R-LETTER

Vol. 4, No. 4, August 2013

IEEE COMMUNICATIONS SOCIETY

CONTENTS

Message from the Review Board

Introduction

Since the launch of R-Letter in October 2010, there have been fifteen publications. Credits should be given to all contributors. In order to deliver timely bi-monthly R-Letter, the Review Board needs to maintain a pool of nominated papers so that board members can have sufficient time to complete the review and editorial process. Therefore, we would like to invite the MMTC community to actively participate in the nomination process. Please refer Paper Nomination Policy at the end of this issue. Nominators of Review articles will be acknowledged in the respective R-Letter.

The Review Board aims at recommending recent (within one and half year) state-of-the-art and emerging publications in the literature. The topics should be of general interest for the MMTC community. In this issue, the two distinguished articles discuss multi-view video encoding and telepresence based on an array of Kinect sensor.

Distinguished Category

The growing need for applications and services which are ranging from stereoscopic telepresence system to multi-view encoded content is increasing rapidly. However, it is not sure how these services may be deployed with commodity hardware like the Kinect. For multi-view video encoding, the actual encoding delay is an important issue which calls for the analysis and optimization thereof.

The first paper, published in IEEE Journal of Selected Topics in Signal Processing, provides a framework for the analysis and optimization of the encoding latency for multi-view video. **The second paper**, published in Elsevier: Computers & Graphics, shows how to use commodity depth cameras in order to provide enhanced personal auto-stereoscopic telepresence.

Regular Category

While more and more social data are available, how to extract visual signals and images and analyze them have been extensively studied in the literature. Meanwhile, how to allocate resources for these services is another important area to study. In this issue, the regular category has assembled six papers on these topics. **The first paper**, published in the *IEEE Transactions on Signal Processing*, Paper proposes the construction of two-channel wavelet filter banks for analyzing graph-signals. The **second paper**, from *IEEE Transactions on Multimedia*, proposes a blind resource allocation scheme by taking into account the fairness among the users. The authors derive the convergence time of the proposed scheme and show that the proposed scheme provides almost the same MoS value as the optimum solution which knows QoE model in advance. The **third paper**, published in *IEEE Transactions on Multimedia*, the authors provide a generalized framework for optimizing the resources needed to support real-time IPTV services in a virtualized architecture, which takes advantage of the different deadlines associated with each service to effectively multiplex these services by time-shifting scheduling. The **fourth paper** is the best paper from *IEEE ICME'2013*, which proposes a novel scheme to automatically summarize and depict human movements from 2D videos without 3D motion capture or manually labeled data. The **fifth paper** is the best student paper award of *IEEE ICME'2013*. The authors propose a unified model to automatically identify visual concepts and estimate their visual characteristics, or visualness, from a large-scale image dataset. We would like to thank all the authors, reviewers, nominators, editors and others who contribute to the release of this issue. The **sixth paper,** published in the *IEEE Wireless Communications*, propose a cross-layer optimization framework for cooperative video summary transmission.

We would like to thank all the authors, nominators, reviewers, editors and others who contribute to the release of this issue.

IEEE COMSOC MMTC R-Letter

IEEE ComSoc MMTC R-Letter

Director:
Irene Cheng, University of Alberta, Canada
Email: locheng@ualberta.ca

Co-Director:
Weiyi Zhang, AT&T Research, USA
Email: maxzhang@research.att.com

Co-Director:
Christian Timmerer
Alpen-Adria-Universität Klagenfurt, Austria
Email: christian.timmerer@itec.aau.at

How to Analyze and Optimize the Encoding Latency for Multiview Video Coding
A short review for "A Framework for the Analysis and Optimization of Encoding Latency for Multiview Video"
Edited by Christian Timmerer

Multiview video with additional scene geometry information, such as depth maps, is a widely adopted data format to enable key functionalities in new visual media systems, such as 3D Video (3DV) and Free Viewpoint Video (FVV) 0. Given that the data size of multiview video grows linearly with the number of cameras, while the available bandwidth is generally limited, new schemes for an efficient compression for multiview video [2] and additional data [3] have been under investigation in recent years.

The authors argue that the design of multiview prediction structures for multiview video coding [4] has been mostly focused on improving rate-distortion (RD) performance, ignoring important differences in the latency behavior of the resulting codecs. These differences in latency may be critical for delay constrained applications such as immersive video conferencing scenarios, in which the end-to-end delay, the communication latency, needs to be kept low in order to preserve interactivity [5]. In hybrid video encoders there is a clear trade-off between RD performance and encoding delay, mainly due to the use of backward prediction and hierarchical prediction structures. In single-view video encoders, the encoding delay can be easily estimated and reduced by simple decisions on the design of prediction structures.

The analysis of the encoding delay in the case of multiview video is more challenging as it requires to handle more complex dependency structures than in single-view video, including not only temporal but also inter-view prediction. Additionally, the fact that the encoder may have to manage the encoding of several frames at the same time (frames from several views), due to the inherent parallel nature of multiview video, makes the characteristics of multi-processor hardware platforms play a significant role in the analysis.

In this paper, the authors propose a general framework for the characterization of the encoding latency in multiview encoders that captures the influence of 1) the prediction structure and 2) the hardware encoder model. This framework allows a systematic analysis of the encoding latency for arbitrary multiview prediction structures in a multiview encoder. The primary element of the proposed framework is an encoding latency model based on graph theory algorithms that assumes that the processing capacity of the encoder is essentially unbounded, i.e., the directed acyclic graph encoding latency (DAGEL) model. It can be seen as a task scheduling model [6] (the encoding of a frame is the task unit) that is used to compute the encoding latency rather than the schedule length. The paper also demonstrates that, despite the assumption of unbounded processing capacity, the encoding latency values obtained with the DAGEL model are accurate for multiview encoders with a finite number of processors greater than a required minimum, which can be identified. Otherwise, results provided by the DAGEL model represent a lower bound to the actual encoding latency of the encoder.

As an example of the applications of the DAGEL model, the authors show how it can be used to reduce the encoding latency of a given multiview prediction structure in order to meet a target value while preserving as much as possible the RD performance. In this approach, the objective is to prune the minimum number of frame dependencies (those that introduce a higher encoding delay in the original structure) until the latency target value is achieved. Therefore, the degradation of RD performance due to removal of prediction dependencies is limited. Finally, the authors demonstrate that the pruned prediction structures still produce a minimum encoding latency, as compared to other pruning options, even in hardware platforms models that

do not meet the minimum requirements in terms of the number of processors of the DAGEL model.

Following this research direction, future work includes the extension of this framework to multiview decoders and the use of graph models to analyze the delay behavior in more realistic encoder/decoder hardware architectures [7].
This paper is nominated by Cha Zhang of the MMTC 3D Processing, Rendering and Communication (3DPRC) Interest Group.

References:

[1] P. Merkle, K. Mueller, and T. Wiegand, "3D video: acquisition, coding, and display," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 2, pp. 946–950, 2010.

[2] A. Vetro, T. Wiegand, and G. Sullivan, "Overview of the stereo and multiview video coding extensions of the H.264/MPEG-4 AVC standard," *Proceedings of the IEEE*, vol. 99, no. 4, pp. 626–642, Apr. 2011.

[3] ISO/IEC JTC1/SC29/WG11, "Call for Proposals on 3D Video Coding Technology," *MPEG output doc. N12036*, Geneva, Switzerland, Mar. 2011.

[4] P. Merkle, A. Smolic, K. Müller, and T. Wiegand, "Efficient prediction structures for multiview video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 11, pp. 1461–1473, Nov. 2007.

[5] G. Karlsson, "Asynchronous transfer of video," *IEEE Communication Magazine*, vol. 34, no. 8, pp. 118–126, Aug. 1996.

[6] Y.-K. Kwok and I. Ahmad, "Static scheduling algorithms for allocating directed task graphs to multiprocessors," *ACM Computing Surveys*, vol. 31, no. 4, pp. 406–471, Dec. 1999.

[7] P. Carballeira, J. Cabrera, F. Jaureguizar and N. García, "Systematic Analysis of the Decoding Delay in Multiview Video", *Journal of Visual Communication and Image Representation, Special Issue on Advances in 3D Video Processing*, (in press) (doi: 10.1016/j.jvcir.2013.04.004).

Christian Timmerer is an assistant professor at the Institute of Information Technology (ITEC), Alpen-Adria-Universität Klagenfurt, Austria. His research interests include immersive multimedia communication, streaming, adaptation, and Quality of Experience with more than 100 publications in this domain. He was the general chair of WIAMIS'08, ISWM'09, EUMOB'09, AVSTP2P'10, WoMAN'11, QoMEX'13 and has participated in several EC-funded projects, notably DANAE, ENTHRONE, P2P-Next, ALICANTE, QUALINET, 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.

Does Kinect Provide a Simply and Cheap Solution for Telepresence?

A short review for "Enhanced Personal Autostereoscopic Telepresence System Using Commodity Depth Cameras"

Edited by Irene Cheng

Telepresence technology enables a user to feel engaged as if he/she is a part of the virtual scene. "Telepresence" and "Kinect" are often associated when either keyword is searched on the Internet. Applications, starting from the entertaining "Kinect Star Wars" to Kinect-based medical image exploration and collaborative telepresence in a social setting, have become commonplace since the launch of Kinect depth sensors in 2010.

The paper presents a low-cost Kinect-based telepresence system that offers real-time 3D scene capturing and head-tracked stereo 3D display without the user wearing any eyewear device. The system is an enhancement of the authors previous version published in ISMAR 2011 [1]. There have been quite a few telepresence systems being developed before, but none of them is based on Kinect and addresses the many issues associated with an array of depth sensors. The appearling features about Kinect lie in its low cost and simplicity. The paper shows a way to develop a cheap telepresence system, which was traditionally of high cost. But Kinect also brings in unique challenges including various artifacts in depth maps such as holes and noises, and the interference among multiple Kinects. Many existing techniques are employed or adapted by the authors for denoising, hole-filling, smoothing, data merger, surface generation, color correction and head tracking. The system takes advantage of a fully GPU-accelerated data processing and rendering pipeline. The main contribution lies in the integration of various existing techniques to deliver a workable solution. The complete software and hardware framework for implementing the system is presented, including GPU-acceleration.

The Introduction gives a flavor of the evolution of 3D data acquisition using depth cameras and visualization using eyewear in a telepresence environment since late 90s'. The proposed system is based on the inexpensive Microsoft Kinect sensor, providing a 58° x 45° field of view with high depth accuracy. After proper sensors calibration, an entire room-sized scene can be captured in real-time. By combining 2D eye detection technique and depth data, Kinect is able to offer a markerless tracking solution. However, there are challenges that the authors encountered in using Kinect sensors for implementing their system. Inter-unit interference is a major problem because each sensor projects a fixed structured light pattern of similar wavelengths. There is also difficulty for presenting seamless integration of color-matched data between cameras. Thus the enhancements include introducing a software solution to the Kinect interference problem and a visibility-based method to merge data between cameras, as well as for dynamic color matching between color-plus-depth cameras. The hardware configuration and software implementation are detailed in the paper. Interested readers can refer to Section 4.2 to understand how the multi-Kinect interference problem is addressed. Color matching is a common problem in many camera systems. Even the same camera model device often exhibits different color gamuts [2] and so as Kinect sensors. The current available Kinect driver (at the time of this paper) allows only automatic color and exposure control. Thus color values can vary dramatically between adjacent sensors. Here the authors argue that applying traditional color matching techniques is ineffective because automatic control may alter color balances. They introduce using depth information to find color correspondences between cameras and build a color matching function. Details are described in Section 4.6. Another enhancement explored in this paper is related to eye position tracking accuracy, speed and latency described in Section 4.7.

Comparison of results shows the good performance of the proposed telepresence system. In the Conclusion, the authors point out that although the system is functional, the output

IEEE COMSOC MMTC R-Letter

image quality still needs improvement, and in particular the temporal noise artifacts present at the edges of objects at depth pixel level. Instead of presenting all the technical discussions and computational analysis, what I like about this paper is its clarity and readability. A short overview suitable for the general readers is given at the beginning of each section followed by a
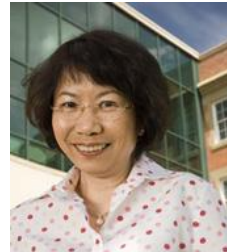
Acknowledgement:

This paper is nominated by Jianfei Cai of the MMTC 3D Processing, Rendering and Communication (3DPRC) Interest Group.

References:

[1] Ilie A, Welch G. Ensuring color consistency across multiple cameras. In: Proceedings of the tenth IEEE international conference on computer vision – volume 2. ICCV '05; Washington, DC, USA: IEEE Computer Society; 2005, p. 1268–75. ISBN 0-7695-2334-X-02. http://dx.doi.org/10.1109/ICCV.2005.88

[1] Maimone A, Fuchs H. Encumbrance-free telepresence system with real-time 3d capture and display using commodity depth cameras. In: Tenth IEEE international symposium on mixed and augmented reality (ISMAR); 2011. p. 137–46. http://dx.doi.org.10.1109/ISMAR.2011.6092379.

more in-depth explanation. The developed system shows a promising way to bring telepresence to common users, which will stimulate more subsequent multimedia communication research.

Irene Cheng, SMIEEE is the Scientific Director of the Multimedia Research Centre, and an Adjunct Professor in the Faculty of Science, as well as the Faculty of Medicine & Dentistry, University of Alberta, Canada. She is also a Research Affiliate with the Glenrose Rehabilitation Hospital in Alberta, Canada. She is a Co-Chair of the IEEE SMC Society, Human Perception in Vision, Graphics and Multimedia Technical Committee; was the Chair of the IEEE Northern Canada Section, Engineering in Medicine and Biological Science (EMBS) Chapter (2009-2011), and the Chair of the IEEE Communication Society, Multimedia Technical Committee 3D Processing, Render and Communication (MMTC) Interest Group (2010-2012). She is now the Director of the Review-Letter Editorial Board of MMTC (2012-2014).

Over the last ten years, she has more than 110 international peer-reviewed publications including 2 books and 31 journals. Her research interests include multimedia communication techniques, Quality of Experience (QoE), Levels-of-detail, 3D Graphics Visualization and Perceptual Quality Evaluation. In particular, she introduced applying human perception – Just-Noticeable-Difference – following psychophysical methodology to generate multi-scale 3D models.

Extending Signal Processing Techniques to Graph Domain

A short review for "Perfect Reconstruction Two-Channel Wavelet Filter Banks for Graph Structured Data"

Edited by Jun Zhou

Graph theory has been successfully adopted in many computer vision and pattern recognition applications. When dealing with large scale data, one of the problems that hinders the wide adoption of graphical models is the very high computational complexity caused by large number of nodes and vertices in graph. To address this challenge, one would expect that only a few nodes in the graph be used to form a compacted representation of the original graph. Then data processing can be performed only on a small neighborhood of each node. Some recent efforts in this direction have explored traditional signal processing techniques, such as wavelet transform, as possible solutions.

The paper published by Narang and Ortega in IEEE TSP is a seminal work on graph sampling and design of critically sampled wavelet filter banks on graphs. It not only provides a comprehensive review of the spatial/spectral representation of graph signals and existing work on two-channel filter banks, but also proposes the important characteristics of the sampling strategy and filter banks for perfect reconstruction of bipartite graphs.

The key idea behind this method is applying a two-channel filter banks that decompose a graph into high-pass and low-pass channels, each containing only part of the nodes in the graph after downsampling and following upsampling operations. When these two channels are combined, they form a perfect reconstruction of the original graph representation. In order to achieve such distortion-free reconstruction, an aliasing component, which is composed of filter banks and downsampling functions, shall be set to zero. Therefore, the goal of this research is to figure out what are the proper filter banks and downsampling functions to meet the above requirement.

To develop the downsampleing strategy, the authors proposed that the decomposed high-pass the low-pass channels shall contain complement node sets of the original graph. This leads to the building of a bipartition of the graph nodes [1]. Based on the graph spectral theory, this strategy generates spectral coefficients at symmetric graph frequencies around a central frequency, which is equivalent to the aliasing component of the reconstruction function.

To design the filter banks, the authors pointed out that they shall meet three conditions, *i.e.*, aliasing cancellation, perfect reconstruction, and orthogonality. Therefore, a quadrature mirror filter bank method [2] (wavelet is one of such method) was chosen and extended to bipartite graph. This method allows a single basis spectral kernel be created, while all other kernels are built on top of the basis kernel.

Whilst it is straightforward to adopt the wavelet filter banks on bipartite graph, the application of this framework to arbitrary graph requires generating a series of bipartite subgraphs from the original graph. Then each subgraph can be processed independently with a cascaded transform being implemented at the end. In this paper, the authors proposed to use the biparticity method from Harary *et al* [3] for subgraph generation.

Two experiments have been performed to demonstrate how the proposed method can be applied to image processing (as an example of regular graph) and traffic graph analysis (as an example of irregular graph). These examples show that the two-channel wavelet filter banks and the sampling method form a practical solution for graph decomposition and reconstruction. It enables efficient graph computation, which has been expected by the research community. I believe this work will generate long-term impact to the development of graph theory because it provides an elegant way

of applying signal processing techniques to solve structured pattern recognition problems.

References:

[2] S. Narang and A. Ortega, "Downsampling graphs using spectral theory," *Proceedings of the International Conference on Acoustics, Speech and Signal Processing,* pages 4208-4211, 2011.

[3] J. Johnston. "A filter family designed for use in quadrature mirror filter banks," *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal*, pages 291-294, 1980.

[4] F. Harary, D. Hsu, and Z.Miller, "The biparticity of a graph," Journal of Graph Theory, vol. 1, no. 2, pp. 131–133, 1977.
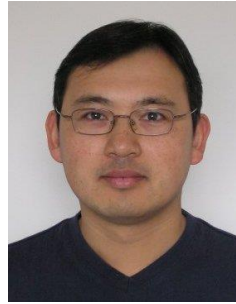
Jun Zhou received the B.S. degree in computer science and the B.E. degree in international business from Nanjing University of Science and Technology, China, in 1996 and 1998, respectively. He received the M.S. degree in computer science from Concordia University, Canada, in 2002, and the Ph.D. degree in computing science from University of Alberta, Canada, in 2006.

He joined the School of Information and Communication Technology in Griffith University as a lecturer in June 2012. Prior to this appointment, he had been a research fellow in the Australian National University, and a researcher at NICTA. His research interests are in statistical pattern recognition, interactive computer vision, and their applications to hyperspectral imaging and environmental informatics.

Fairness Resource Allocation in Blind Wireless Multimedia Communications

A short review for "Fairness Resource Allocation in Blind Wireless Multimedia Communications"
Edited by Koichi Adachi

A scenario where one base station (BS) assigns available resource to multiple multimedia users is considered. The fairness resource allocation problem is formulated with a fairness parameter α to maximize summation of user's quality of experience (QoE). Traditional α-fairness resource allocation in wireless multimedia communication systems assumes the QoE model (or utility function) of each user is available at scheduler. In such a system, balancing the tradeoff between the fairness and performance is an important task for BS as it is known that the introduction of fairness generally has a negative impact on the performance [1-4]. However, the critical assumption in most existing studies is the availability of QoE model at the scheduler, which is may not be practical.

Therefore different from previous works, a *blind* scenario is considered in this paper where BS has no knowledge of the QoE model during the whole resource allocation procedure. The answers to the following two questions are provided in this paper: 1) How to set the fairness parameter α from the perspective of performance-fairness tradeoff? 2) Given a specific fairness parameter α, how to implement the α-fairness resource allocation online?

The main contributions of this paper are lying in: *Qualitative analysis* and *Technical realization*. Some recent works theoretically analyzed the performance with special cases of a value, e.g., proportional fairness ($\alpha \rightarrow 1$) and max-min fairness ($\alpha \rightarrow \infty$) [1-3,5,6]. For more general case of α, only empirical modelling was provided [2]. An exact expression for the upper bound of the performance loss is caused by α-fairness, characterizing the fairness-performance tradeoff (*Theorem 1*). This enables a BS scheduler to choose the appropriate fairness parameter α, which answers the first question.

For technical realization of a specific fairness resource allocation, convex optimization [4], [7-10] and game theory [10-15] are generally used

in previous works. However, both approaches requires the utility function of each user to be available at BS or the controller. In this paper, a blind fairness-aware resource allocation problem is decomposed into two subproblems to describe the behaviors of the users and BS. The second question is answered by proposing a bidding game for the reconciliation between the two subproblems. The authors show that although all the users behave selfishly, any specific a-fairness scheme can be implemented by the bidding game between the users and BS (*Theorem 2*).

In *Theorem 1*, the upper bound of the performance loss incurred by the α-fairness is derived. The derived upper bound connects the number of users in the system and the fairness parameter α and it is independent from the QoE model as long as it satisfies some assumption, which is not too restrictive. The authors propose a bidding game to decompose the optimization problem into two subproblems similar to [16]. The first subproblem describes the behaviors of the users and the second the BS, respectively. Each user tries to maximize its own objective function by bidding game. Then, for given payment from each user, the BS strives to find optimal transmission rate for each user to maximize its own objective function (*control function*). The control function is composed of the bidding money, allocated resource, and fairness parameter α. In *Theorem 2*, the format of control function is given to resolve the original optimization problem. To make the proposed bidding game work smoothly in a realistic blind scenario, the assumption that each user does not cheat during the whole bidding process is required. The authors also provide the counter-measures for this issue. The convergence property of the proposed bidding game is given in *Theorem 3*.

The performance evaluation of the proposed bidding game is performed based on real-world traces consisting of three multimedia applications: audio, file, and video. For

comparison, the resource allocation with full-information, where the BS knows the QoE model of each user in advance, is considered [17]. Firstly, the accuracy of the derived upper bound of the loss function is confirmed by comparing it with the real observed loss values. It is shown that the real loss value is close to the obtained upper bound for different value of α. It is also confirmed that a higher value of α yields a higher performance loss, which is consistent with the previous works [2]. Furthermore, larger number of users also incurs a larger performance loss. This observation suggests a basic operation rule for BS: when the system has a relatively small number of users, BS can achieve fair allocations without significant performance deterioration. However, in the case with a large number of users, BS should be careful to employ fairness as it will easily lead to a large performance loss. Secondly, the total MOS value of the proposed bidding game scheme is compared with those of the full-information case. The proposed bidding game almost is shown to achieve the same performance as the full-information case when α is large. Finally, the convergence property of the proposed bidding game is clarified that 1) it converges within a limited number of iteration and 2) the number of users affects the convergence time.

The derived upper bound of the performance loss incurred by the α-fairness is useful to characterize and understand the tradeoff between the performance loss and the fairness. Since the proposed bidding game does not require knowledge of QoE of each user, it is applicable to practical multimedia communication system.

References:

[1] T. Lan, et al., *An Axiomatic Theory of Fairness in Resource Allocation*, Princeton, NJ, USA: Princeton Univ., 2009.

[2] L. Kaplow and S. Shavell, *Fairness Versus Welfare*, Cambridge, NJ, USA: Harvard Univ., 2002.

[3] L. Zhou, et al., "Distributed media services in P-based vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 60, no. 2, pp. 692–703, Feb. 2011.

[4] L. Zhou, et al., "Distributed scheduling scheme for video streaming over multi-channel multi-radio multi-hop wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 409–419, Mar. 2010.

[5] T. Nguyen, et al., "Efficient multimedia distribution in source constraint networks," *IEEE Trans. Multimedia*, vol. 10, no. 3, pp. 523–537, Apr. 2008.

[6] Y. Li, et al., "Content-aware distortion-fair video streaming in congested networks," *IEEE Trans. Multimedia*, vol. 11, no. 6, pp. 1182–1193, Oct. 2009.

[7] D. Hu, et al, "Scalable video multicast in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 334–344, Mar. 2010.

[8] H.-P. Shiang and M. v. d. Schaar, "Information-constrained resource allocation in multi-camera wireless surveillance networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 4, pp. 505–517, Apr. 2010.

[9] Y. Zhang, et al., "Multihop packet delay bound violation modeling for resource allocation in video streaming over mesh networks," *IEEE Trans. Multimedia*, vol. 12, no. 8, pp. 886–900, Dec. 2010.

[10] H. Hu, et al., "Peer-to-peer streaming of layered video: Efficiency, fairness and incentive," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 8, pp. 1013–1026, Aug. 2011.

[11] Z. Han, et al, "Fair multiuser channel allocation for OFDMA networks using Nash bargaining solutions and coalitions," *IEEE Trans. Communun.*, vol. 53, no. 8, pp. 1366–1376, Aug. 2005.

[12] H. Park and M. van der Schaar, "Fairness strategies for wireless resource allocation among autonomous multimedia users," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 2, pp. 297–309, Feb. 2010.

[13] Y. Chen, et al., "Multiuser rate allocation games for multimedia communications," *IEEE Trans. Multimedia*, vol. 11, no. 6, pp. 1170–1181, Oct. 2009.

[14] Q. Zhang and G. Liu, "Rate allocation games in multiuser multimedia communications," *IET Commun.*, vol. 5, no. 3, pp. 396–407, 2011.

[15] H. Park and M. v. d. Schaar, "Bargaining strategies for networked multimedia resource management," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3496–3511, Jul. 2007.

[16] F. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecommun.*, vol. 8, no. 1, pp. 33–37, 1997.

[17] A. Khan, et al, "Quality of experience-driven adaptation scheme for video applications over wireless networks," *IET*

Commun., vol. 4, no. 11, pp. 1337–1347, 2010.

Koichi Adachi received the B.E., M.E., and Ph.D degrees in engineering from Keio University, Japan, in 2005, 2007, and 2009 respectively. From 2007 to 2010, he was a Japan Society for the Promotion of Science (JSPS) research fellow. Currently he is with the Institute for Infocomm Research, A*STAR, in Singapore. His research interests include cooperative communications. He was the visiting researcher at City University of Hong Kong in April 2009 and the visiting research fellow at University of Kent from June to Aug 2009.

Improved Cloud Resource Utilization for IPTV Transmission

A short review for "Optimizing Cloud Resources for Delivering IPTV Services through Virtualization"
Edited by Carl James Debono

Internet Protocol-based video delivery is increasing in popularity with the result that its resource requirements are continuously growing. It is estimated that by the year 2017 video traffic will account 69% of the total consumer's Internet traffic [1]. Content and service providers typically configure their resources such that they can handle peak demands of each service they provide across the subscriber population. However, this means that the resources are under-utilized during non-peak times. The predominant types of Internet Protocol TeleVision (IPTV) services that the authors of the original paper focus on are Live TV and Video On Demand (VoD) services, as these are the primary capabilities supported by service providers. Live TV provides a very bursty workload profile with tight deadlines, whilst on the other hand VoD has a relatively steady load and is less stringent on delay requirements.

The solution presented takes advantage of the temporal differences in the demands from these IPTV workloads to better utilize the servers that were deployed to support these services. While VoD is delivered via unicast, Live TV is delivered over multicast to reduce bandwidth demands. However, to support Instant Channel Change (ICC) in Live TV, service providers send a unicast stream for that channel for a short period of time to keep a good quality of experience. If a number of users change their channels around the same period of time, this produces a large burst load on the server that has to support the corresponding number of users. Compared to the ICC workload which is very bursty and has a large peak to average ratio, VoD has a relatively steady load and imposes a relatively lax delay requirement. By multiplexing across these services, the resource requirements for supporting the combined set of services can be reduced.

Two services that have workloads which differ significantly over time can be combined on the same virtualized platform. This allows for scaling of the number of resources according to each service's current workloads. It is, however, possible that the peak workload of different services may overlap. Under such scenarios, the benefit of a virtualized infrastructure diminishes, unless there is an opportunity to time shift one of the services in anticipation of the other service's requirements to avoid having to deliver both services at the same time instant. In general, the cloud service provider strives to optimize the cost for all time instants, not necessarily just reducing the peak server load. The authors of the original paper consider a generalized cost function, which can be specialized to a peak server load or tiered pricing as possible options.

Consider a scenario with multiple services, each having its own deadline constraint. The optimization problem tackled is to determine the number of servers that are needed at each time instant by minimizing a generalized cost function, while at the same time satisfying all the deadlines associated with these services. To achieve this, the authors identified the server-capacity region which is formed by servers at each time instant (or server tuple), such that all the arriving requests meet their deadlines. The results show that for any server tuple with integer entries inside the server-capacity region, adopting an Earliest Deadline First (EDF) strategy [2] manages to service all the requests without missing deadlines.

After identifying the server-capacity region, several cost functions were considered, namely: a separable concave function, a separable convex function, and a maximum function. The original authors find that the feasible set of server tuples is all integer tuples in the server-capacity region. This constraint increases the difficulty of finding a solution to the problem. However, for a piecewise linear separable convex function, an algorithm that minimizes the cost function can easily be found. Moreover, only causal information of the requests coming at each time-instant are required. On the other hand, for

concave cost functions, the original paper reports that the integer constraint can be relaxed, since all the corner points of the region of interest have integer coordinates. Therefore, concave programming techniques without integer constraints [3] can be applied. The paper also investigates a "two-tier" cost function, as a basic strategy for cloud service pricing. A closed form expression is found for the optimal number of servers needed in such a scenario. The algorithm developed reduces the run time duration to $O(T^2)$, compared to the $O(T^3)$ complexity required when directly implementing the expression.

The authors of the original paper study two approaches for sharing the resources: (a) postponing and (b) advancing the delivery of VoD. The postponement approach assumes that chunk i is requested at time t, and has a deadline which is d time units after the initial request. Conversely, the advancement technique assumes that all the chunks are requested when the video is first demanded by the user and that each chunk has a deadline for its playout time. A series of simulations were set up for both scenarios to study the effect of varying the ICC durations and the delay tolerance of VoD services on the total number of servers needed for the combined workload. Two cost functions were considered to determine the number of servers, namely, the maximum and the piecewise-linear convex functions. A limit on the downlink bandwidth was also considered for the VoD delivery advancement method. The reported results show that server bandwidth savings between 10% and 32% can be obtained by anticipating the ICC load and shift the VoD load ahead of the ICC bursts.

The reported results show that if peak loads are considered, the algorithm is capable of reducing the peak by around 24%. This has a direct impact on the cost of the infrastructure since 24% fewer servers would be required to serve all the requests in the simulated scenario.

The possibility of predicting and time-shifting IPTV load in wired and wireless networks allows for better utilization of the cloud infrastructure. Further work is needed to improve prediction techniques and include other parameters such that the ever increasing demand of video services can be sustained. Furthermore, storage requirements and other traffic on the network need also to be considered in the optimization

strategy. The solution presented relies on homogeneous servers, something which cannot be guaranteed, and thus heterogeneous systems need to be studied too. Moreover, low-complexity security solutions to avoid eavesdropping of the video data, and its related processing, need to be developed as these services keep proliferating.

References:

[1] Cisco Visual Networking Index: Forecast and Methodology, 2012-2017. [Online] Available: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-481360.pdf

[2] J. A. Stanković, M. Spuri, K. Ramamritham, and G. C. Buttazzo, *Deadline Scheduling for Real-Time Systems : Edf and Related Algorithm,* Norwell, MA, USA: Kluwer, 1998.

[3] N. V. Thoai, and H. Tui, "Convergent algorithms for minimizing a concave function," *Mathematics of Operations Research*, vol. 5, no. 4, pp. 556-566, 1980.

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 an Associate Professor. He is currently the Deputy Dean of the Faculty of ICT at the University of Malta.

Prof. 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 Vice-Chair of Technical Activities for 2013. 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.

Automatic Output Motion Depiction using 2D videos as Direct Input

A short review for "Human Movement Summarization and Depiction from Videos"

Edited by Irene Cheng

Motion analysis is a popular research topic studied in different disciplines including computer vision, image processing, communication and computer graphic. It has diverse applications, such as training, medicine, entertainment, surveillance and navigation. The analysis can be based on 2D videos or 3D motion capture data. Starting from 3D motion data often produces good results [1, 2, 3] because the motion sequence is view invariant and not affected by visual occlusion. However, capturing 3D motion requires the use of sensors to track feature points movements. The apparatus setup can be complex and expensive, and thus the operation is not generally accessible by non-professionals. In contrast, motion analysis based on 2D videos can be affected by view variation and occlusion, but video data is easily be obtained by amateurs.

This paper presents an automatic method to depict human movement using 2D videos as direct input. The method does not require 3D motion capture data or manual intervention. The method analyzes inter-frame as well as frame group trajectories of body feature points based on both body part detection and optical flow adjusted by error correction. The output is color coded arrows and motion particles, which are particularly useful for training and rehabilitation purposes to show how a specific movement can be performed. The compact depiction can also be used for trajectory integration, action recognition and movement analysis. There are three steps in the proposed method: segment videos into sub-actions, track feature points, and depict motion using estimated movement. The authors tested their method on a number of videos with satisfactory results.

One finding is that the number of clusters is difficult to determine when clustering based methods are applied to action segmentation. Thus in the first step, the authors use cluster of streamlines to complement an action boundary detection scheme. Seed points are randomly selected in each frame. The trajectories are generated by linking the points between frames. In the current implementation, a group of 15 frames is used to compute motion trajectories. The authors realize that by using this simple scheme there is no guarantee the motion trajectories will intersect.

However, for motion depiction purpose, only the overall path is needed and thus a rough representation is adequate. The obtained motion trajectories are then shifted so that they all start from point (0,0,0). The three coordinates represent x, y and time. The trajectories are further projected to the xy plane and the 2D coordinates of points on the trajectories are normalized. In order to detect movement boundaries, it requires the action features be stable when body parts keep their motion direction. The distance of streamlines between successive time instants is computed and the results are plotted on a 1D curve. Potential action changes are indicated by local maxima on the curve, which are used in the subsequent motion estimation step.

In the next step, body parts are detected with associated feature points trajectories. Ten body parts including head, torso, four half arms and four half legs are detected. The paper points out that the detector does not distinguish the left and right arms and legs, and there are many detection errors. To obtain accurate and efficient body parts association, a linear *multiple shortest path* following problem is formulated. I find the graphs constructed in this paper for each pair of limbs, and the inclusion of four possible body part assignments at each layer in the graph quite interesting. Graphical models are commonly used in research and are often an efficient mean to solve or simplify complex high dimensional problems. Since the trajectory can have a long span if propagating point location from frame to frame using optical flow, thus the method formulates the point cloud trajectory estimation as an optimization problem based on constraining trajectories by the body part direction, optical flow and object foreground estimation. Error body part movement estimation is then cleaned up.

Static illustration integrated with artwork is adopted in the final step to translate human movement estimation into graphics representations. The authors use directional arrows to depict the body part movement, particles to illustrate the local motion, and ghost images to indicate transitional and ending poses. Mean trajectory is computed and used as the center line of the arrow with predefined width. In order to

remove the error from the mean trajectory, they fit a trajectory to a second-order polynomial, which is sufficient to quantify the general shape of the motion. The color at each point on the arrows is proportional to the speed of the motion. Only directional arrows of significant length are kept. The final image with overlapping arrows is then rendered.

The proposed method was tested on two ballet sequences and two recorded videos containing complex motions. The results demonstrate that even with occlusion and low quality video shots taken with a shaky hand-held camera, the output motion depictions are satisfactory. Though more tests could have further validated the robustness of the approach, I find this paper interesting and can inspire further research on motion data analysis.

References:

[1] J. Assa, Y. Caspi and D. Cohen-or, "Action Synopsis: Pose Selection and Illustration," SIGGRPAH 2005.
[2] S. Bouvier-Zappa, V. Ostromoukhow and P. Poulin, "Motion Cues for Illustration of Skeletal Motion Capture Data," Symposium on Non-Photorealistic Animation and Rendering 2007.
[3] J. Barbic, A. Safonova J. Pan, C. Faloutsos, J.K. Hodgins and N.S. Pollard, "Segmenting Motion Capture Data into Distinct Behaviors," ACM Graphics Interface 2004.

Irene Cheng, SMIEEE is the Scientific Director of the Multimedia Research Centre, and an Adjunct Professor in the Faculty of Science, as well as the Faculty of Medicine & Dentistry, University of Alberta, Canada. She is also a Research Affiliate with the Glenrose Rehabilitation Hospital in Alberta, Canada. She is a Co-Chair of the IEEE SMC Society, Human Perception in Vision, Graphics and Multimedia Technical Committee; was the Chair of the IEEE Northern Canada Section, Engineering in Medicine and Biological Science (EMBS) Chapter (2009-2011), and the Chair of the IEEE Communication Society, Multimedia Technical Committee 3D Processing, Render and Communication (MMTC) Interest Group (2010-2012). She is now the Director of the Review-Letter Editorial Board of MMTC (2012-2014).
Over the last ten years, she has more than 110 international peer-reviewed publications including 2 books and 31 journals. Her research interests include multimedia communication techniques, Quality of Experience (QoE), Levels-of-detail, 3D Graphics Visualization and Perceptual Quality Evaluation. In particular, she introduced applying human perception – Just-Noticeable-Difference – following psychophysical methodology to generate multi-scale 3D models.
.

Identify Visualizable Concepts

A short review for "Mining Visualness"
Edited by Weiyi Zhang

Despite decades of successful research on multimedia and computer vision, the semantic gap between low-level visual features and high-level semantic concepts remains a problem. Instead of generating more powerful features or learning more intelligent models, researchers have started to investigate which concepts can be more easily modeled by existing visual features [1, 2, 3]. To understand to what extent a concept has visual characteristics, i.e. "visualness", has many values. For instance, it can benefit recent research efforts on constructing image databases [4, 5]. These efforts generally attempt to attach images onto pre-defined lexical ontology, while existing ontology were built without taking visual characteristics into consideration. Knowing which concepts are more likely to find relevant images will help save labors and control noises in database construction. Visualness estimation is also useful for image-to-text [2, 6] and text-to-image [7] translation, e.g., words of more visualizable concepts are potentially better annotations for an image.

Albeit the usefulness, a general solution of visualness estimation faces many challenges: 1) It is unknown which concepts or which types of concepts are visualizable, i.e. whether representative images can be found to visualize its semantics. For instance, "dignity" and "fragrant" are both abstract nouns, but the former is more difficult to visualize as "fragrant" is closely related to visual concepts such as flowers and fruits; 2) Different visual concepts have diverse visual compactness and consistency, especially for collective nouns (e.g., "animal") and ambiguous concepts (e.g., "apple", which may represent a kind of fruit or a company); and 3) Even though a concept is highly visualizable, it may still be difficult to capture the visual characteristics due to the semantic gap. Few previous works in the literature have touched this research topic. They either use pre-defined concept list or insufficiently assumed prototypical concepts.

In this paper, the authors attempt to discover and quantify the visualness of concepts automatically from a large-scale dataset. The quantitative measure of a concept is based on visual and semantic synsets (named Visualsets), rather than a single image cluster or keyword as in previous works. Visualsets perform disambiguation on the semantics of a concept and ensures visual compactness and consistency,which is inspired by synsets in the work of ImageNet[4] and Visual Synsets[6]. In this paper's approach, a visualset is a group of visually similar images and related words, both are scored by their membership probabilities. Visualsets contain prototypical visual cues as well as prototypical semantic concepts. Given the visualsets, the visualness of a concept is thus modeled as a mixture distribution on its corresponding visualsets. Moreover, the authors discover both simple concepts (keywords) and compound concepts (combination of unique keywords) simultaneously from the generated visualsets.

The proposed approach contains three steps: 1) build an image heterogeneous graph with attribute nodes generated from multitype features; Given a (noisily) tagged image dataset such as a web image collection, the proposed scheme connects the images into a graph to facilitate the clustering approach for visualsets mining. Specifically, the scheme extracts multiple types of visual features and textual feature for images to generate attribute nodes. The edges of the graph are defined by links between images and attribute nodes instead of image similarities which are generally adopted in previous works [2]. 2) mine visualsets from the heterogeneous graph with an iterative clustering-ranking algorithm; an iterative ranking-clustering approach is applied to form visual and textual synsets, i.e. visualsets. In each iteration, it starts with the guess on image clusters. Based on the guess, the solution scores and ranks each image as well as attribute nodes in each visualset. Images are then mapped to the feature space defined by the visualsets mixture model. Clusters are refined based on the estimated posteriors,

which gives the guess on image clusters for the next iteration. 3) estimate visualness of concepts with visualsets: after the clustering-ranking approach converges, the scheme estimates the visualness of concepts (simple and compound) from the visualsets based on final scores of images and attribute nodes.

The authors also conducted extensive experiments to verify their proposed scheme, using the NUS-WIDE dataset containing 269,648 images and 5,018 unique tag words from Flickr. Two types of global features, 64-D color histogram and 512-D GIST, are extracted. Each type of global features is further clustered into 2000 clusters by k-means clustering, whose centers form the set-based attribute nodes of the image heterogeneous graph. Local SIFT features are also extracted and clustered into 2000 visual words by k-means clustering, based on which word-based attribute nodes are generated. The proposed scheme achieved promising results and discovered 26,378 visualizable compound concepts from NUS-WIDE.

References:

[1] K. Yanai and K. Barnard, "Image region entropy: a measure of visualness" of web images associated with one concept," in ACM Multimedia, 2005.

[2] Y. Lu, L. Zhang, Q. Tian, and W.Y. Ma, "What are the highlevel concepts with small semantic gaps?," in CVPR, 2008.

[3] J.W. Jeong, X.J.Wang, and D.H. Lee, "Towards measuring the visualness of a concept," in CIKM, 2012.

[4] J. Deng, W. Dong, R. Socher, L.J. Li, K. Li, and L. Fei-Fei,"Imagenet: A large-scale hierarchical image database," in CVPR, 2009.

[5] X.J. Wang, Z. Xu, L. Zhang, C. Liu, and Y. Rui, "Towards indexing representative images on the web," in ACM Multimedia Brave New Idea Track, 2012.

[6] D. Tsai, Y. Jing, Y. Liu, H. A. Rowley, S. Ioffe, and J.M. Rehg, Large-scale image annotation using visual synset," in ICCV, 2011.

[7] X. Zhu, A. B. Goldberg, M. Eldawy, C. R. Dyer, and B. Strock, "A text-to-picture synthesis system for augmenting communication," in AAAI, 2007.

Weiyi Zhang is currently a Research Staff Member of the Network Evolution Research Department at AT&T Labs Research, Middletown, NJ. Before join AT&T Labs Research, he was an Assistant Professor at the Computer Science Department, North Dakota State University, Fargo, North Dakota, from 2007 to 2010. His research interests include routing, scheduling, and cross-layer design in wireless networks, localization and coverage issues in wireless sensor networks, survivable design and quality-of-service provisioning of communication networks. He has published more than 70 refereed papers in his research areas, including papers in prestigious conferences and journals such as IEEE INFOCOM, ACM MobiHoc, ICDCS, IEEE/ACM Transactions on Networking, ACM Wireless Networks, IEEE Transactions on Vehicular Technology and IEEE Journal on Selected Areas in Communications. He received AT&T Labs Research Excellence Award in 2013, Best Paper Award in 2007 from IEEE Global Communications Conference (GLOBECOM'2007). He has been serving on the technical or executive committee of many internationally reputable conferences, such as IEEE INFOCOM. He was the Finance Chair of IEEE IWQoS'2009, and serves the Student Travel Grant Chair of IEEE INFOCOM'2011.

Cooperative Video Summary Delivery over Wireless Networks

A short review for "Video Summary Delivery over Cooperative Wireless Networks"
Edited by Xiaoli Chu

It has been widely accepted that cooperative wireless communications can improve node connectivity, increase link throughput, save network power consumption, and so on. Popular cooperative communication schemes include amplify-and-forward (AF), decode-and-forward (DF), and coded cooperation (CC), but none of them has the ability to perform video rate and quality adaptation for video transmissions at the relay node.

Video summarization generates a short summary of the content of a possibly huge volume of video data, in a way that the essential/important information of the original data is delivered to the receiver. Video summarization can significantly reduce the data amount to be transmitted while maintaining video content coverage.

For resource-limited wireless video applications that have stringent requirements on power consumption, video delivery timeliness, video quality and content coverage, such as video surveillance, integrating video summarization with cooperative communications would have significant benefits. However, due to the complexity of combining scene understanding, video coding and wireless communications, there have been limited research efforts reported on video summary transmission over cooperative wireless networks [1-3].

In this paper, the authors propose a cross-layer optimization framework for cooperative video summary transmission. A decode-process-and-forward (DPF) scheme, where a relay node with video processing capability extracts the most useful information from video summary frames and generates a concise version of the summary frame, namely summary of summary (SoS), is proposed for video summary transmission. The destination node then uses the received SoS information to enhance its error concealment capability, resulting in an improved video reconstruction quality. In the proposed cross-layer framework, source coding, relay processing,

power allocation between source and relay, and error concealment strategy are jointly considered.

The SoS can be obtained after a few video processing steps with various levels of computational complexity. Depending on specific system settings and network conditions, the video processing methods can be chosen from down-sampling the image, filtering the high-frequency components of the image, encoding the video frame with a lower bit budget, extracting the region of interest (ROI) information [4], and dropping the current video summary frame.

The video processing methods and error concealment strategies used in the relay node and the destination are known to the system controller, which resides in the source node, controls and optimizes the parameter settings of all modules based on application requirements, channel conditions, and computational complexity. For a practical solution, trade-off needs to be made between the computational capability and the optimality of a solution [5].

The relationship between the video frame loss probability and the packet loss probability of each link depends on the packet encapsulation or packet fragmentation scheme used. In the problem formulation, it is assumed that each video summary frame is compressed into one packet, so the video frame loss probability is the same as the packet loss probability of each link.

Experiments have been carried out to evaluate the performance of the proposed DPF scheme by using H.264/AVC JM 12.2, and through performance comparison (in terms of peak signal-to-noise ratio (PSNR)) with conventional direct transmission, DF, and multipath transmission (MT). The experimental results show that the proposed DPF scheme significantly outperforms the other three schemes. This indicates that the proposed DPF scheme is able to not only exploit the channel fading diversity of cooperative communications but also adapt resource

allocation through flexible video processing at the relay node.

The numerical results also show that excessive power consumption by either the source or relay node will not improve the distortion performance significantly. The power allocation between the source and relay nodes needs to be optimized to achieve a remarkable performance gain with the proposed DPF scheme.

In conclusion, the proposed DPF scheme can achieve a significant improvement in terms of the received video quality, as compared with existing cooperative transmission schemes in the literature.

References:

[1] Z. Li, *et al.*, "Video summarization for energy efficient wireless streaming," *Proc. SPIE Visual Commun. and Image Processing*, Beijing, China, 2005.

[2] P. V. Pahalawatta, *et al.*, "Rate-distortion optimized video summary generation and transmission over packet lossy networks," *Proc. SPIE Image and Video Commun. and Processing*, San Jose, CA, 2005.

[3] D. Wu, S. Ci and H. Wang, "Cross-layer optimization for video summary transmission over wireless networks," *IEEE JSAC*, vol. 25, no. 4, pp. 841-850, May 2007.

[4] D. Wu, *et al.*, "Quality-driven optimization for content-aware real-time video streaming in wireless mesh networks," *Proc. IEEE GlobeCom*, New Orleans, LA, Dec. 2008.

[5] G. M. Schuster and A. K. Katsaggelos, Rate-Distortion Based Video Compression: Optimal Video Frame Compression and Object Boundary Encoding, Kluwer, 1997.

Xiaoli Chu is a Lecturer in the Department of Electronic and Electrical Engineering at the University of Sheffield, UK. She received the B.Eng. degree in Electronic and Information Engineering from Xi'an Jiao Tong University, China, in 2001, and the Ph.D. degree in Electrical and Electronic Engineering from the Hong Kong University of Science and Technology in 2005. From Sep. 2005 to Apr. 2012, she was with the Centre for Telecommunications Research at King's College London. Her current research interests include heterogeneous networks, cooperative communications, cognitive communications, and green radios. She has published more than 60 peer-reviewed journal and conference papers. She is the leading editor/author of the Cambridge University Press book, Heterogeneous Cellular Networks, May 2013. She was a guest editor of the Special Issue on Cooperative Femtocell Networks (Oct. 2012) for ACM/Springer Journal of Mobile Networks & Applications, and is a guest editor of the Special Section on Green Mobile Multimedia Communications (Apr. 2014) for IEEE Transactions on Vehicular Technology.

Paper Nomination Policy

Following the direction of MMTC, the R-Letter platform aims at providing research exchange, which includes examining systems, applications, services and techniques where multiple media are used to deliver results. Multimedia include, but are not restricted to, voice, video, image, music, data and executable code. The scope covers not only the underlying networking systems, but also visual, gesture, signal and other aspects of communication.

Any HIGH QUALITY paper published in Communications Society journals/magazine, MMTC sponsored conferences, IEEE proceedings or other distinguished journals/conferences, within the last two years is eligible for nomination.

Nomination Procedure

Paper nominations have to be emailed to R-Letter Editorial Board Directors:

Irene Cheng (locheng@ualberta.ca),
Weiyi Zhang (maxzhang@research.att.com), and
Christian Timmerer
(christian.timmerer@itec.aauat)

The nomination should include the complete reference of the paper, author information, a brief supporting statement (maximum one page) highlighting the contribution, the nominator information, and an electronic copy of the paper when possible.

Review Process

Each nominated paper will be reviewed by members of the IEEE MMTC Review Board. To avoid potential conflict of interest, nominated papers co-authored by a Review Board member will be reviewed by guest editors external to the Board. The reviewers' names will be kept confidential. If two reviewers agree that the paper is of R-letter quality, a board editor will be assigned to complete the review letter (partially based on the nomination supporting document) for publication. The review result will be final (no multiple nomination of the same paper). Nominators external to the board will be acknowledged in the review letter.

R-Letter Best Paper Award

Accepted papers in the R-Letter are eligible for the Best Paper Award competition if they meet the election criteria (set by the MMTC Award Board).

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

MMTC examines systems, applications, services and techniques in which two or more media are used in the same session. These media include, but are not restricted to, voice, video, image, music, data, and executable code. The scope of the committee includes conversational, presentational, and transactional applications and the underlying networking systems to support them.