MULTIMEDIA COMMUNICATIONS TECHNICAL COMMITTEE
IEEE COMMUNICATIONS SOCIETY
http://committees.comsoc.org/mmc

R-LETTER

Vol. 3, No. 2, April 2012

IEEE COMMUNICATIONS SOCIETY

CONTENTS

Message from R-Letter Co-Director

Welcome to the second issue of the IEEE MMTC Review-Letter (R-Letter) in this year. The R-letter introduces promising new concepts and ideas in multimedia communication to MMTC members by rigorously selecting and reviewing high-impact and innovative papers from recent IEEE Communication Society, MMTC-sponsored publications, and other IEEE publications.

We are pleased to introduce seven high-quality papers in this issue. These papers span four main topics: video delivery, video coding, image processing, and communication in sensor networks.

The first paper, published in the *IEEE Transactions on Multimedia*, considers IPTV multicast in WiMAX networks. The second paper, from the *IEEE Global Telecommunications Conference*, studies dynamic adaptive streaming over HTTP from multiple sources in a cloud environment. The third paper, published in the *IEEE Transactions on Wireless Communications*, deals with streaming scalable videos over multi-hop cognitive radio networks. The fourth paper, from the *IEEE Global Telecommunications Conference*, addresses how a variable-bit-rate video service can be provided in cellular networks. The fifth paper, published in the *IEEE International Conference on Multimedia and Expo*, discusses how depth information can be exploited for efficient coding of multi-view videos. The sixth paper, from the *IEEE Transactions on Image Processing*, addresses the multi-exposure image fusion problem. The last paper, published in the *IEEE Journal of Selected Topics in Signal Processing*, utilizes a selective gossip strategy to tackle the approximation of large vectors of network data.

I would like thank all the editors of this issue for their great work: Carl James Debono, Christian Timmerer, Cheng-Hsin Hsu, Walid Saad, Vladan Velisavljević, Kai-Lung Hua, and Hassan Mansour. I also would like to thank the R-Letter Director Guan-Ming Su for all his great efforts.

Nabil J. Sarhan

Co-Director, IEEE ComSoc MMTC R-Letter
E-mail: nabil at ece.eng.wayne.edu

Towards Optimal IPTV Multicast over WiMAX Systems

A short review for "Optimal layered video IPTV multicast streaming over mobile WiMAX systems"

Edited by Carl James Debono

P.-H. Wu, and Y. H. Hu, "Optimal Layered Video IPTV Multicast Streaming over Mobile WiMAX Systems," IEEE Transactions on Multimedia, vol. 13, no. 6, pp. 1395 – 1403, Dec. 2011.

Multicast transmission offers a solution to provide the same content to multiple users with reduced resource requirements, compared to unicast transmission to each user. This requires the users to connect to one output port of the server compared to a dedicated port for each user in simulcast. Therefore, such a solution avoids saturating the bandwidth of band-limited channels and allows for better scalability of the network.

Wireless Multicast/Broadcast Service (MBS) have been included in the IEEE 802.16e (WiMAX) standard [1]. This allows the possibility of transmitting Internet Protocol Television (IPTV) multicast services over mobile WiMAX channels [2] – [4]. Using the network to provide such a service allows groups of mobile subscribers to receive the same multi-media content in real-time while roaming in metropolitan areas. A number of techniques in the IEEE 802.16e standard have been incorporated in the Physical (PHY) and Media Access Control (MAC) layers to mitigate the impact of the wireless channel impairments on the Quality of Services (QoS). Among these, Orthogonal Frequency Division Multiple Access (OFDMA) plays an important role in alleviating the varying quality of the channel.

Coupled with this, Scalable Video Coding (SVC), which encodes individual video streams in layered video sub-streams, can be used at the application layer to enhance the service. This allows the service provider to offer differential services to different users of the same video stream based on their respective wireless channel conditions, and levels of subscriptions.

The authors introduce WiMAX as a platform for the deployment of the IPTV service. To ensure a high quality Multicast/Broadcast service, the current IEEE 802.16e-2009 and the proposed 802.16m are designated a space time coding zone within each downlink OFDMA frame, known as MBS OFDMA zone [5]. Each MBS zone consists of several Orthogonal Frequency-Division Multiplexing (OFDM) symbols in the successive time intervals and several sub-channels in the frequency domain. All resources within this zone are dedicated to multicast service with no unicast traffic considered.

Due to the limited bandwidth available for radio transmission, multiple TV channels have to compete for radio resources. The goal of the service providers is to maximize the total revenue received for providing IPTV services. On the other hand, subscribers want to get the best deal with highest QoS at the lowest cost. To make the situation worse, different mobile users experience different channel conditions. To receive the same QoS, users with lower channel quality consume more radio resources. Thus, service providers have less incentive to serve high quality video to users with low channel conditions. However, when a video is encoded using SVC, it can be transmitted via different layers. Therefore, users with better channel conditions are more likely to receive more video layers and hence better QoS. Users experiencing worse channel conditions receive fewer video layers and hence inferior QoS.

In their work, the authors formulate the Wireless Radio Resource Allocation (WRRA) problem as a multi-objective optimization problem with a utility function that includes subscription statistics, channel quality data, assigned Modulation and Coding Scheme (MCS) rates as well as pricing of the individual video layers. They further demonstrate that various QoS metrics such as received rate, perceptual quality, or fairness can be easily implemented with an appropriate pricing structure. As such, this utility function is general and flexible.

A new approach was also proposed to solve the multi-objective WRRA optimization problem. Specifically, for a single video program, by imposing the dependence relations among successive layers of a scalable encoded video; it was shown that the total number of different service profiles, and the set of MCS rates

assigned to each layer of a given video stream, can be reduced by orders of magnitudes. Thus, it becomes practical to evaluate the expected utility value for these feasible service profiles. Furthermore, it was observed that only an essential service profile that maximizes the utility for a given amount of radio resources must be considered for the globally optimal solution of the WRRA problem. These profiles and corresponding utility values form a Pareto frontier curve in the utility-resource space of each individual video program. Therefore, to seek a globally optimal solution, one only needs to examine the much smaller set of essential service profiles of individual video streams.

The globally optimal solution of the WRRA problem with multiple video programs will result in one essential service profile from each video stream. Hence, an exhaustive enumeration approach for the globally optimal solution of multi-video WRRA problem is not practical. However, the overall utility of servicing multiple IPTV programs can be defined as a convex function of the utilities of individual video programs. Thus, an approximate solution can be obtained by (a) approximating the Pareto frontier curve of an individual video stream with a quadratic function, and (b) deriving an explicit, closed-form solution to a corresponding constrained convex optimization problem. This yields a good initial estimate of the globally optimal solution. Further refinement of this estimate is then done using a nearest neighbor incremental gradient search which converges within a few of iterations. The presented complexity analysis shows that this approach can be applied to practical real-time system.

The use of multicasting together with SVC can be a solution to allow IPTV over wireless broadband networks. Yet, scalable video coding requires extra bandwidth and more resources at both the server and the client side, compared to standard video coding. This increases power consumption at both the transmitter and receiver. Solutions that reduce the power consumption need to be studied, such as applying multicast trees on top of peer-to-peer networks, where the trees are updated according to the channel quality. Cooperative networks can also provide a solution, where relay nodes can be deployed to enhance the channel conditions for the users. Furthermore, high efficiency video coding is expected to reduce bandwidth requirements by half, leaving space for better error resilience and correction schemes in video applications such as IPTV.

References:

[1] *IEEE Standard for Local and Metropolitan Area Networks Part 16: Air Interface for Broadband Wireless Access Systems*, IEEE Std 802.16-2009.

[2] J. She, F. Hou, P.-H. Ho, and L.-L. Xie, "IPTV over WiMAX: Key success factors, challenges, and solutions," *IEEE Communications Magazine*, vol. 45, no. 8, pp. 87-93, Aug. 2007.

[3] W.-H. Kuo, T. Liu, and W. Liao, "Utility-based resource allocation for layer-encoded IPTV multicast in IEEE 802.16 (WiMAX) wireless networks," in *Proc. of the IEEE Int. Conf. on Communications*, Jun. 2007, pp. 1754 –1759.

[4] S. Deb, S. Jaiswal, and K. Nagaraj, "Real-time video multicast in WiMAX networks," in *Proc. of the 27th Conf. on Computer Communications*, Apr. 2008, pp. 1579–1587.

[5] WiMAX Forum, WiMAX System Evaluation Methodology V2.1, Jul. 2008.

Carl James Debono (S'97, M'01, SM'07) received his B.Eng. (Hons.) degree in Electrical Engineering from the University of Malta, Malta, in 1997 and the Ph.D. degree in Electronics and Computer Engineering from the University of Pavia, Italy, in 2000.

Between 1997 and 2001 he was employed as a Research Engineer in the area of Integrated Circuit Design with the Department of Microelectronics at the University of Malta. In 2000 he was also engaged as a Research Associate with Texas A&M University, Texas, USA. In 2001 he was appointed Lecturer with the Department of Communications and Computer Engineering at the University of Malta and is now a Senior Lecturer. He is currently the Deputy Dean of the Faculty of ICT at the University of Malta.

Dr Debono is a senior member of the IEEE and served as chair of the IEEE Malta Section between 2007 and 2010. He is the IEEE Region 8 Conference Coordination sub-committee chair for 2012. He has served on various technical program committees of

international conferences and as a reviewer in journals and conferences. His research interests are in wireless systems design and applications, multi-view video coding, resilient multimedia transmission and modeling of communication systems

Dynamic HTTP Streaming: Enabling Scalable Video Delivery within Cloud

A short review for "Dynamic adaptive streaming over HTTP from multiple content distribution servers"

Edited by Christian Timmerer

W. Pu, Z. Zou, and C. W. Chen, "Dynamic Adaptive Streaming over HTTP from Multiple Content Distribution Servers", in Proc. of IEEE Global Telecommunications Conference (GLOBECOM), pp. 1 – 5, Dec. 2011.

Nowadays, both users' video playback terminals and their networking access schemes are highly heterogeneous. A user may view videos with high-end desktop terminal supporting 1080p or even higher resolution with tens of Mbps broadband cable Internet access, Quad-core CPU, and high-end GPU. On the other hand, another user may have to view the same videos at a smartphone with only VGA resolution, 500 MHz single core CPU, and several hundred Kbps 3G wireless network access. Recently, Dynamic Adaptive Streaming over HTTP (DASH) was proposed in order to support such heterogeneous video streaming environments. In particular, at the server several versions of the same video program are encoded at different bit rates while the client promises to automatically adapt to bandwidth changes to provide smooth playback with best achievable visual quality.

DASH has been adopted by numerous popular commercial video streaming systems such as Microsoft smooth streaming, Apple HTTP live streaming, Adobe HTTP dynamic streaming, and Akamai HD. Recently, international standard organizations including 3GPP and MPEG are working on standardizing DASH [1][2]. Several unique features of DASH differentiate itself from conventional UDP/RTP-based video streaming. First, it reuses existing Web server rather than deploying dedicated streaming servers. Second, it takes advantage of the benefits in router and firewall optimizations for HTTP traffic.

In this paper, the authors aimed at developing a new architecture for DASH within cloud environments. Instead of streaming from one unique DASH server they are investigating the scenario that the DASH video sources are located in the cloud which is referred to as CloudDASH. That is, the DASH client can simultaneously request video data from multiple content distribution servers (CDN) located within the cloud.

There are several significant challenges in developing CloudDASH. First, in order to relieve CDN server load, reduce HTTP request traffic, and increase TCP throughput, HTTP requests should be as few as possible. However, using too few HTTP requests may result in inefficient tracking of network dynamics which in turn causes playback jitter or large visual quality fluctuation. Second, in the multi-server scenario, mapping video segments and bit rates to different servers is a challenging task. If a large segment is inappropriately assigned to a slow server, the segment is likely to miss its playback deadline. In this case, the bandwidth in transmitting such segment is wasted.

In the new scheme developed in this paper, the authors proposed a collaborative multi-scale scheduling algorithm (CMSS) based on Scalable Video Coding (SVC) [3] to effectively (1) mitigate Web server load and automatically balance loads among servers, (2) adapt to bandwidth dynamics of each server, and (3) optimize aggregated streaming quality. Therefore, SVC bit stream is assumed to be stored segment by segment at the servers. Each segment contains 30 seconds of video and is allocated 30s of transmission time. Packets missed their deadline are dropped directly. As SVC stream contains intrinsic rate adaptation functionality (i.e., CGS, MGS) which allows for relatively long segment lengths. Furthermore, the SVC network abstraction layer units (NALUs) within one segment are rearranged in rate-distortion optimized order. Hence, important NALUs (i.e., with lower priority id) are located in front of less important NALUs within the segment. The actual mapping of NALUs to priority id is not standardized and Amonou et al. [4] presented one practical solution.

With video packet reordered, there is need to map different portions of a segment to different CDN servers. However, it is well known that the TCP throughput is very difficult to predict. In CMSS the authors developed a binary allocation algorithm. At the beginning of each round of video-server mapping only one half of the time budget is allocated. The requested video bits are

based on the previous round of TCP throughput estimation. For example, in the first round each server is allocated to the amount of data estimated to arrive at the user within 15s. In the second round, one half of the available budget is allocated which in the ideal case is 7.5s using the average TCP rate in the first round as the predicted TCP rate. Such TCP rate prediction algorithm can be seen as a two-tap low pass filter.

A prototype has been implemented and deployed on PlanetLab and experimental results verify the effectiveness of the proposed CMSS. CMSS requires a logarithmic number of HTTP requests for each video segment. Comparing with the linear complexity fixed chunk size scheme, CMSS significantly reduced server load without sacrificing bandwidth utilization efficiency.

Other related approaches are presented in [5][6] but CloudDASH is a promising research direction for DASH. This paper provides an initial study including some simplifications. First, it is assumed that every CDN server has a full replica of the SVC coded video program. In practice, sometimes this assumption does not hold. How to optimize DASH users' Quality of Experience (QoE) with this additional dimension of complexity is challenging. Furthermore, only the scenario that CDN servers are organized in a flat topology is considered. Practically, CDN server may be organized in hierarchical structure. Front-end cache servers have small storage space but with high bandwidth and low HTTP request processing cost. The back-end video server farm has infinite storage space, but high query and communication cost. How to optimally allocate SVC video streams between these two different layers to improve DASH response speed, reduce overall system cost, and enhance viewing experience is subject to future work. Another important research topic is DASH service pricing within cloud environments. For example, would a flat rate model be more attractive as the pay-as-you-go model? How to link DASH users' QoE, rather than Quality of Service (QoS), to DASHs' service price? Finally, the QoE in general of such DASH systems is currently not very well understood and is still in its infancy.

References:

[1] I. Sodagar, "The MPEG-DASH Standard for Multimedia Streaming Over the Internet," *IEEE Multimedia*, vol. 18, no. 4, pp. 62-67, Oct-Dec 2011.

[2] T. Stockhammer, "Dynamic adaptive streaming over HTTP: Standards and design principles," in *Proc. ACM MMSys*, San Jose, CA, USA, 2011, pp. 133-144.

[3] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264/AVC Standard," *IEEE Trans. Circ. & Sys. Video Tech.*, vol. 17, no. 9, pp. 1103-1120, 2007.

[4] I. Amonou, N. Cammas, S. Kervadec, and S. Pateux, "Optimized rate-distortion extraction with quality layers in the scalable extension of H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1186–1193, Sep. 2007.

[5] R. Kuschnig, I. Kofler, and H. Hellwagner, "Evaluation of HTTP-based request-response streams for internet video streaming," in *Proc. ACM MMSys*'11, San Jose, CA, USA, Feb. 2011, pp. 245–256.

[6] T. Schierl, Y. S. de la Fuente, R. Globisch, C. Hellge, and T. Wiegand, "Priority-based media delivery using SVC with RTP and HTTP streaming," *J. Multimed. Tools Appl.*, Sep. 2010..

Christian Timmerer is an assistant professor in the Institute of Information Technology (ITEC), Alpen-Adria-Universität Klagenfurt, Austria. His research interests include immersive multimedia communication, streaming, adaptation, and Quality of Experience. He was the general chair of WIAMIS'08, ISWM'09, EUMOB'09, AVSTP2P'10, WoMAN'11 and has participated in several EC-funded projects, notably DANAE, ENTHRONE, P2P-Next, ALICANTE, and SocialSensor. He also participated in ISO/MPEG work for several years, notably in the area of MPEG-21, MPEG-M, MPEG-V, and DASH/MMT. He received his PhD in 2006 from the Alpen-Adria-Universität Klagenfurt. Publications and MPEG contributions can be found under research.timmerer.com, follow him on twitter.com/timse7, and subscribe to his blog blog.timmerer.com

Efficiently Streaming High-Quality Scalable Videos over the White Spaces

A short review for "Streaming scalable videos over multi-hop cognitive radio networks"

Edited by Cheng-Hsin Hsu

D. Hu and S. Mao, "Streaming Scalable Videos over Multi-hop Cognitive Radio Networks", IEEE Transactions on Wireless Communications, vol.9, no.11, pp. 3501 – 3511, Nov. 2010.

Traditionally, wireless spectrums are exclusively licensed to primary users, which can be TV broadcast companies and cellular network operators. Despite that many primary users do not utilize all the licensed spectrums [1], these residue spectrums, often referred to as the white spaces, are simply wasted. In 2008, Federal Communications Commission (FCC) approved the unlicensed use of white spaces by secondary users, which has in turn stimulated the research on cognitive radio.

Cognitive radio [2, 3] is a frequency-agile radio that can sense and adapt to the radio environment. Cognitive radio allows the secondary users to use the underutilized licensed spectrums in order to increase the overall network capacity without imposing significant negative impacts on the primary users. Although the existing cognitive radio studies resolve many issues in spectrum sensing and access techniques, rather few works consider application-level quality of multimedia services in cognitive radio networks.

The authors of this paper consider the problem of streaming videos over multi-hop cognitive radio networks. Scalable videos, such as MPEG-4 Fine Grained Scalable (FGS) and H.264/SVC Medium Grained Scalable (MGS), are employed for graceful quality changes as the spectrum opportunities come and go. The authors consider an infrastructure-less cognitive radio network spanning over one or more fixed primary networks. The objective is to maximize the overall video quality while achieving proportional fairness among video users, with a controlled interference to the primary user. This problem is fairly challenging because many details must be taken into considerations, which include spectrum sensing and sensing errors, spectrum access and primary user protection, video quality and fairness, and channel selections for simultaneous videos.

In particular, the authors formulate the problem of scalable video streaming over a multi-hop cognitive radio network as a Mixed Integer Nonlinear Programming (MINLP) problem. This optimization problem spans across multiple layers, and thus is a cross-layer optimization problem. The authors first relax the MINLP problem into a Nonlinear Programming (NLP) problem. Solving this NLP problem gives an (potentially infeasible) upper bound of the original MINLP problem. The authors also propose a heuristic algorithm, called Sequential Fixing (SF), which iteratively finds the best value of each decision variable, one after another. The SF algorithm gives a lower bound of the MINLP problem. The two aforementioned algorithms are centralized algorithms for the MINLP problem, which were first presented in the authors' earlier work [4].

The centralized algorithms are less applicable for real cognitive radio video systems, and the authors develop distributed algorithms in this paper. The MINLP problem is decomposed into a channel scheduling problem and a path selection problem. The channel scheduling problem is solved using a greedy algorithm, which is provably optimal. The path selection problem is addressed by dual decomposition, which leads to a distributed algorithm based on Lagrange Multipliers and subgradient method. This distributed algorithm is also proved to converge to the optimal solution as long as the step size of the subgradient method is small enough.

The authors evaluate their algorithms using a custom simulator built on C and Matlab. The simulation results indicate that their algorithm converges quickly, only takes 76 iterations in the considered simulation scenario, while imposing negligible overhead, less than 300 broadcast messages. The simulation results also show that the distributed algorithm achieves very similar performance compared to the centralized one. This reveals the efficiency of the distributed algorithm. Finally, the authors demonstrate that the proposed algorithms work for both MPEG-4 FGS streams and H.264/SVC MGS streams.

In summary, this paper jointly considers spectrum sensing in the physical layer and video quality in the application layer in a scalable video streaming system over cognitive radio networks. The proposed algorithm runs in a distributed manner, yet quickly converges to the global optimal solution. The paper clearly shows that cognitive radio network can efficiently transport real-time scalable videos. The paper also points out several practical considerations and includes insightful discussions. I am confident that this work can stimulate many more research on video streaming systems over cognitive radios, both in the theory and systems sides. Hopefully, we will see such systems being implemented and deployed soon.

References:

[1] M. Calabrese, "Broadcast to Broadband: Unlicensed Access to Unused TV Channels?" *IEEE Internet Computing*, vol. 12, no. 2, pp. 71–75, Mar.-Apr., 2008.

[2] Q. Zhao and B. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79-89, May 2007.

[3] I. Akyildiz, W. Lee, M. Vuran, and S. Mohanty, "NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey," *Computer Networks Journal*, vol. 50, no. 9, pp. 2127-2159, Sep. 2006.

[4] D. Hu, S. Mao, and J. Reed, "On Video Multicast in Cognitive Radio Networks" In *Proc. of IEEE INFOCOM'09*, Rio de Janeiro, Brazil, Apr. 2009, pp. 2222-2230.

Cheng-Hsin Hsu received the Ph.D. degree from Simon Fraser University, Canada in 2009, the M.Eng. degree from University of Maryland, College Park in 2003, and the M.Sc. and B.Sc. degrees from National Chung-Cheng University, Taiwan in 2000 and 1996, respectively. He is an Assistant Professor in Department of Computer Science at National Tsing Hua University, Taiwan. He was a Senior Research Scientist at Deutsche Telekom R&D Lab USA, Los Altos, CA between 2009 and 2011. His research interests are in the area of multimedia networking and distributed systems. He has published papers in leading journals, conferences, and workshops. He and his colleagues at Simon Fraser University developed a mobile TV testbed, which won the Best Technical Demo Award in the ACM Multimedia 2008 Conference. He is on the Review Board Committee of IEEE Technical Committee on Multimedia Communications (MMTC) and the Preservation Committee of ACM Special Interest Group on Multimedia (SIGMM). He served as the TPC Co-chair of the ACM Mobile Video Workshop (MoVid'12) and the Proceeding and Web Chair of the ACM International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV'10). He was on the technical program committees of several well-known conferences in his research areas, including ACM Multimedia Conference (Multimedia), ACM International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV), and IEEE International Conference on Multimedia and Expo (ICME).

Towards Variable-Bit-Rate Video Service in Cellular Networks

A short review for "Downlink power control for VBR video streaming in cellular networks: a majorization approach"

Edited by Walid Saad

Y. Huang, S. Mao, and Y. Li, "Downlink power control for VBR video streaming in cellular networks: A majorization approach", in Proc. of IEEE Global Telecommunications Conference (GLOBECOM), pp. 1 – 6, Dec. 2011.

With the ongoing "smartphone revolution," multimedia and video data is becoming a dominant part of the fast growing wireless data. According to a recent study by Cisco, global mobile data traffic will increase 18-fold between 2011 and 2016, with a compound annual growth rate of 78%. Moreover, it is predicted that mobile video data will increase 25-fold from 2011 to 2016, and over 70% of the total mobile data traffic will be video related by the end of period [1]. Such drastic increase in wireless video will significantly stress the capacity of existing wireless networks. Under this context, it is important to revisit the design and operation of existing cellular networks, and to maximize their capacity in carrying real-time video data.

Among different video coding techniques, it has been recognized that variable-bit-rate (VBR) videos can offer constant and better quality over constant-bit-rate (CBR) videos with the same bit budget. In addition, since many stored video are coded in the VBR format, it would be desirable to support VBR video streaming directly, without resorting to transcoding. However, VBR videos also pose great challenges to scheduling, resource allocation, and control, in particular, in wireless networks [2, 3, 4]. This is largely due to the high rate variability and complex autocorrelation structure typically found in VBR video traffic [5, 6].

In this respect, this paper investigates the challenging problem of optimal power allocation for streaming multiuser VBR videos in the downlink of a cellular network. Particularly, the downlink of a cellular network with orthogonal channels, where the base station (BS) streams multiple VBR videos to mobile users in the cell is studied. The authors consider playout buffer underflows and overflows as video performance measure, as well as a deterministic model for VBR video traffic that incorporates video frame and playout buffer characteristics [6]. By assigning transmit power to each video session, a feasible transmission curve can be derived that can minimize both playout buffer overflow and underflow events. A constrained stochastic optimization problem is formulated to minimize the total power consumption at the BS.

The results show that the formulated problem fits well with majorization theory, that deals with partial ordering of real vectors and order-preserving functions [7]. The authors then develop a majorization-based solution for the formulated stochastic programming problem. After proving that the objective function is Schur-convex for the case of relaxed BS peak power constraint, a majorization-based power optimal algorithm with low computational complexity is developed. Furthermore, the authors prove that the power optimal algorithm produces the transmission curve for the VBR video with the smallest rate variations [6]. For the case of multiuser VBR video streaming where the power allocations are coupled by the base station's peak power constraint, a heuristic algorithm is developed to selectively suspend some video sessions, which are least likely to experience playout buffer underflow in the next frame interval. The proposed algorithms are evaluated with trace-driven simulations [8], and are shown to achieve considerable power savings over a conventional "lazy" scheme, which only transmits the minimum amount of video data that is needed in the next frame interval [9].

This paper provides an interesting example of how a classic theory from other fields can be applied to solve a relevant and challenging wireless VBR video streaming problem. The treatment of the problem is thorough and rigorous with excellent technical depth. The finding, that the majorization-based approach outperforms the conventional lazy scheme, is somewhat counter-intuitive, since the lazy scheme seems to be energy efficient by only transmitting the minimal amount of video data in each time slot. This work has the potential of making considerable impact on both theory and

algorithms for cross-layer design and optimization for video communications.

Several future extensions to this work can be envisioned. On one hand, it is natural to apply the majorization based approach to the streaming of multiuser CBR videos. It is also of interest to investigate the potential gain that may result from the reduced rate variability. On the other hand, it would be interesting to study how to extend this approach to the case of live video streaming, where VBR videos are coded online with limited look-ahead information. Finally, an analytical model would be useful to quantify the playout buffer underflow and/or overflow probabilities, which can be incorporated into an admission control scheme for quality of experience (QoE) provisioning for streaming videos in cellular networks.

References:

[1] Cisco Systems, Inc., "Cisco visual networking index: Global mobile data traffic forecast update, 2011-2016," Feb. 2012, [online] Available: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html.

[2] Y. Huang and S. Mao, "Downlink power control for variable bit rate video over multicell wireless networks," in *Proc. IEEE INFOCOM'11*, Shanghai, China, Apr. 2011, pp. 2561—2569.

[3] T. Stockhammer, H. Jenkac, and G. Kuhn, "Streaming video over variable bit-rate wireless channels," *IEEE Transactions on Multimedia*, vol. 6, no. 2, pp. 268—277, Apr. 2004.

[4] G. Liang and B. Liang, "Balancing interruption frequency and buffering penalties in VBR video streaming," in *Proc. IEEE INFOCOM'07*, Anchorage, AK, May 2007, pp. 1406—1414.

[5] M. W. Garrett and W. Willinger, "Analysis modeling and generation of self-similar VBR video traffic," *ACM SIGCOMM Computer Communication Review*, vol.24, no.4, 269—280, Oct. 1994.

[6] J. Salehi, Z.-L. Zhang, J. Kurose, and D. Towsley, "Supporting stored video: Reducing rate variability and end-to-end resource requirements through optimal smoothing," *IEEE/ACM Transactions on Networking*, vol.6, no.4, pp. 397—410, Aug. 1998.

[7] B. C. Arnold, *Majorization and the Lorenz Order: A Brief Introduction.* New York, NY: Springer-Verlag, 1987.

[8] M. Reisslein, *Video Trace Library*, Arizona State University, Tempe, AZ, USA. [online] Available: http://trace.eas.asu.edu.

[9] S. Sen, D. Towsley, Z. Zhang, and J. K. Dey, "Optimal multicast smoothing of streaming video over the Internet," *IEEE Journal on Selected Areas in Communications*, vol.20, no.7, pp.1345—1359, Sept. 2002.

Walid Saad received his B.E. degree in Computer and Communications Engineering from the Lebanese University, Faculty of Engineering II, in 2004, his M.E. in Computer and Communications Engineering from the American University of Beirut (AUB) in 2007, and his Ph.D from the University of Oslo in 2010. From August 2008 till July 2009 he was a visiting scholar in the Coordinated Science Laboratory at the University of Illinois at Urbana Champaign where he pursued research activities during his 2nd year of the PhD. From January 2011 till August 2011, he was a postdoctoral research associate at the Electrical Engineering Department at Princeton University.

Currently, he is an Assistant Professor at the Electrical and Computer Engineering Department at the University of Miami. His research interests include applications of game theory in wireless networks, cognitive radio and femtocell networks, physical layer security, and wireless communication systems (UMTS, WiMAX, LTE, etc.). He was the first author of the paper that received the Best Paper Award at the 7th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), in June 2009 and the paper that received a Best Paper Award at the 5th International Conference on Internet Monitoring and Protection (ICIMP 2010) in May 2010.

Multi-View Video Coding Coupled with Depth Information

A short review for "Exploiting depth information for efficient multi-view video coding"

Edited by Vladan Velisavljević

B. W. Micallef, C. J. Debono, and R. A. Farrugia, "Exploiting Depth Information for Efficient Multi-view Video Coding", in Proc. of the IEEE International Conference on Multimedia and Expo (ICME), July 2011.

With recent advancements in multimedia communications, 3D [1] and free-viewpoint [2] television services are closer to becoming a reality for everyday use. These services provide the depth perception and the free-viewpoint navigation of the presented media, thus, they will provide a better immersive experience which is closer to the natural one. An efficient 3D video representation for these services is the multi-view video-plus-depth format [3]. This represents the captured scene from multiple viewing directions, together with their aligned geometrical depth data. Thus, by using advanced 3D rendering techniques, such as the depth-image-based-rendering [4], it allows viewpoint rendering from any arbitrary direction, which is essential for such services.

To compress the huge amount of data required for these 3D videos, the multi-view video coding (MVC) standard [5] was developed. This gives high compression efficiency by using both the motion and the disparity compensation techniques to compensate the currently being encoded macro-block (MB) from the corresponding already encoded temporal and viewpoint reference frames, respectively. Thus, it is used for efficient encoding of both the color [5] and the depth [6] multi-view videos of the 3D videos.

To improve the compression efficiency of the MVC, the authors of this paper exploit the multi-view geometry together with the geometrical depth information, to provide geometric motion and disparity vector predictors for the color MVC. These predictors are expected to be more accurate to transmit the optimal compensation vectors than the conventional median ones. Thus, they provide a reduction in the encoded data and indicate a better motion and disparity estimations' search areas. This results in more efficient compensation vectors.

Furthermore, the authors described how the SKIP mode of the MVC standard can be extended to predict also the disparity vectors from the neighborhood MBs, to obtain a hybrid temporal/inter-view referenced SKIP mode. Usually, the neighborhood vectors are highly correlated as they are most likely representing the compensation of the same objects, thus, their compensation directions are also correlated. Therefore, the authors found that the SKIP modes' reference frame can be determined using the neighborhood encoded vectors, by using a majority voting scheme.

In the experimental results section, the authors investigate the distribution of the encoded residual vectors of the 16×16 modes, when the original or the proposed predictors are used. They demonstrated that smaller residual vectors actually result with the proposed geometric predictors and that this provides a reduction of about 6% in the encoded bit-rates. Then, when the new temporal/inter-view SKIP mode is applied on top of this to further encode the multi-view videos, it gives a bit-rate reduction of about 15%.

An interesting conclusion derived from these results is that the geometric predictors are in fact more accurate to represent the compensation of the current MB. Moreover, some MBs in the dynamic regions, which are usually disparity compensated, could now be encoded more efficiently as SKIP modes, using the proposed new SKIP mode and the neighborhood encoded disparity vectors. Both of these contribute to highly reduced bit-rates of the encoded color multi-view videos.

The authors investigate these efficient techniques for the low-latency coding structure, to improve the MVC efficiency for real-time applications. For the success of MVC in such applications, it is crucial to increase the coding efficiency, since they are refrained from using efficient encoding techniques that could increase their coding latency, such as the bi-prediction estimation. As the proposed technique does not affect the coding latency, and the depth data is available to support the virtual viewpoint generation, it is

wise to exploit this data to increase the coding efficiency of the real-time MVC.

Potential future work may include the study of how these efficient techniques will behave when the same geometric predictors are used with a reduced search area, for fast MVC, as proposed by the same authors in [7]. This will sacrifice a part of the coding efficiency but it will decrease the encoding computations of the low-latency coding structures, thus making real-time encoding less computational intensive and more adequate for mobile devices.

References:

[1] L. Onural, "Television in 3-D: What are the prospects?," *Proceedings of IEEE*, vol. 95, no. 6, Jun. 2007.

[2] M. Tanimoto, "Overview of free viewpoint television," *Signal Processing: Image communications* 21, 2006.

[3] MPEG & VCEG, "Multi-view video plus depth (MVD) format for advanced 3D video systems," JVT-W100, Apr. 2007.

[4] P. Kauff, N. Atzpadin, C. Fehn, M. Müller, O. Schreer, A. Smolic, and R. Tanger, "Depth map creation and image based rendering for advanced 3DTV services providing interoperability and scalability," *Signal Processing: Image Comm.* Special Issue on 3DTV, Feb. 2007.

[5] P. Merkle, A. Smolic, K. Müller, and T. Wiegand, "Efficient prediction structures for multi-view video coding," *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 17, no. 11, Nov. 2007.

[6] P. Merkle, A. Smolic, K. Müller, and T. Wiegand, "Efficient compression of multi-view depth data based on MVC," in *Proc. of Int. Conference on Image Processing*, Sept. 2007.

[7] B. W. Micallef, C. J. Debono, and R. A. Farrugia, "Exploiting depth information for fast motion and disparity estimation in multi-view video coding," in *Proc. of 3DTV-Conference*, May 2011.

Vladan Velisavljević received the B.Sc. and M.Sc. (Magister) degree from the University of Belgrade, Serbia, in 1998 and 2000, respectively, and the Master and Ph.D. degree from EPFL, Lausanne, Switzerland, in 2001 and 2005.

From 1999 to 2000, he was a member of academic staff at the University of Belgrade. In 2000, he joined the Audiovisual Communications Laboratory (LCAV) at EPFL as teaching and research assistant, where he was working on his Ph.D. degree in the field of image processing. In 2003, he was a visiting student at Imperial College London. From 2006 to 2011, Dr. Velisavljević was a Senior Research Scientist at Deutsche Telekom Laboratories, Berlin, Germany. Since October 2011, he is Senior Lecturer (Associate Professor) at Bedfordshire University, Luton, UK.

He has co-authored more than 40 research papers published in peer-reviewed journals and conference proceedings and he has been awarded or filed 4 patents in the area of image and video processing. He co-organized a special session at IEEE ICIP-2011 on compression of high-dimensional media data for interactive navigation and he is a co-chair of the Multimedia Computing and Communications Symposium (MCC) at IEEE ICNC-2013. His research interests include image, video and multiview video compression and processing, wavelet theory, multi-resolution signal processing and distributed image/video processing.

Towards Scene Referred Experience: Generalized Random Walks Approach

A short review for "Generalized random walks for fusion of multi-exposure images"

Edited by Kai-Lung Hua

R. Shen, I. Cheng, J. Shi, and A. Basu, "Generalized Random Walks for Fusion of Multi-Exposure Images", IEEE Transactions on Image Processing, vol. 20, no.12, pp. 3634 – 3646, Dec. 2011.

In many scenarios, taking a single digital photo is insufficient to clearly record all the details in the scene. For instance, some areas in the photo may appear so bright that details are washed out (over-exposed) and other portions may appear so dark that details can hardly be seen (under-exposed). This is because a natural scene usually contains a wide range of intensity levels that is beyond what a common digital camera is able to capture and also beyond the display capability of a common digital screen. This contradiction between the high dynamic range (HDR) nature of a real-world scene and the low dynamic range (LDR) limitation of current capture and display devices motivates the development of fusion techniques for multi-exposure images. The basic idea is to take a series of photos (LDR images) of the same scene under different exposure settings and then select the well-exposed pixels from each photo to create a single image according to certain quality measures.

An alternative procedure to multi-exposure fusion is the combined process of HDR reconstruction and tone mapping [1], in which an HDR image is recovered from input LDR images and then the dynamic range of this HDR image is compressed (tone-mapped) to produce an output LDR image. This process allows more flexibility in terms of user interaction. However, fusion techniques are normally preferred for quickly generating a well-exposed image from the input image set, especially when the number of input images is small and speed is crucial. Previous multi-exposure fusion methods [2][3] usually define the fusion weights locally without adequate consideration of consistency in a large neighborhood, which may cause unnatural appearance in the fusion results. Some methods partition the input images into different regions, either using uniform blocks or by segmentation techniques, and then try to maximize a certain quality measure within each region. These methods tend to cause artifacts at object/region boundaries, because inter-region information is not effectively exploited. Multi-resolution fusion methods normally work better at region boundaries and are good at enhancing main image features by blending fusion weights at different scales. However, the weights are still mainly determined locally without considering large neighborhood information. This may cause some inconsistencies in the results.

This paper introduces a fresh view of the multi-exposure image fusion problem. The authors propose a probabilistic method to achieve an optimal balance between two quality measures, i.e., local contrast and color consistency. The probabilities that a pixel in the fused image comes from different input images are estimated based on these two measures and then used as fusion weights in the final composition step. The local contrast measure is applied in order to select those input pixels containing more details. Because high local contrast is usually associated with high local variation, for a given input pixel, the local variation around it is computed and then modified using a sigmoid-shaped function to indicate local contrast. The color consistency measure imposes not only consistency in a large neighborhood but also consistency with the natural scene. This measure is based on the assumptions that adjacent pixels having similar colors in most input images will indicate similar colors in the fused image, and that similar colors at the same pixel location in different input images under proper exposures will indicate the true color of the scene. Therefore, for two given adjacent pixels, their similarity is evaluated based on their color differences. These two locally defined quality measures are integrated in their proposed generalized random walks (GRW) framework as compatibility functions to obtain optimal fusion weights.

In order to model the problem in the random walks context, the authors convert it to a labeling problem. The fused image is organized as an undirected graph, where each scene node corresponds to a pixel in the fused image. Within this graph, there is also a set of label nodes, each of which indicates an input image. The proportions of contributions from each input

pixel to their corresponding pixels in the fused image are considered as the probabilities of each scene node being assigned different labels. This probability estimation is further formulated as transition probability calculation in GRW, i.e., to calculate the probability that a random walker starting from a scene node first reaches a label node in the graph. This can be efficiently computed by solving a set of linear systems. Compared to previous random walks algorithms [4][5], GRW provides a more flexible framework that allows more flexible forms of the compatibility functions.

The effectiveness of the presented fusion algorithm was validated using various image sequences, and compared with state-of-the-art multi-exposure fusion methods and tone mapping methods. Experiments demonstrated that the algorithm generates fused images with fine details, good local contrasts, and minimal artifacts with high efficiency. The sensitivity of the free parameters used in the algorithm was also carefully analyzed. Based on the analysis, practical parameter ranges were suggested. The performance of this algorithm could be further improved by reducing its sensitivity to image noise with the development of more effective quality measures. Since GRW is developed as a general framework for solving different labeling problems, it is also worthwhile to explore the effectiveness of GRW on other applications, such as stereo matching [6].

References:

[1] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, W. Heidrich, and K. Myszkowski, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*, 2nd ed. Waltham, MA: Morgan Kaufmann, June 2010.

[2] A. Goshtasby, "Fusion of multi-exposure images," *Image Vision Comput*., vol. 23, no. 6, pp. 611–618, June 2005.

[3] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion," in *Proc. Pacific Graphics*, pp. 382–390, October 2007.

[4] L. Grady, "Multilabel random walker image segmentation using prior models," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, vol. 1, pp. 763–770, July 2005.

[5] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 28, no. 11, pp. 1768–1783, November 2006.

[6] R. Shen, I. Cheng, X. Li, and A. Basu, "Stereo matching using random walks," in *Proc. Int. Conf. Pattern Recognit.*, pp. 1-4, December 2008.

Kai-Ling Hua received the B.S. degree in electrical engineering from National Tsing Hua University, in 2000, and the M.S. degree in communication engineering from National Chiao Tung University in 2002, both in Hsinchu, Taiwan. He received the Ph.D. degree from the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, in 2010. Since 2009, Dr. Hua has been with National Taiwan University of Science and Technology, where he is currently an assistant professor in the Department of Computer Science and Information Engineering. He is a member of Eta Kappa Nu and Phi Tau Phi. His current research interests include digital image and video processing, computer vision, and multimedia networking.

Decentralized Approximation in Wireless Sensor Networks: Selective Gossip

A short review for "Efficient decentralized approximation via selective gossip"

Edited by Hassan Mansour

D. Ustebay, R. Castro, and M. Rabbat, "Efficient Decentralized Approximation via Selective Gossip", IEEE Journal of Selected Topics in Signal Processing, vol. 5, no. 4, pp. 805 – 816, Aug. 2011.

Recent progresses in wireless sensor networks and related applications have brought about the requirement for distributed signal processing algorithms. However, as the connectivity and number of nodes increase, we encounter the network scalability issue. The collection and processing of data at a fusion center in large networks is not perfect since it makes both bottlenecks in the network and a single point of malfunction. Furthermore, communication overhead of such centralized algorithms can be comparable or significantly higher than that of in-network signal processing algorithms. Under this scenario, we remain constrained by the underlying infrastructure lifetime. Effective algorithms to alleviate the aforementioned problems are needed.

One of the promising solutions is to deploy gossip algorithm, which is an iterative decentralized computation scheme. Each node in the network computes the approximation of the quantity and updates their approximations based on asynchronous information exchanged with one-hop neighbors. The updates used in the gossip algorithm asymptotically achieve consensus. As described, gossip algorithm requires neither routing nor coordination. It is clear that the risks discussed above for centralized system are significantly reduced.

Randomized gossip algorithms [1]-[6] intend to solve the average consensus problem asymptotically through local information exchanges between neighboring nodes. At every iteration of randomized gossip, one of the nodes wakes up uniformly at random and the waken node randomly chooses one of its neighbors to exchange values such that both nodes take as their new approximation the average of the values they exchanged. Note that the global average is preserved through iterations.

The main contribution of this paper is to describe selective gossip to specifically tackle the approximation of large vectors of network data instead of scalar values. Note that the significance of elements cannot be obtained before reaching the consensus through the iterations in the regular gossip algorithm. As proposed in this paper, selective gossip adaptively determines which elements are significant and which are insignificant while gossiping. At each round of gossip, two neighboring nodes exchange information for components of the vector that at least one of the nodes believes to be significant. Consequently, the network resources are preserved to compute components which contain significant energy. This paper proves that selective gossip converges asymptotically. In particular, all the nodes in the network reach consensus on the values of the significant components. On the other hand, for the insignificant components, all nodes in the network terminate computation with approximations which are below the threshold in absolute value, and all nodes obtain approximations which are below the threshold after a finite number of iterations. Therefore, all nodes reach a consensus on which components to disregard.

The paper then shows how selective gossip can be used for sparse approximation in a field estimation application. It turns out that selective gossip obtains a network-wide approximation with considerably fewer transmissions compared to naively gossiping in parallel on all coefficients.

Future work in this field will include investigating the rates of convergence for selective gossip. The rates of convergence can be associated with the study of voter models from interacting particle systems [7]. In the voter model each node chooses a random neighbor with some probability and adopts the state of this neighbor. Hence the significance of component values in selective gossip can be seen as the votes in the voter model. For finite graphs, the authors of [8] show that the convergence time of voter model process is related to the hitting times of the random walks on the graph. Further improvements to the selective gossip algorithm

could be achieved using the synchronous distributed averaging algorithm of [9]. This algorithm provides faster rates compared to randomized gossip at the cost of extra memory at nodes.

References:

[1] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508–2530, June 2006.

[2] A. Dimakis, A. Sarwate, and M. Wainwright, "Geographic gossip: Efficient aggregation for sensor networks," in *Proc. Information Processing in Sensor Networks*, Nashville, TN, Apr. 2006.

[3] F. Benezit, A. Dimakis, P. Thiran, and M. Vetterli, "Gossip along the way: Order-optimal consensus through randomized path averaging," in *Proc. Allerton Conference on Communication, Control, and Computing*, Urbana-Champaign, IL, Sep. 2007.

[4] T. Aysal, M. Yildiz, A. Sarwate, and A. Scaglione, "Broadcast gossip algorithms: Design and analysis for consensus," in *Proc. IEEE Conference on Decision and Control*, Cancun, Mexico, Dec. 2008.

[5] D. U□stebay, B.N.Oreshkin, M.J.Coates, and M.G.Rabbat,"Greedy gossip with eavesdropping," *IEEETransactionson Signal Processing*, vol. 58, no. 7, pp. 3765 – 3776, July 2010.

[6] E. Yildiz and A. Scaglione, "Differential nested lattice encoding for consensus problems," in *Proc. Information Processing in Sensor Networks*, Cambridge, MA, Apr. 2007.

[7] T. M. Liggett, *Interacting Particle Systems*. Springer-Verlag, 1985.

[8] D. Aldous and J. A. Fill, "Reversible markov chains and random walks on graphs," manuscript, 1994. Available: http://www.stat.berkeley.edu/ ~ aldous/RWG/book.html

[9] B. Oreshkin, M. Coates, and M. Rabbat, "Optimization and analysis of distributed averaging with short node memory," *IEEE Transactions on Signal Processing*, vol. 58, no. 5, pp. 2850 –2865, May 2010.

Hassan Mansour received his B.E. degree in Computer and Communications Engineering from the American University of Beirut in 2003, his M.A.Sc. and Ph.D. in Electrical and Computer Engineering from the University of British Columbia in 2005 and 2009, respectively. Since January 2010, he has been a postdoctoral research fellow in the Mathematics and Computer Science departments at the University of British Columbia.

His research interests include the theory and application of compressed sensing, sparse signal reconstruction algorithms utilizing prior information, seismic imaging, scalable video coding (SVC), stereo image processing, and high dynamic range (HDR) image and video compression.

Paper Nomination Policy

IEEE MMTC R-letter welcomes review paper nomination. Any paper published in an IEEE ComSoc journal/magazine or in the MMTC sponsored proceedings: IEEE GLOBECOM, ICC and ICME, in the two years preceding the next award board's election, is eligible.

The paper nomination is always open. Paper nominations have to be sent to the IEEE MMTC Review Board Director by email. The nomination should include the complete reference of the paper, author information, a brief supporting statement (maximum one page), the nominator information, and an electronic copy of the paper when possible. Only papers published in the two years preceding the nomination will be considered.

Each nominated paper will be reviewed by two members of the IEEE MMTC Review Board, according to the area of expertise, and avoiding any potential conflict of interest. The reviewer names will be kept confidential. If both members agree that the paper is of award quality, they will recommend publishing the review of the paper (partially based on the nomination supporting document) in the IEEE MMTC Review Letter.

For more details, please refer to http://committees.comsoc.org/mmc/awards.asp