Available online at www.sciencedirect.com
ScienceDirect
Signal Processing: *Image Communication* 22 (2007) 505–514
SIGNAL PROCESSING: IMAGE COMMUNICATION
www.elsevier.com/locate/image

Design of frame dependency for VCR streaming videos

Kai-Chao Yang*, Chun-Ming Huang, Jia-Shung Wang

Department of Computer Science, National Tsing Hua University, Hsinchu 30043, Taiwan

Received 12 October 2006; received in revised form 13 March 2007; accepted 27 March 2007

Abstract

The structure of *Group of pictures* (GOP) is favorable for GOP-based random-access operations, while inside the GOP, the sequential coding dependency between frames is unfavorable for frame-based random-access or fast-scan operations. In a video streaming system, when some VCR operations such as fast-scan operations are performed, they may cause serious server load or network load due to the speed-up of transmitting frames. In this paper, we formulated the fast-scan and random-access operations of VCR functionality, and proposed a novel frame dependency to eliminate extra server load and minimize requirements of the network bandwidth when performing fast-scan and random-access operations. With the proposed structure, the server does not have to speed up the transmission rate of frames to fulfill the requested playback rate. The number of frames to be transmitted is also significantly reduced, and thus the network load can be decreased.

Keywords: Video coding; Motion estimation; Streaming video; GOP; VCR

1. Introduction

With video being a popular part of modern communications and information, many multimedia delivering applications such as interactive TV and video-on-demand (VOD) are more feasible. Today, the rapid development of high-speed access network makes the streaming video over the Internet be achievable and practical, actually it is increasingly becoming one of the most important and successful services. Streaming applications allow numerous clients to access as well as playback remote video programs from the server concurrently. A streaming system should provide some user-friendly services to make users more comfortable. For example, in a video streaming system, it is important to support *fast-scan* and *random-access* operations, so that a user can take a quick preview of the video content. However, these operations need transferring video content at a very high transmission rate corresponding to the playback rate. This certainly cause serious server load and bandwidth load for the video server and the involving networks, which restricts the service scale. Therefore, it is highly desirable that a server can afford to provide VCR operations. To reduce the server load and the network load, lots of multicast [1,16] or broadcast [9] methods have been proposed so that many clients can share the same video

*Corresponding author.
E-mail addresses: kcyang@vc.cs.nthu.edu.tw (K.-C. Yang), cmhuang@cic.org.tw (C.-M. Huang), jswang@cs.nthu.edu.tw (J.-S. Wang).

0923-5965/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.image.2007.03.005

stream. Nevertheless, a dedicated stream still has to be delivered from a video server to subscribed clients to provide a true VOD service [13]. To support VCR operations, several methods have also been presented, but most previous works solve the problem via large storage or network overhead. In [15], fast-scan operations are implemented at the client side using pre-fetched video frames. In [12], a GOP-skipping-based dynamic transmission scheme was proposed. In [14], a dual-bitstream approach was presented that combines a normal encoded and a reverse encoded bitstream to reduce the number of frames needed to be transmitted during a fast-scan period. However, the system cannot avoid delivering redundant frames because of the intrinsic property of MPEG-like temporal coding dependency among frames. Because of the increasing network load during fast-scan periods, multiple bitstreams encoded at different playback rates are usually provided in a low bit-rate video streaming system to support different fast-scan operations. However, this causes serious storage overhead if there are numerous videos in the server. Instead of the above schemes, a compressed-domain approach is proposed in [2] to support reverse playback. Recently, SP/SI frames proposed in H.264/AVC [10] provide immediately switching between two streams coded at multiple playback rates, but multiple bitstreams are still needed for different playback rates.

In [5], we have successfully reduced the number of redundant transmitted frames when performing specific fast-scan operations by reorganizing coding dependency among frames. In this paper, we propose another coding dependency among frames in a GOP for reducing the server load and the network load. Instead of the traditional coding dependency, frames are encoded with a special order, which avoids transmitting redundant frames when performing fast-scan and jump operations. With the new coding dependency, the transmission rate needs not to be speeded up during a fast-scan period, and thus saving bandwidth. As for the server side, only one bitstream is required for each video to support all playback rates. Besides, the proposed coding dependency can be easily modified from *Multiframe motion compensation* in H.264 [18].

In the remaining part of this paper, we address the problem of VCR functionality on the VOD systems in Section 2. A feasibility study on this problem is discussed in Section 3. Next, the proposed algorithm to reconstruct the temporal coding dependency is presented in Section 4. In Section 5, a proximity approximation is proposed to improve the transmission efficiency. Then we give several experimental results in Section 6 to demonstrate the proposed scheme. Finally, conclusions are given in Section 7.

2. Conventional GOP structure

An ordinary video coder usually compresses a video clip with MPEG format, where a video clip consists of a series of groups of pictures (GOP), and each GOP contains three types of frames, I-, P-, and B-frames. For video streaming applications, however, B-frames may cause unexpected delay, and thus PB-frames are used instead of B-frames in H.263 standard [8]. By the same policy in streaming video, we also suppose that a GOP contains only I- and P-frames, which is additionally supported by H.261 [7], MPEG-4 *Simple* profile [6], and H.264 *Baseline* profile.

In the conventional GOP structure, frames are encoded and decoded linearly. Excluding the I-frame, each frame is decoded by referring to its previous frame, as illustrated in Fig. 1, where $r \rightarrow t$ means frame t *refers* to frame r (or frame t is *compensated* from r). The conventional GOP structure can be viewed as a completely skewed tree, where each frame cannot be decoded until all of its ancestors are decoded.

If someone performs a fast-scan operation, due to the conventional coding dependency between frames, the server needs to transmit redundant frames to guarantee that all requested frames could be decoded successfully. For example in Fig. 1, suppose the requested frames in the GOP are 0, 3, 6, 9, and 12 (Triple speedup playback). All frames before frame 12 need to be sent for decoding of the requested frames, which causes a 2.6 times transmission overhead. According to our experimental results, the transmission rate and the number of redundant frames are roughly proportional to the displayed speed factors by the conventional GOP.

Random-access operation (or *jump* operation) is the special case of fast-scan operations. When a client requests for randomly accessing a frame, the

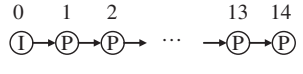

Fig. 1. The conventional frame dependency without B-frames. Each P-frame refers to the previous frame.