Perception-Aware Low-Power Media Processing for Portable Devices

Ye Wang (IEEE Member), National University of Singapore, Singapore
wangye@comp.nus.edu.sg

1. Introduction

Smart mobile devices are proliferating, at increasingly smaller sizes. Many of these are capable of capturing and playing a significant quantity of audio and video, as well as uploading and downloading media to social networks. However, there has been relatively slow progress in improving the power to weight ratio in batteries for portable devices. As such, research efforts have been directed at developing techniques to reduce the power consumption of applications running on portable devices. Our group has explored novel methods and techniques addressing the fundamental tension between computational workload (and hence power consumption) and user perceived quality of service (QoS). In this letter, I summarize our findings, which are the outcome of a research grant (R-252-000-236-112) from the Ministry of Education in Singapore.

2. Techniques explored

Remove irrelevant computation

The computational workload of a portable multimedia device can be effectively reduced by identifying and removing information irrelevant to human perception. We identified elements in the media that were irrelevant or redundant, thereby also identifying unnecessary processes in the associated computations. Then we designed a new workload-scalable audio decoding scheme that would enable users to control the tradeoff between playback quality and power consumption in battery-powered portable audio players [1][2]. Our objective was to give users a control at the decoder side, similar to the Long Play (LP) recording mode on the encoder side in many media recording devices.

A standard single-layer audio format, such as that found in MP3, decodes the entire frequency band, corresponding to a high computational workload. We have developed a Bandwidth and Stereo-image Scalable (BSS) decoding scheme. The algorithm is based on an analysis of the perceptual relevance of the various audio components in the compressed bitstream. The bandwidth and stereo-image scalability directly translates into scalability of the computational workload generated by the decoder. This can be exploited by a voltage/frequency scalable processor to save energy and prolong battery life.

The workload scalable audio decoding scheme allows the user to switch between multiple output quality levels for a single-layer audio format such as MP3. Each level is associated with a different degree of power consumption, and hence battery life.

Remove redundant computation

An important trend in the consumer electronics and mobile handset industries is that the number of components and functions being integrated into a single device are steadily increasing. One could expect that the computational workload, and thus the energy consumption, of a mobile device increases linearly with the number of integrated components and functions. Our work has shown that system level computational workload and energy consumption can be effectively reduced, for example by using motion sensors to decrease the workload of a video encoder [3][4].

We believe that there is plenty of redundant computation in a multi-component multimedia device such as a videophone that can be removed by exploiting the methodology presented in these papers.

Decoding workload modeling and its applications to video encoding and transcoding

Dynamic voltage scaling (DVS) algorithms for multimedia applications have recently been a subject of intensive research. Many of these algorithms use control-theoretic feedback techniques to predict future execution demand of an application based on past execution demand. Such techniques suffer from two major disadvantages: (i) they are computationally expensive; and (ii) it is difficult to provide performance or QoS guarantees, since predictions can be incorraect. To address these shortcomings, we proposed a new approach for dynamic voltage and frequency scaling [5]. Our technique is based on an offline bitstream analysis of multimedia files. The analysis is used to insert metadata describing the computational demand that will be generated when decoding the file. Such bitstream analysis and metadata insertion can be done when the multimedia file is being downloaded into a portable device from a desktop computer.

As a proof-of-concept, we implemented this technique using an MPEG-2 decoder. Our experiments showed that the amount of metadata needed is a small fraction of the total size of the video clip, and that this technique can lead to significant energy savings. The inserted metadata will typically describe the frequency values at which the processor needs to be run at different points in time during the decoding process. In contrast to runtime prediction-based techniques, our scheme can provide performance and QoS guarantees, while avoiding runtime computation overhead.

We developed this research further by proposing a new workload prediction model, which is fast, accurate and generic enough to apply to different video formats, decoder implementations, and target platforms [6]. The model was used to design a new workload-scalable transcoding scheme that converts a prerecorded video bitstream into a new video bitstream. The new bitstream can satisfy a target decoding device's workload constraint, while keeping transcoding distortion minimal. We also extended this work by designing a novel decoding-workload-aware video encoding scheme [7]. The scheme takes the raw video data and decoding workload limit of a mobile client as input, and generates a video bitstream that strives to achieve best video quality while matching the client's constraints.

Media Server supported offline workload analysis and optimal speed control for reducing power consumption

To further reduce the power consumption of mobile devices, we leverage the computational power of the media server [8]. Most existing DVS techniques are suboptimal in achieving energy efficiency when providing the guaranteed playback QoS, mainly due to the inherent limitations of client-only approaches. To address this problem, we investigated the possibility of using smoothing mechanisms with media server supported DVS techniques. We proposed a generic offline bitstream analysis framework, and an optimal speed control algorithm, that achieve maximal energy savings among all feasible speed profiles for given buffers. The proposed scheme enables us to achieve a theoretical lower bound of buffer size for a given media clip. There are four significant aspects to our scheme that support practical implementation. First, it does not require a priori information about a client's configuration, rendering it particularly suitable for broadcast or multicast applications. Second, a speed profile based on buffer sizes can provide satisfactory energy efficiency. Third, the required buffer sizes are so small that they can be met by most mobile devices. Fourth, the additionally generated information (i.e., the speed profile information) is negligible compared to the size of the media content. These properties solve the diversity and feasibility issues of media server supported DVS schemes. Experimental results show that, in comparison with other existing techniques, our scheme improves the performance of DVS significantly.

Joint encoder-decoder framework (JEDF)

In a typical application scenario, music audio is produced once but may be played back multiple times across many devices. To reduce computation and power consumption of the decoders, we introduced techniques to shift a portion of the computation to the encoder or server side.

In comparison with the relatively slow progress of battery technology, semiconductor memory has improved rapidly, making storage a less critical factor when designing low power mobile devices. To exploit this trend, we proposed a joint encoder-decoder framework (JEDF), which allows the decoder to trade off energy and memory consumption without sacrificing playback quality [9]. We employed a sum-of-powers-of-two (SOPOT) technique, an approximate signal processing (ASP) technique, in an MPEG AAC decoder to reduce its computational workload. However, the SOPOT introduces ASP noise (in the decoder) along with the quantization noise introduced by lossy compression (in the encoder). The sum of these may become audible if it exceeds the masking threshold. We approached this problem from a new perspective: the proposed JEDF allows the ASP and quantization noises to be shaped jointly to match the masking threshold. In the case that the perceptual room between the masking threshold and the quantization noise is insufficient for the ASP noise, the JEDF can reduce the quantization noise level, which results in an increase in bit rate. Towards implementing this idea, we have developed two new techniques with promising experimental results: 1) SOPOT truncation noise shaping and 2) truncation noise allocation based on a perceptual model.

Priority-based burst packet transmission to save power

It is well-known from the literature that the Wireless Network Interface (WNI) is typically the most power consuming component in multimedia streaming to mobile devices. A standard approach for saving power is to force WNI to a sleep state when no network activity is expected. In a streaming application, however, this approach is problematic because data is typically received continuously. A possible solution is to transmit data packets in bursts, allowing WNI more time in between bursts to sleep, thus saving power. Unfortunately, bursty packet transmission is undesirable; it may compromise network performance, causing network congestion and packet loss. To solve this problem, we proposed an adaptive energy-saving streaming mechanism that adjusts burst length according to congestion conditions [10]. Experimental data showed that this approach provides a desirable tradeoff between power efficiency (average burst length and low dispersion of bursts) and congestion avoidance (packet loss and network problems observed by other users).

Our approach employed a layered multimedia coding scheme, in which a customized packet scheduler arranges bursty traffic such that priority packets with perceptually significant media data are transmitted first. When the high priority packets are delivered reliably, the lower priority packets can be easily skipped by the receiver without significant loss of perceived QoS. Fundamentally, this is an application of the perception-aware principle to wireless transmission.

Compressed domain beat detector

Music processing tasks, such as beat detection, can be implemented using low-complexity algorithms to handle the constraints of small mobile devices. To this end, we demonstrated a scheme of complexity scalable beat detection of popular (pop music) recordings that could run on battery-powered handheld devices [11].

The algorithm provides both theoretical and practical contributions, as we use the number of Huffman bits from the compressed bitstream, without requiring any decoding, as the sole feature for onset detection. Furthermore, we provide an efficient and robust graph-based beat induction algorithm. By applying the beat detector to compressed rather than uncompressed audio, the computational workload can be reduced by almost three orders of magnitude.

3. Conclusions

Depending on the application scenario, there are various strategies available to reduce the power consumption of battery-powered mobile devices. We have studied the entire signal processing chain, from the encoder, server, and wireless network to the decoder. As a result, we have proposed a number of approaches, which are summarized in this paper. Fundamentally, we seek desirable tradeoffs between processing workload (and hence power consumption) and human perceived quality of service by exploiting such parameters as inter-modality synergy, latency, and memory consumption. Until a revolutionary breakthrough in battery technology occurs, low-power media processing undeniably remains an important research problem.

References

[1] Wendong Huang, Ye Wang and Samarjit Chakraborty, "Power-Aware Bandwidth and Stereo-Image Scalable Audio Decoding," ACM Multimedia Conference, November 06-11, 2005, pp. 291 – 294, Hilton, Singapore, ACM.

[2] Wendong Huang and Ye Wang, "Efficient Partial Spectrum Reconstruction usinag an Asymmetric PQMF Algorithm for MPEG-Coded Stereo Audio," IEEE International Conference on Multimedia & Expo (ICME-06), July 9-12, 2006, pp. 901-904,Toronto, Canada.

[3] Guangming Hong, Ahmad Rahmati, Ye Wang and Lin Zhong, "SenseCoding: Accelerometer-assisted Motion Estimation for Efficient Video Encoding," ACM Multimedia Conference, October 27-31, 2008, pp. 749-752, Vancouver, Canada, ACM.

[4] Xiaoming Chen, Zhendong Zhao, Ahmad Rahmati, Ye Wang and Lin Zhong, "Sensor-Assisted Motion Estimation for Efficient H.264/AVC Video Encoding," ACM Multimedia 2009, pp. 381-390, Beijing, China, ACM.

[5] Yicheng Huang, Samarjit Chakraborty and Ye Wang, "Using Offline Bitstream Analysis for Power-Aware Video Decoding in Portable Devices ," ACM Multimedia Conference, November 06-11, 2005, pp. 299 – 302, Hilton,

Singapore, ACM.

[6] Yicheng Huang, Vu An Tran and Ye Wang, "A Workload Predication Model for Decoding MPEG Video and its Application to Workload-scalable Transcoding," ACM Multimedia Conference, September 23-29, 2007, pp. 952 – 961, Augsburg, Germany, ACM.

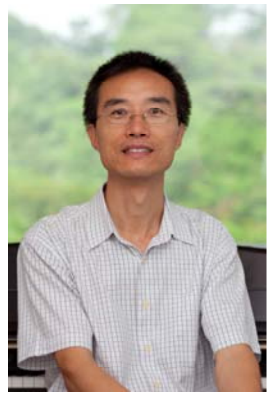
[7] Yicheng Huang, Guangming Hong, Vu An Tran and Ye Wang, "Decoding-workload-aware Video Encoding," ACM 2008 International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV-08), May 28-30, 2008, pp. 45-50, ACM

[8] Wendong Huang and Ye Wang, "An Optimal Speed Control Scheme Supported by Media Servers for Low-Power Multimedia Applications," Multimedia Systems, Volume 15, Number 2, pp. 113-124, April, 2009.

[9] Wendong Huang and Ye Wang, "A Joint Encoder-Decoder Framework for Supporting Energy Efficient Audio Decoding," Multimedia Systems, Volume 15, Number 2, pp. 101-112, April, 2009.

[10] Jari Korhonen and Ye Wang, "Power-Efficient Streaming for Mobile Terminals," ACM 2005 International Workshop on Network and Operating System Support for Digital Audio and Video (NOSSDAV-05), June 12-14, 2005, pp. 39 – 44, Oregon, USA, ACM

[11] Jia Zhu and Ye Wang, "Complexity-Scalable Beat Detection with MP3 Audio Bitstreams," Computer Music Journal (CMJ), pp. 71-87, January, 2008, MIT Press



Ye Wang worked at the Nokia Research Center in Tampere, Finland for 9 years as a research engineer and senior research engineer. Since 2002, he has been a faculty member in the Computer Science Department at the National University of Singapore (NUS). While a faculty member, he has established and directed the Sound and Music Computing (SMC) Lab in the NUS School of Computing. His main research projects include: 1) multimodal mobile music retrieval (M3R) and its applications to Healthcare, and 2) the NUS Mobile Music Group (NuMOG) for edutainment. In the recent past, he has worked on scalable and error robust audio streaming, perception-aware low-power media processing for portable devices. For more information concerning Dr. Wang's and the SMC lab's research activity, interested parties may visit his webpage:
http://www.comp.nus.edu.sg/~wangye