Applications of Multimedia Framework and Digital Video Watermarking on Internet

☑

91 10 22

Applications of Multimedia Framework and Digital Video Watermarking on Internet

MPEG-21

FGS

FGS

FGS

FGS

MPEG-21 MPEG-4 MPEG-7

FGS(Fine Granularity Scalable)

，

Digital Item Adaptation (DIA)

ABSTRACT

MPEG-21 provides a unified solution, Universal Multimedia Access (UMA), for constructing a multimedia content delivery and rights management framework. Based on the concepts of UMA, we build a simplified UMA model on the Internet. In this framework, the source video material is encoded and archived as FGS bitstreams. The video features such as motion vectors are recorded in the format of MPEG-7 descriptors. To support video contents of different formats, we create a transcoder to convert the bitstream from the FGS format to an MPEG-4 simple profile format that fits with the terminal capability. Moreover, a novel FGS coding scheme is created to improve the robustness of the FGS bistreams in an Internet transmission scenario. Consequently, the multimedia information can be streamed through the networks without networks jitters and significant quality degradations existed in the current commercial implementations. To have a more strict evaluation methodology according to the specified common conditions for scalable coding, an FGS-based unicast streaming system as a test bed of scalability over the Internet.

Keywords:
MPEG-21, Universal Multimedia Access, MPEG-4, MPEG-7, Digital Watermarking, FGS, Transcoding, RFGS (Robust FGS), Vide streaming test bed, Digital Item Adaptation (DIA).

I. Introduction

Recently, multimedia technology provides the different players in the multimedia value and delivery chain with an excess of information and services. Access to information and services can be provided with ubiquitous terminals and networks from almost anywhere at anytime. However, no solution enables all the heterogeneous communities to interact and interoperate with one another so far.

MPEG-21 provides a solution named as Universal Multimedia Access (UMA) [1]. The major aim is to define a description of how these various elements such as network, terminal, user preference, and natural environments, etc. can fit together. It defines a multimedia framework to enable transparent and augmented use of multimedia resources across a wide range of
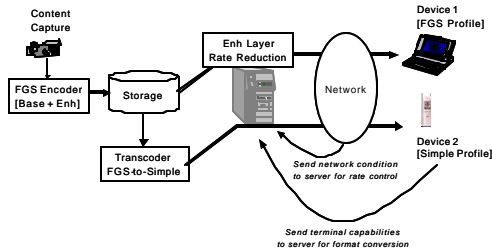
(a) Proposed architecture
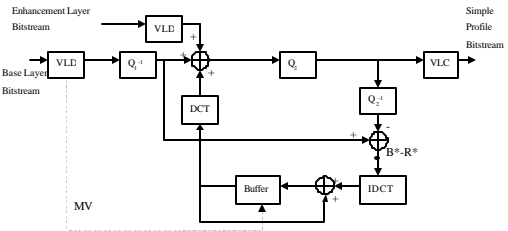
(b) FGS-to-Simple Transcoder

Figure 1. The application scenario of the proposed UMA multimedia delivery system that employs the archived FGS bitstreams. Figure (b) is the block diagram of the FGS-to-Simple Transcoder in Figure (a).

networks and devices employed by different communities. Additionally, it allows content adaptation according to terminal capability.

II. The ARCHITECTURE of UMA MULTIMEDIA DELIVERY SYSTEM

For achieving UMA, we propose a video server that contains the key modules described in MPEG-21 [1]. In this model, we combine the tools as referred to MPEG-4 Fine Granularity Scalability (FGS) [2], MPEG-4 Simple Profile, MPEG-7 [3], Digital Watermarking techniques [4], and Internet protocols.

A. Efficient FGS-to-Simple Transcoding

To fit with the issue of content adaptation according to terminal capability, we propose a real-time transcoding system that converts the FGS bitstreams into Simple Profile bitstreams covers five modules as shown in Figure 1(a).

1) Video Content Capture

The source video is inputted from the image acquisition device and saved in a digital form like YUV or RGB formats.

2) FGS Encoder and Bitstream Archive

Each video signal is encoded as FGS bitstreams and stored in mass storages. When any request from users, the server will send bitstreams directly to the terminals, or pass them through the Transcoder when format conversion is necessary.

3) Real-time Transcoder:

The transcoding depends on the channel conditions, terminal capabilities, and video content features. Assuming the terminals only support MPEG-4 Simple Profile. When justifying the transmission conditions, the server will ask the transcoder to convert the archived FGS bitstreams into Simple Profile bitstreams with specified formats, sizes, and qualities in an adaptive manner.
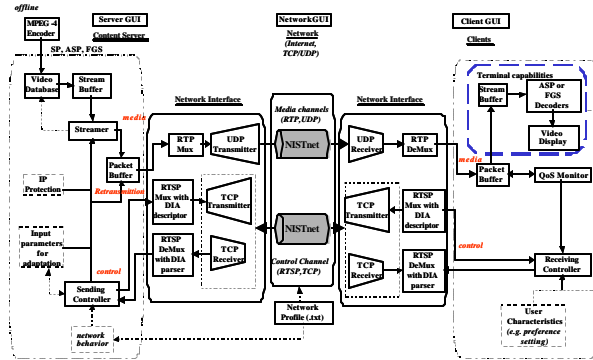
(a) System architecture

Figure 2. Architecture of the proposed FGS-based video streaming test bed.

The real-time FGS-to-Simple transcoder [5] is as illustrated in Figure 1 (b).

The reference method to perform FGS-to-Simple transcoding adopts a cascaded architecture that connects a FGS decoder and a Simple Profile encoder. Our primary objective is to simplify this reference architecture and demonstrate the high quality results of our FGS-to-Simple transcoder.

In the proposed architecture, the motion vectors within the FGS base layer bitstream are reused in MPEG-4 Simple Profile encoder. Additionally, the transcoding is processed in DCT-domain that can provide a low-complexity transcoder.

4) Channel Monitor:

It accepts feedback information from the terminals and also estimates the characteristics of channels, which mean the round-trip time, packet lost ratio, bit error rate, and bandwidth. All obtainable information is passed to the server for adapting the content delivery.

5) Terminals:

Priori to receiving the bitstream, the terminal exchanges its capabilities with the server. As shown in Figure 1(a), the terminals of different decoding capabilities including FGS and Simple

Profile are supported. Consequently, the terminals receive and reconstruct the demanded video signals.

Thus in our framework, the source video is encoded and archived as FGS bitstreams, which can provide various QoS service like SNR scalable video coding schemes [2]. With the FGS bitstreams saved in FGS BitStream Archive module, the proposed system can serves heterogeneous terminals through the Internet. Moreover, according to Internet and Terminal devices capabilities, the Channel Monitor can adapt the different resources to each Terminal.

B. *FGS Streaming on the Internet*

The delivery of multimedia information to mobile device over wireless channels and/or Internet is a challenging problem because multimedia transportation suffers from bandwidth fluctuation, random errors, burst errors and packet losses [2]. However, it is even more challenging to simultaneously stream or multicast video over Internet or wireless channels under UMA framework. The compressed video information is lost due to congestion, channel errors and transport jitters. The temporal predictive nature of most compression technology causes the undesirable effect of error propagation.

To address the broadcast or Internet multicast applications, we propose novel techniques to further improve the temporal prediction at the enhancement layer so that coding efficiency is superior to the existing FGS. Our approach, which is called as RFGS (Robust FGS) [6], utilizes two parameters, the number of bitplanes b ($0 \le b \le$ Maximal number of bitplanes) and the amount of predictive leak a ($0 \le a \le 1$), to control the construction of the reference frame at the enhancement layer. These parameters a and b can be selected for each frame to provide tradeoffs between coding efficiency and error drift. Our approach offers a general and flexible framework that allows further optimization. It also includes several well-known motion compensated FGS techniques as special cases with particular sets of a and b. We analyze the theoretical advantages when parameters a and b are used. We provide an adaptive technique to select a and b, which yields an improved performance as compared to that of fixed parameters. The identical technique is applied to the base layer for further improvement.

C. *FGS-Based Video Streaming Test Bed for MPEG-21 UMA with Digital Item Adaptation*

As shown in **Figure 2**, we are developing an FGS-based unicast video streaming test bed, which is now being considered by the MPEG-4/21 committee as a reference test bed [7], [8]. The proposed system supports MPEG-21 DIA scheme which leads to a more strict evaluation methodology according to the MPEG committee specified common test conditions for scalable video coding. It provides easy control of media delivery with duplicable network conditions. To provide the best quality of service for each client, we will propose relevant rate control, error protection, and transmission approaches in the content server, network interface, and clients, respectively.

III. System Simulations

A. *Efficient FGS-to-Simple Transcoding*

To demonstrate the performance of our proposed UMA multimedia delivery system in Figure 1(a), we build an FGS-to-Simple transcoder in Figure 1(b). In the proposed system, each sequence is pre-encoded and stored in the FGS Bitstream Archive.

Three methods are considered for comparison:
1. A simple profile encoder using the original video sequence (SP_ME)
2. A cascaded transcoding using a complete FGS enhancement bitstream and motion vectors from the base layer bitstream (SP_MV)
3. An efficient transcoding with complete FGS enhancement video and motion vectors from the base layer bitstream (FGS-to-SP).

The test video sequences, named as Foreman, News, and Container, are in CIF and YUV format. The first frame is coded as an I-VOP and the others are coded as P-VOP's at 30Hz. For the FGS encoding, the quantization step size (QP) used in the base layer is set at 10 for I-VOP's and 12 for P-VOP's. The MPEG-4 Simple Profile encoder employs constant quantization, where the set of QP used is {5, 7, 14, 21, 28}. As shown in Figure 3, our transcoding schemes (FGS-to-SP) have neglected quality loss in PSNR at low and medium bit-rates and have about 0.5~0.9 dB loss in PSNR at high bit-rate.

B. *FGS Streaming on the Internet*

The coding efficiency of the RFGS is compared with the base-line FGS coding. The results of our
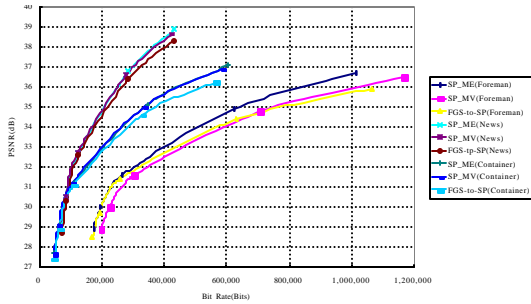
Figure 3. The performance of transcoding with the luminance components of the three video sequences and using various sources of motion vectors and different enhancement information.

proposal are denoted as RFGS (a, b). All the tests adopt the test condition B as in the core experiments testing conditions as specified by the MPEG-4 committee. The sequences used includes Akiyo, Foreman, and Coastguard, in CIF picture size and YUV format. The GOV size is 60 frames that consist of one *I*-picture, 19 *P*-pictures, and two *B*-pictures between each pair of *P*-pictures. To derive the motion vectors for *P*- and *B*-pictures, a simple half-pixel motion estimation scheme using linear interpolation is used. The search range of the motion vectors is set to ±31.5 pixels. The bitrate of the base layer is 256 kbits/sec with TM5 rate control, and the frame rate is 30 Hz. The total bit rate of the enhancement layer bitstream is truncated to bitrates ranging from 0 kbits/sec to 2048 kbits/sec with an interval of multiple of 128 kbits/sec. A simple frame-level bit allocation with a truncation module is used in the streaming server to obtain the optimized quality under the given bandwidth budget.

The simulation results are illustrated in Figure 4. The overall performance of the fast motion sequences such as Coastguard and Foreman is about 2 dB over the baseline FGS The overall performance of the slow motion sequence such as Akiyo is less improved up to 1.1 dB over the baseline FGS. To verify the error recovery capability of the RFGS, we assume the network bandwidth is sharply dropped only when the first *P*-picture is transmitted within a GOV and the bit budget for the other frames is set as 1024 kbits/sec. Under this bandwidth scenario, Figure 5 shows that the value of a strongly affects the error attenuation capability of the RFGS framework. For a small a of 0.5, the error is attenuated very fast. After forth *P*-pictures, the PSNR differences are
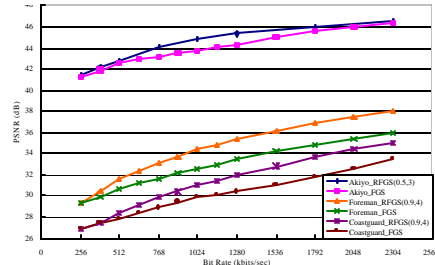

Figure 4. PSNR versus bitrate comparison between FGS and RFGS coding schemes for the Y component.
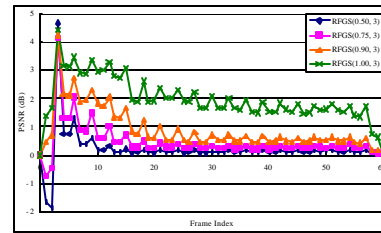

Figure 5. The error attenuation in PSNR for the Y component of the Akiyo sequence under different α in RFGS framework, where the pair of the values indicates the prediction mode parameters (α, β).

reduced to about 0.1 to 0.3 dB. When a equals to unity the drift lasts for a long period of time.

Reference

[1] A. Vetro, "AHG report on describing terminal capabilities for UMA, " MPEG01/m7275, Sydney, Australia, July 2001.

[2] W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," *IEEE Trans. Circuits Syst. Video Technol*., pp. 301-317, Mar. 2001.

[3] J. Magalhaes and F. Pereira, "Describing user environments for UMA applications," MPEG01/m7312, Sydney, Australia, July 2001.

[4] F. Hartung and M. Kutter, "Multimedia watermarking techniques, " *Proc. IEEE*, vol. 87, no.7, pp.1079-1107, July 1999.

[5] Y.-C. Lin, C.-N. Wang, Tihao Chiang, A. Vetro and H.F. Sun, "Efficient FGS to Single layer transcoding, " Published by ICCE2002.

[6] H.-C. Huang, C.-N. Wang, and Tihao Chiang, "A robust fine granularity scalability using trellis based predictive leak," Published by ISCAS2002.

[7] MPEG Video Group, "Draft Testing Procedures for evidence on scalable coding technology", *ISO/IEC JTC1/SC 29/WG 11 N4927*, July 2002.

[8] MPEG Requirements Group, "Draft Applications and Requirements for Scalable Coding," *ISO/IEC JTC1/SC 29/WG 11 N4984*, July 2002.