithm shown in Fig. 3 may use a 3-D down-sampling with a compression ratio of 8 to 1 as the pre-processing stage of the MPEG-4 encoder and a 3-D up-sampling with a ratio of 1 to 8 as the post-processing stage of the MPEG-4 decoder. Using the same procedure, described in the 2-D DRC MPEG-4 [3], the details of the encoding procedure in the reduced resolution mode are illustrated as follows.

- The size of a block is 16×16, and the size of a macroblock is 32×32. That is, both the size of a macroblock and a block are quadrupled compared to the original spatial resolution. For the 3-D spatial and temporal down-sampling, the number of macroblocks between two neighbouring frames is one eighth of the normal resolution mode.
- In two neighbouring frames, motion estimation is carried out for each 32×32 macroblock and the range of motion vectors is enlarged to double size.
• Prediction of each 16×16 block is created based on the estimated motion vector. Then 16×16 prediction error block is calculated.

• The 16×16 prediction error block is down-sampled to the half resolution. The size of down-sampled prediction error block is 8×8, which is equal to that of the standard MPEG-4.

• DCT (discrete cosine transform) and Q (quantization) are applied to down-sampled 8×8 prediction error block. And VLC (variable length coding) is assigned to each coefficient just in the same way as MPEG-4.

• Coefficients are inverse-quantized (IQ) and IDCT (inverse DCT) is applied. Then, ),(  ),( is minimized for a 2

• 64 binary searches are used, where and

• Coefficients are inverse-quantized (IQ) and IDCT (inverse DCT) is applied. Then 8×8 reconstructed prediction error block is 8×8, which is equal to that of the standard MPEG-4.

• MPEG-4.

• Using the 3-D down-sampled video as an input video and produces an output video by a factor of 2 in each dimension as follows:

\[ t_1, t_2, t_3 \text{ are also integers.} \]

The 3-D down-sampling takes \( X(t_1, t_2, t_3) \) as an input video and produces an output video of \( \hat{Y}(t_1, t_2, t_3) \) by a factor of 2 in each dimension as follows:

\[
\hat{Y}(t_1, t_2, t_3) = \text{avg}(\sum_{i=0}^{1} \sum_{i=0}^{1} \sum_{i=0}^{1} X(2t_1 + i_1, 2t_2 + i_2, 2t_3 + i_3)),
\]

for \( 0 \leq t_i \leq n_i - 1, \ i = 1,2,3 \). (3)

Here \( \text{avg}(•) \) is the average (arithmetic mean) of a set of numeric values.

3-D Up-Sampling Filter

Using the 3-D down-sampled video \( \hat{Y}(t_1, t_2, t_3) \) obtained by (3), the 3-D up-sampled video can be calculated by a 3-D linear interpolation and given by

\[
\hat{X}(t_1, t_2, t_3) = \sum_{k_1=0}^{1} \sum_{k_2=0}^{1} \sum_{k_3=0}^{1} Y(k_1, k_2, k_3) R(t_1 - 2k_1, t_2 - 2k_2, t_3 - 2k_3),
\]

for \( 0 \leq t_i \leq 3, \ i = 1,2,3 \). (4)

where \( R(t_1 - 2k_1, t_2 - 2k_2, t_3 - 2k_3) \) is the 3-D linear function defined by

\[
R(t_1, t_2, t_3) = R(t_1) \cdot R(t_2) \cdot R(t_3)
\]

and \( R(t) \) is the 1-D linear function given by

\[
R(t) = \begin{cases} 1 - |t|/2, & \text{if } |t|/2 \leq 0, \\ |t|/2, & \text{otherwise} \end{cases}
\]

B. Adaptive Quantization Selection

The R-D theory is a lossy compression goal into that of minimizing coding distortion \( D \), which is a measure of distance between the original frame and the reconstructed frame according to a chosen metric, subject to a constraint in the bit-rate \( R \) for coding the frame data [7]. In the 3-D DRC MPEG-4 algorithm, the operational R-D problem that minimizes the distortion of quantization is done while using the 3-D DRC method.

Let \( Q \) be an adaptive quantization step and \( S \) be the 3-D DRC down/up sampling step. For a given bit-rate constraint \( R_c \), the R-D optimization [8] problem for the 3-D DRC MPEG-4 algorithm can be defined by

\[
D(Q,S) \text{ subject to } R(Q,S) \leq R_c
\]

(7)

where \( D(Q,S) \) and \( R(Q,S) \) are the distortion and rate of the 3-D DRC MPEG-4 algorithm, respectively, and given by

\[
D(Q,S) = \sum_{n=0}^{63} D_n(Q(n)) + D(S)
\]

and

\[
R(Q,S) = \sum_{n=0}^{63} R_n(Q(n)) + R(S).
\]

Using the Lagrange multiplier method, the rate-constrained problem in (7) can be converted into the unconstrained problem, namely,

\[
\min_{Q,S}(J(\lambda)) = D(Q,S) + \lambda R(Q,S)
\]

(10)

where \( \lambda \geq 0 \) and \( J(\lambda) \) are Lagrange’s multiplier and cost, respectively. In other words, the Lagrangian method is choosing each \( q = Q(n) \) and \( S \) so as to minimize \( D(Q,S) + \lambda R(Q,S) \). This algorithm is called the optimal R-D selection and described as follows.

Using the same procedure, described in [8], firstly, we suppose all the possible pairs \( (R_n(Q(S),D_n(Q(S))) \) for \( 0 \leq n \leq 63 \) as points on an x-y plane. Then, we get 64 half-convex hulls of R-D points by the use of the Graham scan algorithm [9]. In addition, we define the values \( (q, S) \) to be \( (q_{n}(1), S) \) through \( (q_{n}(h_n), S) \) for \( 0 \leq n \leq 63 \), where \( h_n \) are the remaining points on the hull. Then we determine the slope of the nth coefficient at these \( h_n \) points, that is,

\[
\lambda_n(k) = \frac{D_n(q_{n}(k+1), S) - D_n(q_{n}(k), S)}{R_n(q_{n}(k+1), S) - R_n(q_{n}(k), S)}
\]

for \( 1 \leq k \leq h_n - 1 \),

(11)

where \( \lambda_n(h_n) = 0 \) and \( \lambda_n(k) > \lambda_n(k+1) \geq 0 \). Finally, we find the least index \( k \) for which \( \lambda \geq \lambda_n(k) \) for any given \( \lambda \geq 0 \). Then, \( D_n(Q(S), \lambda R_n(Q(S)) \) is minimized for a solution of \( (q, S) = (q_{n}(k), S) \), where \( k \) is chosen to be the least index such that \( \lambda \geq \lambda_n(k) \).

Therefore, for given \( \lambda \geq 0 \), 64 binary searches are used to efficiently find \( Q \) and \( S \) such that
\[ D(Q,S) + \lambda R(Q,S) \]
\[ = \sum_{n=0}^{63} [D_n(Q(n),S) + \lambda R_n(Q(n),S)] \quad (12) \]

is minimized. In other words, the adaptive quantization \( Q \) and the down/up sampling \( S \) are thus obtained. However, for the normal resolution mode, \( D(S) = 0 \), the proposed 3-D DRC MPEG-4 is just the standard MPEG-4 \([2]\), and for the reduced resolution mode or low bit-rate video, the proposed video coding uses the 3-D DRC down/up sampling method with the adaptive quantization.

IV. EXPERIMENTAL RESULTS

A. Simulation Procedure

Consider an \( M \times N \) video sequence with \( K \) frames. The size of each original video frame in the RGB color space is \( M \times N \times 3 \) pixels. For the best compression result, the original RGB frame is converted into another preliminary frame in YUV color space, in which most of the information is concentrated in the luminance component (Y) and less in the chrominance components (U, V). After color space conversion, one set of \( M \times N \) pixels is used for Y, and two sets of \( \lceil M/2 \rceil \times \lceil N/2 \rceil \) pixels, where \( \lceil \cdot \rceil \) denotes the least integer greater than or equal to \( \cdot \), are used for the U and V color components. The proposed 3-D DRC MPEG-4 algorithm shown in Fig. 3 can then independently process each color component Y, U, and V.

The proposed MPEG-4 algorithm utilizes the 3-D DRC down sampling with adaptive quantization as the pre-processing stage of the MPEG-4 encoder and also uses the 3-D DRC upsampling as the post-processing stage of the MPEG-4 decoder for a low bit rate. There are three processing steps in the encoder phase. The first step is the pre-processing that uses the 3-D DRC down sampling with a compression ratio of 8 to 1 for each of Y, U, and V frames. The second step is to get a new quantization parameter using an adaptive quantization selection illustrated in Section III-B. The third step is to use the MPEG-4 encoding algorithm. The frame after this step is called the MPEG-4 bitstream. This bitstream now has a very small number of bits when compared with the original frame. The resulting file still has the standard MPEG-4 format. As a consequence, this bitstream can be used as the standard MPEG-4 decoder, save on storage, and decrease the transfer time for a communication.

In addition, in the proposed 3-D DRC MPEG-4 decoder, there are two processes. The first step is the MPEG-4 standard decoding algorithm. After this step, the MPEG-4 bitstream is separated into three separate Y, U, and V frames. The second step is the post-processing step that uses the 3-D DRC up sampling with a ratio of 1 to 8 for the Y, U, and V frames. After this process, three Y, U, and V frames are therefore converted to the original resolution. Then, the three Y, U, and V frames are combined again into one YUV format. Finally, this YUV frame is converted into the reconstructed RGB frame.

B. Simulation Result

The proposed MPEG-4 with 3-D DRC method is implemented in Microsoft visual C++ program and compared with the standard MPEG-4 and the 2-D DRC MPEG-4 methods. These three algorithms are based on simple profile (SP). I-frame (without temporal prediction) and P-frame (with prediction from the previous frame) are used according to a periodic structure of the form IPP…, first with I-frame period \( N = 30 \) frames. Performance comparisons are carried out on three movie video sequences of Rush Hour, Girl 1, and Girl 2, shown in Fig.4, and three standard CIF video sequences of Dancer, Table, and Mother and Daughter, shown in Fig 5, respectively. All the reported results (bit rates and PSNR performance) are computed from the reconstructed frame.

The results of these three algorithms are summarized in Table I and Table II. As shown in these two tables, at various bit rates, for Y component of frame sequences with 352×288 and 640×480 resolutions, the proposed 3-D DRC MPEG-4 algorithm is a little better than both 2-D DRC and standard MPEG-4 methods. However, for U and V components, the proposed method is then much better than the standard MPEG-4 and 2-D DRC MPEG-4 method.

Finally, the coding results for the 10th frame of the other standard CIF video sequences of Stefan and the 2nd frame of the 640×480 video sequence of Girl 1 are shown in Figs. 12 and 13, respectively. The Stefan sequence is coded at 30 fps, 900 kbps, and the Girl 1 sequence is coded at 30 fps, 1000 kbps. Both figures include the original and decoded frames from the above three algorithms. The superior visual quality of the proposed 3-D DRC MPEG-4 method is clearly visible in both sequences. That is, the proposed MPEG-4 with 3-D DRC and adaptive quantization obtains a better subjective quality and objective PSNR of the
reconstructed frame than the standard MPEG-4 and the MPEG-4 DRC algorithm.

### TABLE I PERFORMANCE COMPARISON OF THREE ALGORITHMS FOR 352 × 288 AT VARIOUS LOW BIT RATES

<table>
<thead>
<tr>
<th>Video (352 × 288)</th>
<th>Bitrate (Kbps)</th>
<th>PSNR (dB) Frame</th>
<th>MPEG-4</th>
<th>2-D DRC</th>
<th>3-D DRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dancer</td>
<td>220</td>
<td>Y</td>
<td>33.06</td>
<td>32.45</td>
<td>33.06</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td>38.61</td>
<td>40.22</td>
<td>40.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>V</td>
<td>36.40</td>
<td>37.88</td>
<td>37.88</td>
</tr>
<tr>
<td>Table</td>
<td>200</td>
<td>Y</td>
<td>29.33</td>
<td>24.80</td>
<td>29.33</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td>37.13</td>
<td>36.96</td>
<td>37.65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>V</td>
<td>35.46</td>
<td>35.50</td>
<td>35.96</td>
</tr>
<tr>
<td>Mother and Daughter</td>
<td>100</td>
<td>Y</td>
<td>34.43</td>
<td>32.70</td>
<td>34.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td>40.98</td>
<td>41.90</td>
<td>43.01</td>
</tr>
<tr>
<td></td>
<td></td>
<td>V</td>
<td>42.27</td>
<td>42.68</td>
<td>44.06</td>
</tr>
</tbody>
</table>

### TABLE II PERFORMANCE COMPARISON OF THREE ALGORITHMS FOR 640 × 480 AT VARIOUS LOW BIT RATES

<table>
<thead>
<tr>
<th>Video (640 × 480)</th>
<th>Bitrate (Kbps)</th>
<th>PSNR (dB) Frame</th>
<th>MPEG-4</th>
<th>2-D DRC</th>
<th>3-D DRC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rush Hour</td>
<td>300</td>
<td>Y</td>
<td>35.49</td>
<td>36.45</td>
<td>36.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td>41.61</td>
<td>43.62</td>
<td>44.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>V</td>
<td>40.45</td>
<td>42.90</td>
<td>43.50</td>
</tr>
<tr>
<td>Girl 1</td>
<td>220</td>
<td>Y</td>
<td>36.35</td>
<td>36.60</td>
<td>36.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td>38.91</td>
<td>41.68</td>
<td>42.73</td>
</tr>
<tr>
<td></td>
<td></td>
<td>V</td>
<td>41.55</td>
<td>43.79</td>
<td>45.00</td>
</tr>
<tr>
<td>Girl 2</td>
<td>200</td>
<td>Y</td>
<td>38.59</td>
<td>40.19</td>
<td>41.25</td>
</tr>
<tr>
<td></td>
<td></td>
<td>U</td>
<td>42.85</td>
<td>44.27</td>
<td>46.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>V</td>
<td>39.81</td>
<td>41.74</td>
<td>43.57</td>
</tr>
</tbody>
</table>

Fig. 6 Comparison of MPEG-4 (Standard), 2-D DRC MPEG-4 and 3-D DRC MPEG-4 for Mother and Daughter (Y component)

Fig. 7 Comparison of MPEG-4 (Standard), 2-D DRC MPEG-4 and 3-D DRC MPEG-4 for Mother and Daughter (U component)

Fig. 8 Comparison of MPEG-4 (Standard), 2-D DRC MPEG-4 and 3-D DRC MPEG-4 for Mother and Daughter (V component)

Fig. 9 Comparison of MPEG-4 (Standard), 2-D DRC MPEG-4 and 3-D DRC MPEG-4 for Rush Hour (Y component)
C. Computation Complexity

In order to compare the computation complexity of the proposed 3-D DRC MPEG-4 with both standard MPEG-4 and 2-D DRC MPEG-4 methods, the number of multiplication and addition/subtraction/shift are estimated in this section. In addition, it follows from [3] that the number of operations in each function unit is shown in Table III for estimation basis. In this table, the operations are also categorized into two groups: one is multiplication operation and the other is addition, subtraction, and shift operation. Furthermore, it is also assumed that the operations are performed to all the blocks in video frames. To illustrate this, Table IV shows the estimated number of multiplication and addition/subtraction/shift operations to encode/decode two neighbouring CIF frames. It also follows from [3], “Operations in Encoder” is the summation of all the six function units, and “Operations in Decoder” is the summation of “inverse quantization”, “inverse DCT” and “3-D up-sampling”. Obviously, the estimated operation numbers of the proposed 3-D DRC MPEG-4 algorithm in both encoder and decoder are less than those of both standard MPEG-4 and 2-D DRC MPEG-4 algorithms.
TABLE III NUMBER OF OPERATIONS FOR EACH FUNCTION UNIT

<table>
<thead>
<tr>
<th>Function Unit</th>
<th>Multiplication</th>
<th>Addition, Subtraction, Shift</th>
<th>Estimation Basis</th>
</tr>
</thead>
</table>
| 3-D down-sampling   | (16×16)×2 blocks →8×8 block | 0                            | 576 3-D down-sampling of spatial (vertical, horizontal) and temporal directions are performed independently. vertical direction: (16×8)×2=256 addition  
horizontal direction: (8×8)×2=128 addition  
temporal direction: 8×8=64 addition  
rounding: 8×8=64 addition divided by 8: 8×8=64 3-bit shift |
| DCT                 | 8×8 block      | 256                          | 352 For 8 points, 16 multiplication and 22 addition/subtraction are necessary.     |
| Quantization        | 8×8 block      | 64                           | 64 1 multiplication and 1 addition per coefficient                               |
| Inverse Quantization| 8×8 block      | 64                           | 64 1 multiplication and 1 addition per coefficient                               |
| Inverse DCT         | 8×8 block      | 256                          | 352 For 8 points, 16 multiplication and 22 addition/subtraction are necessary.     |
| 3-D up-sampling     | 8×8 block →(16×16)×2 blocks | 0                            | 1920 3-D up-sampling of spatial (vertical, horizontal) and temporal directions are performed independently. vertical direction: 16×8=128 addition  
horizontal direction: 16×16=256 addition  
temporal direction: (16×16)×2=512 addition  
rounding: (16×16)×2=512 addition divided by 8: (16×16)×2=512 3-bit shift |

TABLE IV NUMBER OF OPERATIONS FOR TWO NEIGHBOURING CIF FRAMES

<table>
<thead>
<tr>
<th>Multiplication</th>
<th>3-D down-sampling</th>
<th>DCT</th>
<th>Quantization</th>
<th>Inverse Quantization</th>
<th>Inverse DCT</th>
<th>3-D up-sampling</th>
<th>Operations in Encoder</th>
<th>Operations in Decoder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard</td>
<td>0</td>
<td>1216512</td>
<td>304128</td>
<td>304128</td>
<td>1216512</td>
<td>0</td>
<td>3041280</td>
<td>1520640</td>
</tr>
<tr>
<td>2-D DRC</td>
<td>0</td>
<td>304128</td>
<td>76032</td>
<td>76032</td>
<td>304128</td>
<td>76032</td>
<td>456192</td>
<td>1216512</td>
</tr>
<tr>
<td>3-D DRC</td>
<td>0</td>
<td>152064</td>
<td>38016</td>
<td>38016</td>
<td>152064</td>
<td>0</td>
<td>380160</td>
<td>190080</td>
</tr>
<tr>
<td>Addition, subtraction, and shift</td>
<td>Standard</td>
<td>0</td>
<td>1672704</td>
<td>304128</td>
<td>1672704</td>
<td>0</td>
<td>3953664</td>
<td>1976832</td>
</tr>
<tr>
<td>2-D DRC</td>
<td>380160</td>
<td>418176</td>
<td>76032</td>
<td>76032</td>
<td>418176</td>
<td>1064488</td>
<td>2413024</td>
<td>1558656</td>
</tr>
<tr>
<td>3-D DRC</td>
<td>342144</td>
<td>209088</td>
<td>38016</td>
<td>38016</td>
<td>209088</td>
<td>1140480</td>
<td>1976832</td>
<td>1387584</td>
</tr>
</tbody>
</table>

V. CONCLUSIONS

In this paper, a novel 3-D DRC method is used with the MPEG-4 standard to improve both the 2-D DRC MPEG-4 and the standard MPEG-4 methods for low bit-rate video coding. The proposed method uses a 3-D down sampling as a pre-processing stage of the MPEG-4 encoder and a 3-D up sampling as a post-processing stage of the MPEG-4 decoder. In addition, the proposed 3-D DRC down sampling with a compression ratio of 8 to 1 is implemented by the use of a cubic down-sampling procedure to the half size of spatial (vertical, horizontal) and temporal directions, and the proposed 3-D DRC up sampling with a ratio of 1 to 8 is calculated by a 3-D linear interpolation. Furthermore, an optimal algorithm is utilized to solve the rate-distortion (R-D) problem that joints the 3-D DRC down/up sampling and an adaptive quantization for MPEG-4 video coding. Finally, experimental result shows that the proposed 3-D DRC MPEG-4 method achieves a better subjective quality and objective PSNR of the reconstructed frame than the standard MPEG-4 and the 2-D DRC MPEG-4 methods for low bit-rate video coding.

ACKNOWLEDGMENT

This work was supported in part by the National Science Council, R.O.C., under Grant NSC 94-2213-E-251-004 and NSC 95-2221-E-251-003. The author would also like to thank Mr. Cheng-Ying Wu for his valuable contributions.

REFERENCES


---

**Lung-Jen Wang** received the Ph.D. degree in Computer Science and Engineering from the National Sun Yat-Sen University, Taiwan, 2001. From 2005 to 2008, he was a Chairman in the Department of Information Technology at National Pingtung Institute of Commerce, Taiwan. From 2008 to 2010, he was the Secretary General of National Pingtung Institute of Commerce, Taiwan. From 2010 to 2011, he was the Acting President of National Pingtung Institute of Commerce, Taiwan.

Currently, he is an Associate Professor in the Department of Computer Science and Information Engineering at National Pingtung Institute of Commerce, Taiwan. His primary research interests are in the areas of image processing, multimedia network and real-time system.