System Modeling and Implementation of MPEG-4 Encoder under Fine-Granular-Scalability Framework

Final Report

Embedded Software Systems
Prof. B. L. Evans

by

Wei Li and Zhenxun Xiao

May 8, 2002

Abstract

Stream media over the Internet is now the mainstream of media communications due to the booming of Internet usage as well as the increasing demanding. Internet is very heterogeneous, in the sense of the dynamic distribution of bandwidth and different receivers with different processing ability. How to ensure different quality of service under different environment falls under the category of scalability.

Fine Granularity Scalability, known as FGS, is a new amendment for MPEG-4 standard to address the scalability issue, and it is generally applicable to all video-encoding schemas.

In this project, we started from the FGS amendment for MPEG-4 and proposed an abstract model for video encoder using Homogeneous Synchronous Data Flow (HSDF) specification. As an example, we successfully ported the free source code of MPEG-2 from MPEG Simulation Software Group (MSSG) using this model in Ptolemy. Also, the further improvements of this model are extensively discussed.

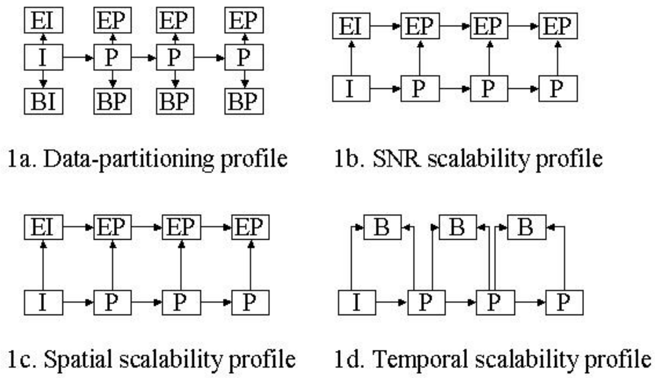
1. Introduction

MPEG-1, MPEG-2 and MPEG-4 are a set of standards developed by moving pictures experts group (MPEG) for compression of motion video at various rates. Standards developed by MPEG only define the syntax of encoded bit streams, and they have been applied to many different applications, such as video CD (MPEG-1) and DVD (MPEG-2).

Due to the booming of the Internet and increasing demands, streaming videos over the Internet has become the mainstream of media communications [1, 6]. Providing streaming media services over the Internet poses higher requirements than traditional video coding and decoding techniques, which are designed to provide optimized services for certain video quality at given bit rates [4]. One special requirement is to solve the scalability issue and provide scalable service. The scalable streaming service should be not only efficient in using network bandwidth, but also capable of adapting to the dynamic distribution of bandwidth due to the variable network configuration and network congestion in IP network and capable of providing services to receivers with different receiving ability as well. MPEG-4 developed by moving pictures experts group (MPEG) proposes such a mechanism – Fine Granularity Scalability with bit plane coding to achieve this goal.

In this report, we will first investigate the approaches that addressing the scalability issue in MPEG-2. Following that, we will introduce the amendment of FGS framework in MPEG-4 and the coding method being adopted in the amendment, known as bit plane coding. Then we will present a homogeneous synchronous data flow (HSDF) model for MPEG-4 with FGS support as well as the software implementation using the SDF domain of Ptolemy. And then, we will introduce some fundamental experimental results. Finally, we will present our conclusions and the future work.

2. Different Scalable Profiles in MPEG-2:



1a. Data-partitioning profile 1b. SNR scalability profile

1c. Spatial scalability profile 1d. Temporal scalability profile

Figure 1. Illustration of different scalable Profiles, data-partitioning, SNR scalability, spatial scalability and temporal scalability, in MPEG-2, with only two layers, known as base-layer and enhancement-layer.
I, P and B in figure refer to I-frame, P-frame and B-frame in MPEG standard. BI, BP, EI and EI denote base-layer I-frame, base-layer P-frame, enhancement-layer I-frame and enhancement-layer P-frame in scalable profiles in MPEG-2 respectively. The arrows connecting different frames show the reference relationship between them.

The scalability issue has been extensively investigated in MPEG-2, and there are four basic profiles proposed to provide the scalable service, known as data partitioning, SNR scalability, spatial scalability and temporal scalability [5], as illustrated in Fig. 1.

The basic idea of above four scalability profiles is to separate the media information into different layers, with a base-layer containing coarse quality video and enhancement layers containing fine information. However, a major disadvantage of these methods is that the decoder needs to get all the information in the enhancement layer to improve the image quality or unable to use it at all.

3. FGS Framework under MPEG-4:

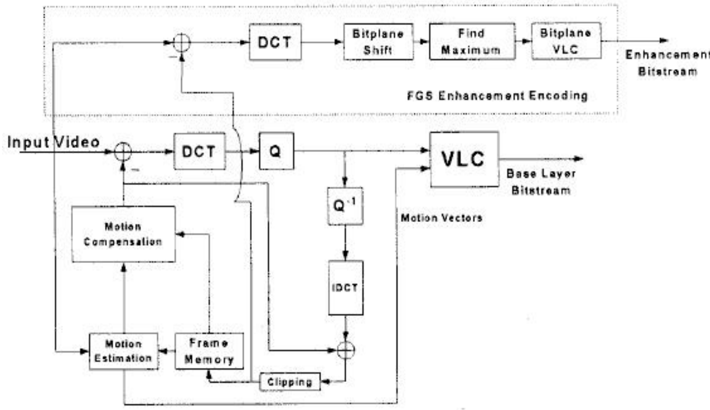


Figure 2. Illustration of video coding performance [2]

FGS also codes media information into a base layer and an enhancement layer. The basic layer is encoded using non-scalable coding to reach the minimal bandwidth requirement.

The enhancement layer is used to code the residue image between the reconstructed image from the base layer and the original image. The enhancement layer is coded by using bit plane coding of the DCT coefficients of the residue image so that the bit streams in enhancement layer can be truncated into any number of bits planes and the decoder is able to reconstruct the enhancement layer from the truncated bit streams and the base layer. The structure of FGS encoder is shown in Fig 2 [2].

For the sake of proving scalable service, fine-granular-scalability (FGS) framework was proposed as an amendment in MPEG-4 [2, 3]. The key issue of FGS framework is the introducing of a new encoding schema, known as bit plane coding (Fig. 3), in substitution of run-length entropy encoding schema which is used in MPEG-2 profiles for the encoding enhancement layer.

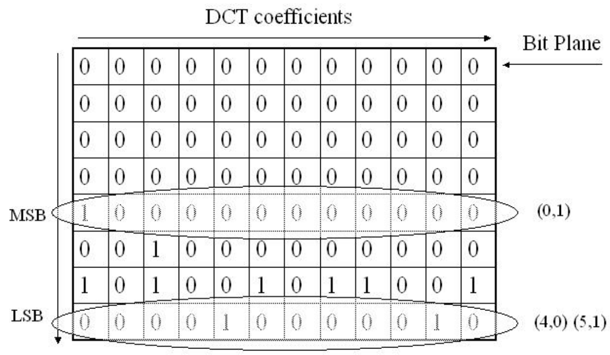


Figure 3. An example of bit plane encoding schema [9]

In the figure, each column is the 8-bit representation of a DCT coefficient. Traditional run-level encoding schema is to encode each DCT coefficient individually. In bit-plane schema, all DCT coefficients are arranged as a matrix, each row with the bit from all DCT coefficients at the same significant position are encoded separately. The bit-plane with all zero bits are not coded.

As illustrated in Fig. 3, for a (8, 8) DCT block, a bit plane is composed of 64 bits, with one bit from each of the 64 DCT coefficients at the same significant position. Different from the conventional run-level coding in which DCT coefficients are coded one after another by using variable length code, bit plane coding applies variable length code one bit plane by one bit plane. As illustrated in Figure 3, the encoding order for bit-plane coding schema is from the most significant bit (MSB) to the least significant bit (LSB). When enhancement layer is encoded in this way, decoder is able to utilize partial information of the enhancement layer.

4. Modeling and Implementation

Our initial goal is to model and implement a MPEG-4 encoder with FGS scalability profile. The modification of MPEG-4's commercial model during the course motivates us to abstract a model that can capture the essential features of video encoder, and have the flexibility to accommodate various standards.

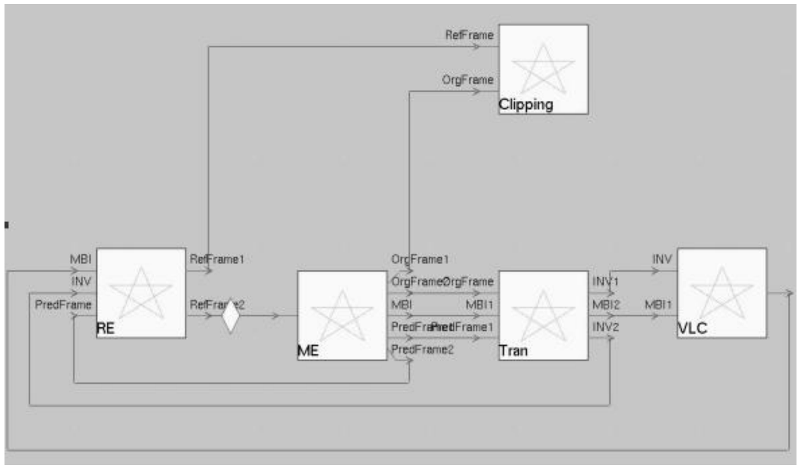


Figure 4. HSDF model and implementation of video encoder in SDF domain of Ptolemy. In the figure, *RE*, *ME*, *Tran*, *VLC* and *Clipping are implemented as five* dynamically linked stars. Among the five stars, *RE* is responsible for image reconstruction, *ME* is responsible for motion estimation, *Tran* is responsible for DCT transformation, *VLC* is responsible for varied length coding and *Clipping* is the functional block responsible for getting the residual image between original and reconstructed images for enhancement layer output in FGS profile.

By investigating the free MPEG-2 codec source code from MPEG Software Simulation Group (MSSG) [10], we generalized three basic data structures, *Frame*, *Blocks* and *MBI* (Macroblock information). Among the three data types, *Frame* contains the information about the video clips in real space, *Blocks* contains the information about video clips in Fourier space after discrete-cosine-transformation (DCT), *MBI* is an accessory data structure which contains the information associate with every marcoblock, such as the frame type for the macroblock, macroblock type, motion type, quantitization factor, motion vectors, etc. The basic data structures are very coarse, and we believe that it is beneficial for separating the modeling and implementation.

In accordance with the FGS encoder structure, as shown in Fig. 2, we constructed a homogeneous SDF model, as illustrated in Fig. 4, in the SDF domain of Ptolemy, and the five dynamic linked stars are implemented by porting the MSSG's implementation.

5. Experiment Results

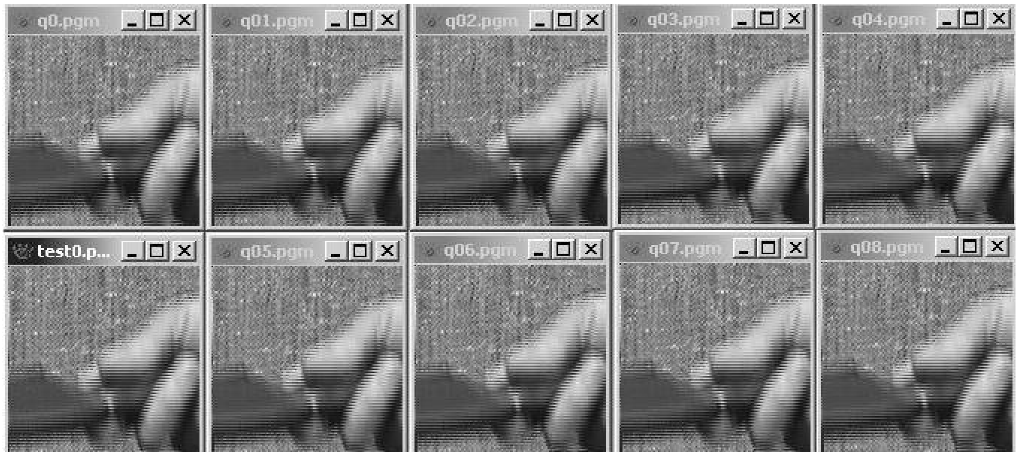


Figure 5. FGS effect on the 128x128, 8-bit grayscale picture
In the figure, q0.pgm is the base plane, test0.pgm is the original picture. Pictures with name of q0x.pgm is the combination of base plane with the x bit planes from the most significant bit plane, in this case is 7.

With our implementation of the video encoder using Ptolemy, we tested the effect of FGS using the encoder, as shown in Fig. 5.

As illustrated in the set of pictures in Fig. 5, reconstructed pictures with additional bit planes from enhanced bit planes are not obviously better than the picture with only base plane. The facts that the picture itself is in low resolution, and there is not sharp variation in the picture, which is reflected in high frequency domain, may lead to this obscure result. We are going to do some more tests with different quality pictures to further address this issue.

6. Conclusions and Future Work:

In this project, we constructed an abstract model for video encoder with the FGS feature in SDF domain. By porting the MSSG implementation of MPEG-2 codec in this model using Ptolemy proved that the model is feasible.

As demonstrated in our results, FGS does not have obvious enhancing effects for all pictures. What is the applicable range of the FGS feature is an interesting question. Also, further investigation about the trade-off between the video qualities versus the size of enhanced information is also a challenge.

Also, from the point view of implementation, the model presented here is scalable. The model depends on message passing between different actors, and it can be easily ported to multiple-processor environments. Look closely, for the ME star, which is responsible for motion estimation and has the highest computation requirement, hierarchy implementation in different domains, such as Process Network can make it further scalable.

With the coarse data structure and flexible implementation, the proposed model is also extensible to other video encoding standards, such as the newly appearing H.26L with FGS feature [11].

References:

[1] D. P. Wu, Y. T. Hou, W. W. Zhu, Y. Q. Zhang, J. M. Peha, "Streaming Video over the Internet: Approaches and Directions", *IEEE Transactions on Circuit and Systems for Video Technology, vol*. 11, no. 1, pp. 1-20, February 2001.

[2]. W. P. Li, "Overview of Fine Granularity Scalability in MPEG-4 Video Standard:, *IEEE Transactions on Circuit and Systems for Video Technology*, vol. 11, no. 3, pp. 310-317, March 2001.

[3] H. Radha, M. van der Schaar and Y. W. Chen, "The MPEG-4 Fine-Grained Scalable Video Coding Method for Multimedia Streaming Over IP", *IEEE Transactions on Multimedia*, vol. 3, no. 1, pp. 53-68, March 2001.

[4] H. Sun, W. Kwok, and J. W. Zdepski, "Architectures for MPEG compressed bitstream scaling", *IEEE Transactions on Circuit and Systems for Video Technology*, vol. 6, no. 2, pp.191-199, April 1996.

[5] B. G. Haskell, A. Puri, and A. N. Netravali, "*Digital Video: An Introduction to MPEG-2*", New York: Champion and Hall, ISBN 0-412-08411-2, 1997.

[6] G. J. Conklin, G. S. Greenbaum, K. O. Lillevold, A. F. Lippman, and Y. A. Renzik, “Video Coding for Streaming Media Delivery on the Internet”, *IEEE Transactions on Circuit and Systems for Video Technology*, vol. 11, no. 3, pp.269-281, March 2001.

[7] F. Wu, S. Li, and Y.-Q. Zhang, “A framework for efficient progressive fine granularity scalable video coding”, *IEEE Transactions on Circuit and Systems for Video Technology*, vol. 11, no.3, pp. 332–344, Mar. 2001.

[8] F. Wu, S. Li, X. Sun, R. Yan and Y,-Q. Zhang, “ Macroblock-based progressive fine granularity scalable coding”, ISO/IEC JTC1/SC29/WG11, MPEG2001/N6779, Pisa meeting, January, 2001.

[9] F. Ling, W. Li, and H. Sun, “Bitplane coding of DCT coefficients for image video compression”, *Proc. SPIE Visual Communications and Image Processing (VCIP),* vol. 3653, no. 1, San Jose, CA, Jan. 25-27, 1999.

[10] MPEG Software Simulation Group, *Free MPEG Software*, http://www.mpeg.org/MPEG/MSSG/.

[11] Y. W. He, F. Wu, S. P. Li, Y. Z. Zhong, S. Q. Yang, "H.26L-based fine granularity scalable video coding", *ISCAS* 2002.