Multiple Description Motion Compensation Video Coding for MPEG-4 FGS over Lossy Packet Networks

Chih-Ming Chen, Yung-Chang Chen*, and Chien-Min Chen

Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan 30013, R.O.C., ycchen@ee.nthu.edu.tw

Abstract

A novel error resilience coding technique, named Multiple Description Scalar Quantization for Fine Granularity Scalability (MDSQ-FGS), is presented to improve the temporal prediction efficiency for video coding over lossy packet networks and wireless channels. Despite easily adapting to the channel bandwidth fluctuation, the coding efficiency of scalar coding technique is low since only base-layer is used in its motion prediction. But to achieve higher coding efficiency by using high quality reference frames in enhancement-layer will make “drift” error propagate, caused by the mismatch between the reference frames used in encoding and decoding when reference data are lost. MDSQ-FGS video coding scheme based on MD coding with modified multiple description scalar quantization (MDSQ) and partial predictions are proposed to control drifting error without too much reduction in efficiency. Simulation results indicate that the proposed coder outperforms the normal MPEG-4 FGS coder for coding efficiency. MDSQ-FGS coder will also be applicable for error prone wireless networks for a mobile station moving between two access points.

1. Introduction

MPEG-4 FGS [1] video coding scheme is a scalar coding technique with bit plane coding in the enhancement-layer to adapt to the channel bandwidth fluctuation. But it has a lower coding efficiency with motion prediction only from the lowest quality base-layer. Using high quality reference frames in enhancement-layer can improve coding efficiency, but will make “drift” error propagate. It is difficult to make some trade-off between coding efficiency and drift error. To handle the problem of packet loss, the most popular approach is to add redundancy among packets. MD [2] coding is suitable for multimedia transmission, since its decoder can estimate the original data with tolerable distortion when only part of packets are received. We propose MDSQ-FGS with partial predictions coding scheme to achieve more coding efficiency with low drift error probability. MDSQ-FGS is also a MD coding which constructs a base part of enhancement-layer to be used for motion prediction.

The rest of this paper is organized as the following. The basic idea of MDSQ coding will be reviewed in section 2. In section 3, MDSQ-FGS coder combined with partial prediction is proposed to reduce drift. In section 4, experimental results are presented. In section 5, we modify the index assignment of MDSQ-FGS for wireless application. The conclusions are stated in section 6.

2. Base MDSQ coding schemes

In this paper, we propose a new MD video coder using the base MDSQ and partial prediction. First, we will give a short review of MD coding. Then we introduce a new concept of the base part of MDSQ that is referred to as “Base MDSQ”.

2.1. Review of MD coding

MD coding integrates source and channel coding methods to increase error resilience of data transmission under channel failures, by employing the diversity of channels. Data are encoded into several parts that are called “descriptions”, and sent over multiple independent channels. The probability of failures of all channels is greatly decreased, so decoder will have a large probability to receive correct data from at least one channel at one time, and can estimate the original data with acceptable quality. MD coding intentionally add some redundancy among these descriptions, so it always has a trade-off between coding efficiency and error resilience.

Several MD video coders have been proposed recently. Vaishampayan [4] proposed a video coder based on MDSQ with two independent prediction

2.2. Concept of base MDSQ

In this paper, MDSQ [2] is considered. The MDSQ encoder conceptually extracts the common part of all descriptions, which will be referred to as the “base part” of multiple descriptions. The base part is very useful in video coding, since encoder/decoder may use it as the prediction reference to improve the prediction efficiency, and drift can be avoided if the base part can always be received in the decoder. Fig. 1 shows an example of a modified index assignment of base MDSQ. In the index assignment matrix, we group all the central indexes to form a number of “base part groups”. Each group corresponds to different reconstruction value and is bounded by a rectangle, which is a square generally. We refer the width or height of the square rectangle as “base part size”.

Assume that the scalar quantizer at the encoder partitions the source value into \( N \) central indexes, and the partitioning thresholds are \( \tau = (\tau_0, \tau_1, \ldots, \tau_N) \), where \( \tau_0 \leq \tau_1 \leq \ldots \leq \tau_N \). Let the probability distribution function of the source be \( p(x) \). The \( j \)-th base part group \( B_j \) contains a set of central indexes \( \{i_{j,1}, i_{j,2}, \ldots, i_{j,N_{j,b}}\} \) and corresponds to a set of side-1 indexes \( \{i_{j,1}^1, i_{j,2}^1, \ldots, i_{j,N_{j,s1}}^1\} \) and a set of side-2 indexes \( \{i_{j,1}^2, i_{j,2}^2, \ldots, i_{j,N_{j,s2}}^2\} \), where \( N_{j,b} \) is smaller than or equal to the square of the base part size, \( N_{j,s1} \) and \( N_{j,s2} \) are equal to the base part size. The reconstructed value of the base part \( \tilde{x}_{B_j} \) is calculated as the mean value of cell intervals of all the central indexes \( \{i_{j,1}^b, i_{j,2}^b, \ldots, i_{j,N_{j,b}}^b\} \) as follows:

\[
\tilde{x}_{B_j} = \sum_{i_{j,1}^b \leq x < i_{j,2}^b} \int_{i_{j,1}^b}^{i_{j,2}^b} x \cdot p(x) dx
\]

Where a possible pair of side indexes \( (i_j, i_j') \) is mapped from a central index \( i_j \). We may describe the reconstruction procedure of base part from central and side indexes by three mappings. The first mapping \( b_0: \{1, 2, \ldots, N\} \rightarrow \Re \) maps a central index to the base part, while the second \( b_1: \{1, 2, \ldots, M\} \rightarrow \Re \) and the last \( b_2: \{1, 2, \ldots, M\} \rightarrow \Re \) map the side-1 and side-2 indexes to the base part, respectively. Now we can formally define the three mappings as follows:

\[
b_0(i) = \tilde{x}_{B_j}, \quad \forall i \in \{i_{j,1}, i_{j,2}, \ldots, i_{j,N_{j,b}}\}
\]

\[
b_1(i) = \tilde{x}_{B_j}, \quad \forall i \in \{i_{j,1}^1, i_{j,2}^1, \ldots, i_{j,N_{j,s1}}^1\}
\]

\[
b_2(i) = \tilde{x}_{B_j}, \quad \forall i \in \{i_{j,1}^2, i_{j,2}^2, \ldots, i_{j,N_{j,s2}}^2\}
\]

As long as the decoder receives at least one description, we can obtain the base part by the mapping \( b_0(\cdot) \) or \( b_1(\cdot) \) or \( b_2(\cdot) \) depending on which descriptions are received. The base part of the source is available with high probability at the decoder, so we may use it as the prediction reference for video coding. By this method, there is almost no drift in the video decoding over lossy networks.

3. MDSQ-FGS with partial prediction

Inspired by the concept of RFGS [7], we propose the idea of partial prediction in MDSQ-FGS coding. Partial prediction is a well-known technique to combat error propagation in the predictive coding, at the expense of increased bit rate. The prediction reference \( \tilde{x}[n] \) of the source sequence \( x[n] \) is given by \( \tilde{x}[n] = \tilde{x}[n-1] + \alpha \cdot \Delta x[n-1] \). Error is fixed no matter how large the temporal distance \( (n-j) \) between the current sample and the lost sample is. Hence the error effect of the lost residue remains constant with the time elapsed.

The encoder architecture of MDSQ-FGS with partial prediction is shown in Fig.2 and the decoder in Fig.3. The enhancement-layer bit streams are produced by first calculating the difference between the original DCT coefficients and the inverse quantization, then encoded by MDSQ coder to split into two descriptions (i.e. side indexes). We further apply run-length coding and variable-length coding on the side indexes prior to be transmitted. The resulting bit streams are packetized to form MEPG-4 video packets. The base part also is extracted from base MDSQ coder and added into the feedback loop of the base-layer to get higher quality reference for motion prediction. The partial factor \( \alpha \) is used to attenuate the drift error when both descriptions are not received correctly, but it still can be used to
reconstruct partial enhancement reference. However, the smaller $\alpha$ will lead to lower performance when base part of MDSQ coder can be reconstructed correctly. In this paper, the proposed algorithm just considers the case of only two descriptions in order to reduce the complexity of implementation and theoretical calculation. The partial factor $\alpha$ is set as unity for base MDSQ.

4. Simulation results

In our simulation, we implement the algorithms over the MPEG-4 video coder with FGS-scalability. QCIF 4:2:0 test sequences are coded at 30 frames per second. Only the first frame is IVOP, and the others are PVOP. Base-layer bit stream is coded at a fixed bite rate of 38.4Kbps by “TM5” rate control to carry the minimum fundamental data, including headers and motion vectors. And we assume this base-layer bit stream can be received correctly at decoder without any loss, and this is also the assumption of MPEG4 video coding scheme. For the enhancement-layer bit streams, the video package length is 512 bytes and six packet random loss rates, including 0%, 3%, 15%, 30%, 50% and 100%, are taken for both independent channels. The width of diagonal lines in the index assignment matrix of base MDSQ is 3.

In our experiments, 100 frames are encoded. It needs 15K bytes to encode the base-layer bit stream, and 2,175K bytes for the enhancement-layer bit stream for normal MPEG-4 FGS coder to encode video without large quantization error. For MDSQ-FGS coder, we use the same data rate as used in normal MPEG4 FGS coder to encode the base-layer bit stream, and the enhancement-layer bit streams need a total of 2,454K bytes, that is, 1,227K bytes for each description. Fig.4 shows that MDSQ-FGS coder can greatly reduce the distortion at loss rates lower than 50% since the base part (also the prediction reference) are seldom lost. But when the loss rate is higher than 50%, the probability that both descriptions are lost at same time will increase, that make more drift error. Fig.4 also shows that MDSQ-FGS coder has lower variance of end-to-end distortion.

![Fig. 2. Encoder architecture of MDSQ-FGS video coder with partial prediction](image)

![Fig. 3. Decoder architecture of MDSQ-FGS video coder with partial prediction](image)

![Fig. 4. PSNR comparison for MDSQ-FGS and normal MPEG-4 FGS coder](image)

![Fig. 5. PSNR comparison for MDSQ-FGS and normal MPEG-4 FGS coder at different loss rates](image)
In order to fairly compare the video coders, MDSQ-FGS coder doesn’t encode some least significant bit planes of enhancement-layer for both descriptions to match the data rate as used in normal MPEG4 FGS to meet a fixed available bit rate. In this simulation, the enhancement-layer bit streams of MDSQ-FGS needs a total of 2,028K bytes, that are less but almost equal to 2,175K bytes used for normal MPEG-4 FGS. Fig.4 shows that MDSQ-FGS gets lower PSNR at lower loss rates below 3%, but has higher PSNR benefit about 2.41dB at loss rates among 3%~50%. When the loss rate is higher than 50%, the PSNR is lower than normal MPEG-4 FGS coder again.

5. Multi-path wireless application

Video streaming over wireless networks is challenging because the channel links are unreliable. Since the endpoint’s mobility cause the variation of the radio linking quality, that will affect the bandwidth of a single channel. We believe that multiple path transport has more potential in wireless LANs. Multiple path transport schemes have been proposed in the past for wireless networks for increased connection reliability [8, 9]. In most cases of MD coding, the descriptions are designed to be equivalently important. But for wireless application, mobile computing device will move among base stations. As shown in Fig.6, device staying at position “A” that will has better linking quality with AP1 than AP2. We suppose that the channel with better linking quality will have lower error probability. MDSQ-FGS can also be applied in this situation by modifying the index assignment that the “base part groups” are bounded not by a square rectangle as drawn in Fig.7. The shape of this rectangle is variable to match the channel condition. If the side of width is more than height, it means that the bit streams of D2 can be transmitted over the better channel from AP2 to device at position “B” in Fig. 7 to carry more precise index data. Applying this modified MDSQ-FGS concept for multi-path wireless application is still a future work.

6. Conclusions

In this paper, MDSQ-FGS coder is proposed. A novel base/enhancement part of MDSQ, with modification of the index assignment matrix, is applied to MPEG-4 FGS enhancement-layer coding to improve the temporal prediction efficiency for video coding over lossy networks. Simulation results tell us that the proposed MDSQ-FGS coder can reduce drift for PSNR improvement, and have lower variation of end-to-end distortion. And we also described that MDSQ-FGS is suitable for wireless multimedia service over un-balanced multiple channels.

7. References