Chapter 3

2-D Scalable Multiple Description Coding for Robust H.264/SVC Video Communications

Wei Xiang
University of Southern Queensland, Australia

Ce Zhu
Nanyang Technological University, Singapore

Chee Kheong Siew
Nanyang Technological University, Singapore

Yuanyuan Xu
Nanyang Technological University, Singapore

Minglei Liu
Nanyang Technological University, Singapore

ABSTRACT

In this chapter, we investigate two popular techniques for error-resilient H.264/SVC video transmission over packet erasure networks, i.e., layered video coding (LVC) and scalable multiple description coding (SMDC). We compare the respective advantages and disadvantages of these two coding techniques. A comprehensive literature review on latest advancement on SMDC is provided. Furthermore, we report new simulation results for the novel two-dimensional scalable multiple description coding (2-D SMDC) scheme proposed in our previous work (Xiang et al., 2009). The 2-D SMDC scheme allocates multiple description sub-bitstreams of a two-dimensionally scalable bitstream to two network paths with unequal loss rates. We formulate the two-dimensional scalable rate-distortion problem and derive the expected distortion for the proposed scheme. To minimize the end-to-end distortion given the total rate

DOI: 10.4018/978-1-61692-831-5.ch003
INTRODUCTION

In recent years, the increasing demand for mobile broadband access to multimedia applications and services has motivated enormous interest in media streaming over the Internet. Media streaming is characterized by high data transmission rate and low delay constraint. Therefore, its applications are highly susceptible to packet delay and delay jitter. However, packet erasure networks such as the Internet only offers best-effort service, and thus cannot guarantee bandwidth and delay. Conventional approaches for tackling packet loss such as retransmission is not effective for streaming media due to the real time nature of the service. Thus, additional mechanisms are needed to provide streaming media delivery over packet erasure networks.

Traditionally, layered video coding (LVC), which is also known as scalable video coding (SVC) with transport prioritization (Chakareski, 2005), is the most popular and effective approach for video transmission over error-prone networks. In LVC, a raw video sequence is coded into a base layer that provides a coarse level of visual quality and can be decoded independently, and multiple enhancement layers that refine the base-layer visual quality and are nevertheless useless alone. SVC has the advantage of enabling media providers to generate a single embedded bitstream from which appropriate subsets can be extracted to meet various requirements of a broad range of clients. SVC is thus essential for multicast applications, where a variety of end users with different capabilities such as bandwidth and processing power receive different presentations of the same media content. In general, temporal scalability, spatial scalability, and signal-to-noise-ratio (SNR) scalability are among the most common scalability mechanisms for SVC. These mechanisms provide for frame rate, spatial resolution, and video quality adaptation to adjust the video source in accordance with device capability and channel bandwidth.

However, the fact that enhancement layers provide little refinement information if the base layer is not received or decoded correctly, is the root cause that scalable video bitstreams are extremely vulnerable to transmission errors. Therefore, in an error-prone environment such as the Internet, providing error resilience against packet errors is as important as enabling scalability. LVC is usually used in conjunction with unequal loss protection (ULP) (Mohr et al., 2000), which provides the base layer with the highest level of channel error protection through the use of forward error correction (FEC) coding.

The major weakness of LVC lies in its excessive reliance on the correct receipt and decoding of the base layer. As an alternative, multiple description coding (MDC) (Gamal & Cover, 1982), (Goyal, 2001) has recently emerged as a promising alternative technique for providing graceful degradation of performance in the presence of channel noise. The essential idea underlying MDC is to generate multiple (> 2) independent descriptions of a source such that each description independently describes the source with a certain desired fidelity. When more than one descriptions are received, they can be synergistically combined to enhance the quality. Therefore, the distortion in the reconstructed signal decreases upon the receipt of any additional descriptions, and is lower bounded by the distortion attained by single description coding.

In recent years, a plethora of different MDC methods have been proposed in the literature. The
Related Content

Semantic-Based Video Transcoding Architectures for Quality of Service Applications in Mobile and Wireless Video Communication
Ashraf M.A. Ahmad (2009). Multimedia Transcoding in Mobile and Wireless Networks (pp. 142-159).
www.igi-global.com/chapter/semantic-based-video-transcoding-architectures/27199?camid=4v1a

Constructing and Utilizing Video Ontology for Accurate and Fast Retrieval
Kimiaki Shirahama and Kuniaki Uehara (2011). International Journal of Multimedia Data Engineering and Management (pp. 59-75).
www.igi-global.com/article/constructing-utilizing-video-ontology-accurate/61312?camid=4v1a

Modelling eCRM Systems with the Unified Modelling Language
www.igi-global.com/chapter/modelling-ecrm-systems-unified-modelling/17313?camid=4v1a

Multimodal Information Fusion of Audiovisual Emotion Recognition Using Novel Information Theoretic Tools
Zhining Xie and Ling Guan (2013). International Journal of Multimedia Data Engineering and Management (pp. 1-14).
www.igi-global.com/article/multimodal-information-fusion-of-audiovisual-emotion-recognition-using-novel-information-theoretic-tools/103008?camid=4v1a