

FGS Enhancement Layer Truncation with Minimized Intra-Frame Quality Variation

Jian Zhou, Huai-Rong Shao, Chia Shen and Ming-Ting Sun

TR2003-16 July 2003

Abstract

This paper proposes an enhancement layer truncation scheme for the Fine-Granularity-Scalability (FGS) video. Our target is to minimize the quality variation of different parts within each frame when the last transmitted enhancement layer is truncated according to the available network bandwidth. This paper proposes an enhancement layer truncation scheme for the Fine-Granularity-Scalability (FGS) video. Our target is to minimize the quality variation of different parts within each frame when the last transmitted enhancement layer is truncated according to the available network bandwidth.

2003 IEEE International Conference on Multimedia and Expo (ICME 2003)

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

FGS ENHANCEMENT LAYER TRUNCATION WITH MINIMIZED INTRA-FRAME QUALITY VARIATION

*Jian Zhou**, *Huairong Shao†*, *Chia Shen†* and *Ming-Ting Sun**

*Department of Electrical Engineering, University of Washington, Seattle, WA, 98195

†Mitsubishi Electric Research Laboratory, 201 Broadway, Cambridge, MA, 02139

ABSTRACT

This paper proposes an enhancement layer truncation scheme for the Fine-Granularity-Scalability (FGS) video. Our target is to minimize the quality variation of different parts within each frame when the last transmitted enhancement layer is truncated according to the available network bandwidth. We propose to redistribute the bits in the enhancement layer that can only be partially kept with the available bit-budget so that it is able to cover the whole frame area to raise the quality of different parts uniformly. Simulation results confirm the effectiveness of the proposed method in improving the decoded visual quality and reducing the intra-frame quality variation.

Keywords: Fine Granularity Scalability, constant quality, enhancement layer truncation.

1. INTRODUCTION

For today's Internet video streaming applications, one important concern is to intelligently deliver compressed video streams to the end-users with heterogeneous environments. Fine Granularity Scalability (FGS) [1] has been adopted as an amendment to the MPEG-4 standard to address this concern.

The FGS encoder generates two bit streams: one is the base-layer stream to provide the basic visual quality, and the other is the enhancement-layer stream to improve the base-layer quality. The enhancement-layer stream is encoded into several layers with bit-plane coding [2] scheme, and this enables FGS to provide continuous rate-control, in that the enhancement layers can be truncated at any point to achieve the target bit-rate. The corresponded quality of the reconstructed frames is proportional to the amount of enhancement-layer bits received. However, the standard does not specify how to truncate the enhancement layers; it only specifies how to decode the truncated bit stream.

Some research efforts have been introduced on how to truncate the FGS enhancement-layer bit-stream. One simple method, called "even truncation", evenly allocates

the available bit-budget to each enhancement-layer frame [3]. This scheme has two defects. (1) The decoded visual quality varies from frame to frame, since the complexity of every frame may be different. We call it inter-frame quality variation. (2) The quality of different parts in a same frame also varies from place to place, since the truncated enhancement layer only covers partial frame area. We call this intra-frame quality variation.

In order to solve the first problem, Zhao [4] uses the nearest feather line to evaluate the importance of each frame, and allocate bits to the enhancement-layer by this criterion. Zhang [5] and Zhao [6] propose to use the optimal rate allocation to truncate the enhancement-layer bit-stream, where the rate-distortion (R-D) curves for each enhancement-layer frame are interpolated during the encoding time to determine the amount of bits that should be truncated. This algorithm minimizes inter-frame quality variation. However, none of the schemes mentioned above has considered the intra-frame quality variation.

For the second defect, Cheong [7] uses water-ring scan order together with selective enhancement to transmit the bit-planes in the area of interest. However, the decoder needs to be modified to decode the water-ring scanned enhancement layers. Another problem is that, for many video sequences with natural scenes, it is hard to define the area of interest, or there might be more than one area of interest. Lim [8] proposes to re-order the enhancement-layer macroblocks according to the quantization values and the coded DCT coefficients of corresponding base-layer macroblock. However, this method does not solve the non-uniform quality in the frame when the enhancement layer is truncated, and the decoder also needs to be modified to decode the enhancement layer.

In this paper, we propose a standard-compatible scheme to reduce the intra-frame quality variation, where the last enhancement layer that can be transmitted for each frame is encoded using the available bit-budget so that it can cover the whole frame area to raise the quality of different parts uniformly. Simulation result confirms the effectiveness of the proposed method.

The rest of the paper is organized as follows. Section 2 presents our enhancement-layer truncation approach.

Simulation results are shown in Section 3, and the conclusion is drawn in Section 4.

2. PROPOSED SCHEME

2.1 Problem Statement

The current MPEG-4 FGS uses a normal scan order to encode the enhancement-layer macroblocks from the upper-left corner down to the bottom-right corner in a frame. As a result, the last bit-plane (layer) of will usually cover part of the frame after the truncation. This is shown in Figure 1. At the decoder side, the upper part covered by the transmitted bit-plane will be enhanced, and the lower part of the frame will not get the enhanced quality, thus the intra-frame quality variation arises.

2.2 Enhancement Layer Truncation

Obviously, if the last bit-plane to be transmitted can cover the whole frame, the quality of the whole frame can be enhanced uniformly. However, the channel bandwidth is often not wide enough to transmit the whole bit-plane.

We solve the above problem by re-encoding the last bit-plane for each frame. Compared with the original last bit-plane, each transcoded block has fewer bits than the original one, but the total amount of bits of the last bit-plane will be the same as the original one. The effect is that, the transmitted bit stream is now able to cover the whole frame area, thus the quality of every block will be uniformly enhanced. The whole process to transcode the enhancement layer of one frame is explained below:

1. Encode the current bit-plane as described in the standard, record the amount of bits generated by each block as R_i where $i = 0, 1, \dots, N-1$, (N is the number of block in the frame) and the total amount of bits R_{BP} for the whole bit-plane.
2. If bandwidth allowed, transmit the current bit-plane and go to step 1 to encode the next bit-plane.
3. If not, it means the remaining bit-budget R_{Budget} will not be enough for the whole bit-plane, then the following steps are taken to reduce the number of bits generated in the bit-plane:

- Shrink the bit budget for each block as:

$$R_i' = R_i - \frac{R_i}{\sum_{i=1}^N R_i} \times (R_{BP} - R_{Budget}) \quad (1)$$

where R_i' is the new bit-budget to encoder each block. Equation 1 indicates the over-shot bit budget $(R_{BP} - R_{Budget})$ is allocated to each block by its original bits contribution to the whole frame.

- Re-encode the symbols in each block until the new bit-budget R_i' is met. Each enhancement layer block have 64 bits, either “0” or “1”, corresponding to the residual errors of DC coefficient to the highest AC coefficients. The encoding procedure with new bit budget means some of the “1” applied to enhance the high frequency DCT coefficients will be dropped.
- Process the next block until the end of the frame.

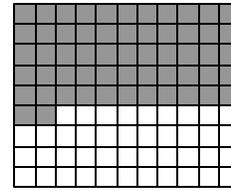
Our proposed scheme can be combined with other algorithms aiming at reducing the inter-frame quality variation, such as [5], so that both the inter-frame and the intra-frame quality variation can be reduced. This scheme is fully standard compatible, and it only introduces a small portion of extra processing on the enhancement layer, and it can be realized in real-time.

2.3 Rate-Distortion Optimization

In the proposed enhancement layer truncation scheme above, what we do is to drop the “1” bits in the enhancement layer block from which corresponds to the highest AC frequency in the DCT domain. Yet this scheme is not optimized from the rate-distortion point of view. For example, assume the first two consecutive coefficients, 8 and 15, will be encoded in the enhancement layer block, and they can be represented as “1000” and “1111” in the binary form. The MSB, or the first enhancement layer contains two “1”. Suppose part of the MSB can be transmitted. If we decide to transmit the bit “1” corresponds to “8”, the overall distortion will be 225 in terms of sum of square difference (SSD), while if to transmit the bit “1” corresponds to “15”, the overall distortion will be as large as 113 in terms of SSD. On the other hand, to erase the bit “1” related to “15” will generate fewer bits to encode the MSB compared with only erasing the “1” related to “8”. So, some balance should be made to decide which “1” bits in the current block should be erased, or, not to be transmitted. The enhancement layer transcoding problem can be generalized as to select some “1” bits from the original block so that the encoded bit-stream will conform to the restricted bit-budget and offer an optimized quality.

Joint rate-distortion optimization method can be utilized to solve this problem. If every block is optimized, then the whole frame is optimized [9]. In one block, for a certain I , we can minimize the cost function of $J(I) = D(R_i) + IR_i$, where R_i is the number of bits used to encode the current block, and $D(R_i)$ is the distortion corresponded to the rate of R_i . As we have mentioned above, when we compute the distortion occurred by erasing the bit “1” in the current bit-plane, the bits associated with the same DCT coefficient in the higher enhancement layer should also be taken into consideration.

- [8] C. Lim and T. Tan, "Macroblock reordering for FGS", ISO/IEC JTC1/SC29/WG11, MPEG 2000/m5759, March, 2000.
- [9] A. Ortega and K. Ramchandran, "Rate-Distortion Methods for Image and Video Compression", IEEE Signal Processing Magazine, Nov. 1998, pp 23-50
- [10] Y. Shoham and A. Gersho, "Efficient Bit Allocation for an Arbitrary Set of Quantizers", IEEE Trans. on ASSP, vol.36, No.9, Sept. 1998, pp1445-1453



The Macroblock that can be transmitted
 The Macroblock that will be truncated

Figure 1– Effect of FGS enhancement layer bit-plane truncation with normal scan order

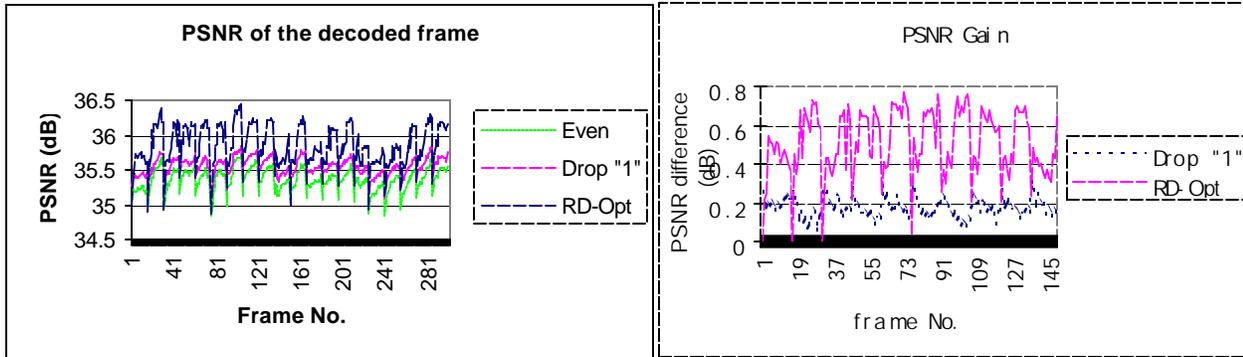


Figure 2– PSNR for each frame (left) and PSNR improvement (right) for each frame in the Akiyo sequence

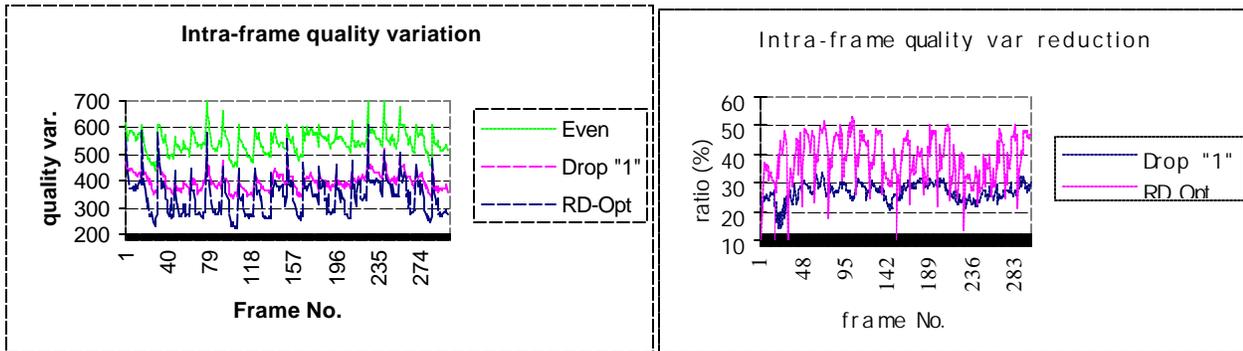


Figure 3– Intra-frame quality variance (left) and variation deduction (right) for each frame in the Akiyo sequence

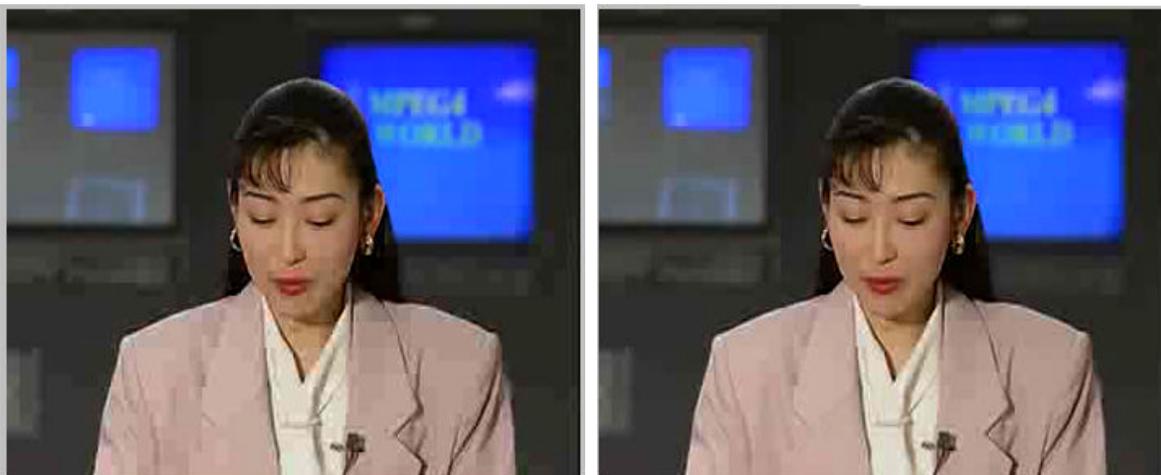


Figure 4– Subjective quality of the decoded frame 61 from even truncation (left) and R-D optimization (right)