

Video Communication Systems with Heterogeneous Clients

Cheng-Hsin Hsu and Mohamed Hefeeda
School of Computing Science
Simon Fraser University
Surrey, BC, Canada

ABSTRACT

Modern wireless mobile devices have evolved to small computers that can render multimedia content, while desktop/laptop computers have become more computationally powerful with faster Internet access. As these computing devices getting more popular, users demand for more and higher quality videos in many communication applications, where clients are heterogeneous in terms of network bandwidth and computing power. The goal of this thesis is to improve client perceived-quality of various video communication systems by adopting scalable video coding tools that enable efficient rate adaptation. We seek to understand scalable coding standards and design optimization and streaming algorithms to make the best possible use of them in practical systems. We consider practical problems of video communication systems in three different environments: Internet streaming systems, TV broadcast networks, and mobile video communication systems. We propose efficient algorithms to solve the considered problems. We evaluate the proposed algorithms using numerical methods and/or simulations. Most importantly, we design and implement testbeds to validate our algorithms. The expected results of applying our algorithms to video communication systems are better video quality and higher user satisfaction as well as better bandwidth utilization and lower processing overhead.

Categories and Subject Descriptors: C.2[Computer Systems Organization]: Computer-Communication Networks

General Terms: Design

1. INTRODUCTION

With recent hardware advances, wireless mobile devices, such as personal digital assistants (PDAs), smart cellular phones, and mobile media players have evolved to small mobile computers that can render multimedia content, while desktop/laptop computers have become computationally powerful with increasingly faster Internet access connections. As these computing devices getting very popular, users are encouraged for getting more and higher quality videos in a full spectrum of communication applications including video streaming (on-demand or live), broadcasting, and video conferencing. In such systems, clients are heterogeneous in terms of both

network bandwidth and processing power, and this heterogeneity is increasing as new technologies are made possible everyday. Therefore, concurrently supporting heterogeneous clients and networks is very challenging when designing video communication systems, and is critical to the success of such systems. The goal of this thesis is to improve client perceived-quality of video communication systems by adopting modern video coding tools.

Traditional, non-scalable, video coders optimize each coded stream at a single bit rate and are not suitable to video communication applications because clients have to meet a minimum requirement to receive and decode the video while clients with more resources cannot get better quality. In contrast, scalable coders produce a single stream that can easily be customized to serve heterogeneous clients over dynamic networks. Scalable video coders compress video data into a base layer that provides basic quality, and one or more enhancement layers that add incremental quality refinements. Modern scalable video coders allow the enhancement layers to be either coarse-grained scalable (CGS), medium-grained scalable (MGS), or fine-grained scalable (FGS), where CGS layers provide the most-limited rate scalability and FGS layers enable very flexible rate adaptations. Scalable coding supports wider ranges of clients, and it can fully utilize available bandwidth of individual clients, which results in better video playback quality and higher user satisfaction.

While scalable coding improves rate scalability, it introduces many new challenging problems that need to be addressed to be usable in practical video communication systems. We seek to understand scalable coding standards and to design optimization and streaming algorithms to make the best possible use of them in practical systems. We consider video communication systems in three different environments. First, we study Internet streaming applications where most clients are powerful personal computers. We study the behaviors of scalable coding in general and of FGS coding in particular. We consider two problems in Internet streaming using scalable coding: optimal rate-allocation in multi-sender to a single-receiver video streaming and optimal structuring for scalable coded streams. Second, we study the mobile TV broadcast networks where a common air medium is shared among many TV channels. We consider two unique problems in mobile TV broadcast networks: optimal energy saving for mobile devices and bounding channel switching delay without compromising energy saving and network utilization. Third, we study mobile video communication systems where handheld, resource limited, devices can capture, encode, and stream videos over the air among each other. We consider two problems in mobile video communication: resource allocation for multiple mobile devices streaming videos over a shared single-hop wireless network and mobile video conferencing using distributed video coding.

For each of these considered problems, we propose efficient algorithms to solve it. We rigorously evaluate the proposed algorithms using numerical methods and/or simulations. Most importantly, we design and implement testbeds to validate our algorithms. The expected results of applying our algorithms to video communication systems are better video quality and higher user satisfaction as well as better bandwidth utilization and lower processing overhead.

2. RESEARCH DIRECTIONS

We consider various challenging problems in three types of video communication systems where scalable video coding can be adopted to improve user experience.

2.1 Internet Streaming Systems

Internet video streaming to personal computers is getting increasingly popular, and users continuously demand for better video streaming quality. We study the potential of adopting scalable coding, in particular FGS coding, to improve video streaming quality in such systems. In [5], we study rate-distortion (R-D) models, which are functions describing the relationship between streaming rates and expected levels of distortion in reconstructed videos. R-D models enable optimization of the received video quality in video communication systems. We quantitatively evaluate FGS R-D models by considering several performance metrics including accuracy, range of applicability, and time complexity. Based on our experimental results, we present recommendations on selecting the best R-D models for various applications. The insights developed while analyzing R-D models enable us to design efficient algorithms to optimize video quality.

In [1,2], we consider optimal streaming of scalable video streams from multi-senders to a single receiver. Streaming from multiple senders is necessary in peer-to-peer (P2P) streaming environments, because of the limited capacity and unreliability of peers. Multiple senders are also desired in distributed streaming systems to achieve disjoint network path streaming and hence better quality. We first formulate the problem for individual frames, then we generalize it for a block of multiple frames. We use R-D models in our formulations to transform nonlinear problems to linear programming (LP) ones, which can be solved efficiently using common optimization techniques such as the Simplex method. We also propose efficient algorithms for the single-frame allocation problem that is provably optimal, and for the multiple-frame case that produces near-optimal solutions in real time. Our experimental study on a diverse set of real video streams validates our analysis and shows the effectiveness of our allocation algorithms in improving the video quality: an improvement of up to 8 dB could be achieved.

In [3, 4, 6, 7], we study optimal structuring of scalable video streams for Internet video streaming to many receivers with heterogeneous bandwidth. Recent video coders, such as H.264/SVC, can encode a video sequence into multiple layers, each with a different rate. Moreover, each layer can either be CGS or FGS. FGS layers support wider ranges of client bandwidth than CGS layers, but suffer from lower coding efficiency. There were no systematic ways in the literature to determine the optimal stream structure that renders the best average quality for all clients. We formulate an optimization problem to determine the optimal rate and encoding granularity (CGS or FGS) of each layer in a scalable video stream that maximizes a system-defined utility function for a given client distribution. We design an efficient, yet optimal, algorithm to solve this optimization problem. Our algorithm is general in the sense that it can employ arbitrary utility functions for clients. We implement our algorithm and verify its optimality. We show how dif-

ferent structures of scalable video streams affect individual client utilities. We compare our algorithm against a heuristic algorithm that has been used before in the literature. We show that our algorithm outperforms the other one in all cases: e.g., our algorithm achieves on average more than 95% bandwidth utilization while the other one only achieves less than 60%.

Our Internet streaming algorithms are being implemented and field-tested using a large-scale content distribution system, called peer-assisted content distribution networks (pCDN) [8]. pCDN is developed in partnership with the Canadian Broadcasting Corporation (CBC/Radio-Canada) to leverage on the P2P computing paradigm beyond file-sharing applications and provide a new generation of content delivery platforms. This system, which distributes multimedia objects to millions of clients scattered all over the Internet, enables us to validate our optimization and streaming algorithms using actual video content and real user demands.

2.2 Mobile TV Networks

Mobile TV allows users to watch their favorite TV shows and games on small handheld devices on the move. Mobile TV trial or commercial services have been deployed in more than 40 countries and Mobile TV market is expected to grow to 20 billion Euros with 500 million customers by 2011, which motivates the European Union to adopt Digital Video Broadcast-Handheld (DVB-H), one of the most widely used mobile TV standards, as its official standard. We study two fundamental problems in mobile TV networks: maximizing energy saving and bounding channel switching delay for mobile devices.

A common issue in all mobile wireless devices is the limited energy supply since they are battery powered. Thus, minimizing energy consumption in mobile TV networks becomes a critical problem for the success and wide adoption of such systems. We consider energy optimization in mobile TV networks in which a base station concurrently broadcasts multiple TV channels to mobile devices over a common wireless medium. In these systems, the base station broadcasts TV channels in bursts with bit rates much higher than the encoding rates of the video streams. Thus, mobile devices can receive a burst of video data and then turn off their radio frequency circuits till the next burst in order to save energy, which is referred to as time slicing. To achieve this energy saving without scarifying streaming quality, the base station must carefully construct the burst schedule for all TV channels. Our problem is to determine optimal burst schedules for broadcasting multiple TV channels in order to minimize the energy consumption of mobile devices. We prove that this burst scheduling problem is NP-complete. We propose an optimal and efficient burst scheduling algorithm for mobile TV networks, where TV channels are classified into multiple classes and each class has a different bit rate. The bit rate of class c , r_c , can take any value in the form of $r_c = 2^i \times r_1$ ($r \in \mathbb{N}$) and r_1 is the bit rate of the lowest class. Our algorithm is provably optimal and runs in $O(S \log S)$ time, where S is the number of TV channels. We also solve the general burst scheduling problem where TV channels can have arbitrary bit rates. Since the general problem is NP-complete, our solution is an approximation algorithm that is efficient and achieves near-optimal energy saving. We validate correctness and optimality of our algorithms using numerical methods and simulations. We are integrating our algorithms into a real mobile TV testbed, which is described later.

Although time slicing enables mobile devices to save energy, it also increases channel switching delay, which is the time a user waits before s/he starts viewing a new channel when a change of channel is requested by that user. Short channel switching delay is important to user experience because many users quickly flip

through many TV channels before they decide to watch a specific one. Our problem is to bound channel switching delay of a mobile TV network that employs time slicing technique to save energy. Our goal is to compose time slicing schemes for switching delay guarantee while achieving optimal energy saving. Simultaneously achieving short switching delay and high energy saving is no easy task because there is a trade-off between them: higher energy saving in general leads to longer switching delay. To address this issue, we propose to simulcast each TV channel over two burst trains, where a burst train is a series of data bursts. One train, called bootstrap train, is optimized for fast channel switching, and the other train, called primary train, is optimized for energy saving. Once a user switches to a new TV channel, his/her mobile device first tunes to the bootstrap train for shorter delay. The mobile device then switches to the primary train upon the first primary burst is available for higher energy saving. Based on this idea, we design new time slicing schemes that ensure a given maximum switching delay is not exceeded. Our schemes are optimal in terms of energy saving. We implement our time slicing schemes in a mobile TV testbed. The experimental results validate our theoretical analysis and show that our proposed schemes meet the given maximum delay and achieves significant energy saving for mobile devices.

We implement a mobile TV testbed for DVB-H networks. This testbed consists of a Linux workstation the runs DVB-H and IP Datacast stacks and acts as the base station. IP Datacast standard is complementary to DVB-H and is designed for end-to-end mobile TV solutions. Some of the software components in our DVB-H base station are being developed by us and others are modified from open-source projects. We use cellular phones and a DVB-H packet analyzer as the receivers, which allow us to monitor the client and network performance. Furthermore, we are integrating H.264/SVC streaming servers and players into our mobile TV testbed for a complete system. We use this testbed to evaluate energy saving achieved by our burst scheduling algorithms and effectiveness of our time slicing schemes on bounding channel switching delays.

2.3 Mobile Video Communication Systems

As wireless networks provide better quality-of-service and higher throughputs, video communication using wireless networks has become an important application. We study video communication in wireless networks with resource limited handheld devices.

We consider a resource allocation problem to optimize the quality of video communication in single-hop wireless networks that are composed of multiple resource limited wireless devices. We address the video optimization problem in wireless networks in two steps. First, we formulate an abstract optimization problem for wireless networks in general. This formulation reveals the important interaction among different layers in the network stack. Then, we instantiate and solve the general problem for IEEE 802.11e WLANs. We show how the calculated optimal solutions can efficiently be implemented in the distributed mode of the IEEE 802.11e standard. We evaluate our proposed solution using OPNET simulator, which captures most features of realistic wireless networks. We also implement our solutions in the driver of an off-the-shelf wireless adapter to show its practicality, and we conduct experiments on a WLAN testbed. Our algorithm results in significant quality improvement for streamed videos.

While conducting the above experiments on our WLAN testbed, we notice that modern video encoders are too complex for mobile wireless devices. This is because these video coders assume video sequences are encoded once and decoded many times as in TV broadcast, Internet streaming, video-on demand, and multimedia storage applications, where complex and lengthy encoding

process for better coding efficiency is justified. In the past few decades, video coding standards have gradually put more computationally complex algorithms, especially for motion estimation, into encoders for better coding efficiency. These video encoders eventually become too complex and are not suitable to mobile applications, where mobile devices are resource limited.

Consider a mobile surveillance system as an example, where several wireless video cameras capture surveillance videos and transmit these videos to an archive server. Typical wireless video cameras have inadequate memory size, computational power, and battery capacity and cannot handle complex video encoders. In contrast, the archive server have much more computational power as they are stationary. Therefore, it is rationale and desirable to push the video coding complexity from the mobile video cameras toward the archive server. More precisely, mobile surveillance systems need video coding systems with simple encoders and complex decoders, which is the reverse of current video coding standards. Video coding systems that can simplify video coder complexity by offloading some or all computations to decoder, are referred to as Winer-Ziv, or distributed, video coders in the literature. Winer-Ziv coders provide another format of scalability, which we call it complexity scalability. This is because Winer-Ziv coders allow video communication systems to dynamically redistribute computational complexity between senders and receivers.

While Winer-Ziv coders appear to be ideal for mobile applications, very few quantitative results have been reported on their practicality. In fact, many of the proposed decoders are excessively complex and are not deployable on commodity computers. We are investigating the decoder complexity of Winer-Ziv video coding algorithms in the literature. Our goal is to design and implement a working (rather than quality optimized) Winer-Ziv coding system, in which video streams can be encoded by handheld devices and decoded by desktop/laptop computers in real-time. The following research steps have been identified, for constructing such a coding system: (1) surveying practical Winer-Ziv coding designs in the literature, (2) implementing representative Winer-Ziv coders, (3) quantitatively evaluating the performance of these coders in terms of coding efficiency and encoder/decoder complexity. (4) conducting a system profiling to identify bottleneck components in terms of computational complexity, and (5) proposing and implementing our Winer-Ziv coder for real mobile communication systems.

3. REFERENCES

- [1] M. Hefeeda and C. Hsu. Rate-distortion optimized streaming of fine-grained scalable video sequences. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 4(1):2:1–2:28, January 2008.
- [2] C. Hsu and M. Hefeeda. Optimal bit allocation for fine-grained scalable video sequences in distributed streaming environments. In *Proc. of ACM/SPIE Multimedia Computing and Networking (MMCN'07)*, pages 1–12, San Jose, CA, January 2007.
- [3] C. Hsu and M. Hefeeda. Optimal partitioning of fine-grained scalable video streams. In *Proc. of ACM International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV'07)*, pages 63–68, Urbana-Champaign, IL, June 2007.
- [4] C. Hsu and M. Hefeeda. Structuring multi-layer scalable streams to maximize client-perceived quality. In *Proc. of IEEE International Workshop on Quality of Service (IWQoS'07)*, pages 182–187, Evanston, IL, June 2007.
- [5] C. Hsu and M. Hefeeda. On the accuracy and complexity of rate-distortion models for FGS-encoded video sequences. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 4(2):15:1–15:22, May 2008.
- [6] C. Hsu and M. Hefeeda. Optimal coding of multi-layer and multi-version video streams. *IEEE Transactions on Multimedia*, 10(1):121–131, January 2008.
- [7] C. Hsu and M. Hefeeda. Partitioning of multiple fine-grained scalable video sequences concurrently streamed to heterogeneous clients. *IEEE Transactions on Multimedia*, 10(3):457–469, April 2008.
- [8] B. Jules and M. Hefeeda. pCDN: Peer-assisted content distribution network. *CBC/Radio-Canada Technology Review Magazine*, 4:1–14, July 2007.