

Interactive Playout of Digital Video Streams

Samuel Rivas

LambdaStream Servicios Interactivos, Spain

Miguel Barreiro

LambdaStream Servicios Interactivos, Spain

Víctor M. Gulías

University of A Coruña, MADS Group, Spain

INTRODUCTION

Interactive Playout

Even though digital systems have many advantages over traditional analogue systems, end users expect that they will not lose any functionality in the transition. Concretely, capabilities of video cassette recorders (VCR) should be ported to stand alone digital media players (e.g., DVD players) and streaming services (e.g., digital TV and VoD services). Those capabilities are usually known as interactive playout or, simply, digital VCR functionality. Typical VCR features are random access, pause/resume, reverse play, fast forward/backward, and slow motion.

Random access and pause/resume functionalities are relatively easy to implement in digital systems. On the other hand, video coding techniques and bandwidth restrictions severely complicate the implementation of the other VCR operations.

Usually, VCR capabilities apply only to video streams. In interactive playout mode, audio streams are commonly discarded. Also, the quality of video streams in interactive mode may be downgraded due to system limitations.

Architecture Constraints

Both stand-alone and client-server systems have architectures similar to Figure 1. There is a server side, where video data is stored, a data path, and a client that receives and plays the video streams. The client has a decoder and a small buffer to temporarily store data received from the server until the decoder decodes and presents them. The client buffer is very small compared

to the server storage; it can only store a few seconds of video data.

For example, in a stand-alone DVD player, the data server is the DVD reader, the storage the DVD disk, the data path the internal memory bus, and the client is the internal decoder. In a **VoD** service for digital TV, the server and storage are the VoD server, the data path is the digital TV network, and the client is the set top box.

There are two main differences between stand-alone and client-server systems. First, the bandwidth is much higher for data paths in stand-alone systems. Second, in client-server systems, many clients share the data server computational power.

Decoders are usually based on standards. This imposes some restrictions on the behaviour they can implement to support VCR functionality. For MPEG-2 (ISO/IEC, 1995) based decoders, an important restriction is that frame rate is fixed to a set of available values. Later MPEG standards (ISO/IEC, 1999, 2005) support arbitrary frame rates.

For the data path, the main constraint is bandwidth, especially in client-server applications.

In the server side, the main constraints are storage size and computational power.

Motion Compensation and Temporal Dependencies

Due to the high bandwidth and storage requirements of uncompressed multimedia data, compression techniques (Shi & Sun, 2000) are usually mandatory in digital multimedia systems.

Modern video compression techniques use a combination of lossy image compression and motion

compensation. **Motion compensation** introduces dependencies among coded frames that impose restrictions in the decoding process.

There are two dependency types: **forward prediction**, if they refer to a frame previous in time, or **backward prediction**, if they refer to a frame in the future. MPEG standards classify coded pictures in three types:

- **I pictures or intra coded pictures:** Pictures that do not have any coding dependency with other pictures.
- **P pictures or predictive coded pictures:** Pictures that may have coding dependencies with past pictures.
- **B pictures or bidirectionally predictive coded pictures:** Pictures that may have coding dependencies with past pictures, future pictures, or both.

Figure 2 depicts frame dependencies for a typical MPEG-2 sequence. In the example, a decoder must decode P_6 and P_3 before decoding B_5 . To decode P_6 , the decoder must have decoded P_3 , and to decode P_3 , it must have decoded I_0 . Because only I pictures are independent, a decoder must receive an I picture to start decoding.

The main difference between P and B pictures is that B pictures make decoding order different from presentation order. For the sequence in Figure 2, the decoding order is $I_0, P_3, B_1, B_2, P_6, B_4, B_5, P_9, B_7, B_8, I_{10}, \dots$

Any P or B picture depends either directly or indirectly on one I picture. An I picture and all the pictures that depend on it form a closed group of related pictures. This is what MPEG-2 calls **group of pictures (GOP)**. For example, a MPEG-2 sequence IBBPBBPIBBPBBP (in display order) is divided into two GOPs with the same structure: IBBPBBP.

For simplicity, we use MPEG-2 stream structure as reference through the rest of this article. That is, I pictures with no coding dependencies, P pictures that depend on their previous I or P picture, and B pictures that depend on their previous and next nonB pictures. We write all sequences in display order.

RANDOM ACCESS

Locating Access Points

Random access functionality enables end users to select a point in time to start playing a requested media. It is possible, and often required, to implement random

Figure 1. Multimedia service architecture

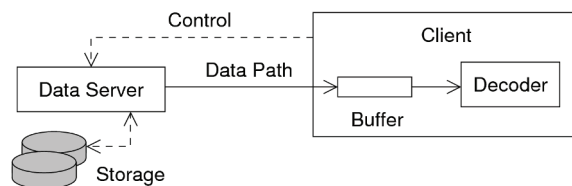
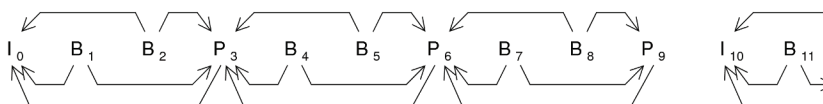


Figure 2. Motion composition frame dependencies in MPEG-2



6 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/interactive-playout-digital-video-streams/17475

Related Content

Spatio-Temporal Analysis for Human Action Detection and Recognition in Uncontrolled Environments

Dianting Liu, Yilin Yan, Mei-Ling Shyu, Guiru Zhao and Min Chen (2015). *International Journal of Multimedia Data Engineering and Management* (pp. 1-18).

www.irma-international.org/article/spatio-temporal-analysis-for-human-action-detection-and-recognition-in-uncontrolled-environments/124242

Improving Online Readability and Information Literacy

John Paul Loucky (2008). *Handbook of Research on Digital Information Technologies: Innovations, Methods, and Ethical Issues* (pp. 284-305).

www.irma-international.org/chapter/improving-online-readability-information-literacy/19849

Designing and Randomising Multiple-Choice Questions for E-Learning in Mathematics Using MVGEN

Irene Brito, Gaspar J. Machado and José João Almeida (2022). *Online Distance Learning Course Design and Multimedia in E-Learning* (pp. 205-224).

www.irma-international.org/chapter/designing-and-randomising-multiple-choice-questions-for-e-learning-in-mathematics-using-mvgen/299838

The Research on Shape Context Based on Gait Sequence Image

Rong Wang, Yongkang Liu and Mengnan Hu (2018). *International Journal of Multimedia Data Engineering and Management* (pp. 21-35).

www.irma-international.org/article/the-research-on-shape-context-based-on-gait-sequence-image/201914

Emoticon Recommendation System to Richen Your Online Communication

Yuki Urabe, Rafal Rzepka and Kenji Araki (2014). *International Journal of Multimedia Data Engineering and Management* (pp. 14-33).

www.irma-international.org/article/emoticon-recommendation-system-to-richen-your-online-communication/109076