Reducing Processing Demands for Multi-Rate Video Encoding: Implementation and Evaluation

Håvard Espeland, University of Oslo and Simula Research Laboratory, Norway
Håkon Kvale Stensland, University of Oslo and Simula Research Laboratory, Norway
Dag Haavi Finstad, University of Oslo and Simula Research Laboratory, Norway
Pål Halvorsen, University of Oslo and Simula Research Laboratory, Norway

ABSTRACT

Segmented adaptive HTTP streaming has become the de facto standard for video delivery over the Internet for its ability to scale video quality to the available network resources. Here, each video is encoded in multiple qualities, i.e., running the expensive encoding process for each quality layer. However, these operations consume both a lot of time and resources, and in this paper, the authors propose a system for reusing redundant steps in a video encoder to improve the multi-layer encoding pipeline. The idea is to have multiple outputs for each of the target bitrates and qualities where the intermediate processing steps share and reuse the computational heavy analysis. A prototype has been implemented using the VP8 reference encoder, and their experimental results show that for both low- and high-resolution videos the proposed method can significantly reduce the processing demands and time when encoding the different quality layers.

Keywords: Multimedia, Multi-Rate, Performance, Quality Adaption, Video Encoding, VP8

INTRODUCTION

The amount of video data available in the Internet is exploding, and the number of video streaming services both live and on-demand, is quickly increasing. For example, consider the rapid deployment of public available Internet video archives providing a wide range of content like newscasts, movies and scholarly videos. In this respect, Internet users uploaded one hour of video to YouTube every second in January 2012 (YouTube, 2012). Furthermore, all major (sports) events like European soccer leagues, NFL Hockey, NBA basketball, NFL football, etc. are streamed live with only a few seconds delay, e.g., bringing the 2010 Winter Olympics (Zambelli, 2009), 2010 FIFA World Cup (Move Networks, 2008) and NFL Super Bowl (Move Networks, 2008) to millions of concurrent users over the Internet, supporting a wide range of devices ranging from mobile phones to HD displays. The number of videos streamed from such services are in the order of tens of billions

DOI: 10.4018/jmdem.2012040101
per month (Flosi, 2010; YouTube, 2012), and leading industry movers conjecture that traffic on the mobile-phone networks will also soon be dominated by video content (Cisco Systems, Inc., 2010).

Adaptive HTTP streaming is frequently used in the Internet and is currently the de facto video delivery solution. For example, Move Networks (2008) was one of the first providers of segmented adaptive HTTP streaming, and has later been followed by major actors like Microsoft (Zambelli, 2009), Apple (Pantos et al., 2010) and Adobe (2010). Recently, there are also standardization efforts by MPEG (Stockhammer, 2011). In these systems, the bitrate (and thus video quality) can be changed dynamically to match the varying bandwidth, giving a large advantage over non-adaptive systems that are frequently interrupted due to buffer underruns or data loss. The video is thus encoded in multiple bitrates matching different devices and different network conditions.

Today, H.264 is the most frequently used codec. However, an emerging alternative is the simpler VP8 which is very similar to H.264’s baseline profile and supposed to be well suited for web-streaming with native support in major browsers, royalty free use and similar video quality as H.264 (Seeling et al., 2010; Ozer, 2010). Nevertheless, for both codecs, the challenge in the multi-rate scenario is that each version of the video requires a separate processing instance of the encoding software. This may be a challenge in live scenarios where all the rates must be delivered in real-time, and in the YouTube case, it will just take an enormous data center to keep the upload rate. Thus, the process of encoding videos in multiple qualities and data rates is both time and resource consuming.

Our initial idea was presented by Finstad et al. (2011) and in this paper, we expand our evaluation and further discuss our results. In particular, we analyze the processing overheads of multi-rate scenarios, and to reduce resource requirement, we investigate possibilities for reusing the output from different steps in the encoding pipeline as the same video elements are processed multiple times with only slightly different parameters. We use the VP8 processing pipeline as a case study, and we present a prototype that supports running multiple parallel VP8 encoder instances. The main contributions of our work are:

- Inspired by several transcoding approaches trying to reuse motion vectors (Kuo & Jayant, 2003; Zhou, Zhou, & Xia, 2008; Senda & Harasaki, 1999), we propose a way of reusing decisions from intra and inter prediction in a video encoder to avoid computational expensive steps that are redundant when encoding for multiple target bitrates, i.e., macroblock mode decision, intra prediction and inter prediction between the instances. The method can be used in any video codec comprising an analysis and encoding step with similar structure as H.264 and VP8.
- A prototype and demonstrator has been implemented using the VP8 reference encoder as a case study. We have performed a wide range of experiments using both low- and high-resolution videos with various data rates and content types. The results show that the computational demands are significantly reduced at the same rates and approximately the same qualities as the VP8 reference implementation.

RELATED WORK

Our proposed multi-rate encoder is based on the idea of sharing and reusing the results from the computational expensive steps. The idea is inspired by the Scalable Video Coding technique (Schwarz et al., 2007), but differs in key aspects: First, no special client player support is necessary. SVC has very little market penetration in comparison to VP8 player availability, which in contrast makes the multi-rate technique useful without requiring changes for the users. Second, the layers of SVC affect the output quality (Wien et al., 2007), which also may be a reason why it has not been widely adopted for online streaming. In our approach
Automatic Pitch Type Recognition System from Single-View Video Sequences of Baseball Broadcast Videos
Masaki Takahashi, Mahito Fujii, Masahiro Shibata, Nobuyuki Yagi and Shin’ichi Satoh (2012). Methods and Innovations for Multimedia Database Content Management (pp. 119-142).
www.igi-global.com/chapter/automatic-pitch-type-recognition-system/66691?camid=4v1a

A Real-Time 3D Visualization Framework for Multimedia Data Management, Simulation, and Prediction: Case Study in Geospatial-Temporal Biomedical Disease Surveillance Networks
www.igi-global.com/article/real-time-visualization-framework-multimedia/54459?camid=4v1a