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

MMTC Communications – Review

IEEE COMMUNICATIONS SOCIETY

Vol. 10, No. 1, February 2019

TABLE OF CONTENTS

Message from the Review Board Directors

Welcome to the February 2019 issue of the IEEE ComSoc MMTC Communications – Review.

This issue comprises three reviews that cover multiple facets of multimedia communication research including depth map up-sampling, parallel view synthesis, task scheduling and resource allocation. These reviews are briefly introduced below.

The first paper is published in IEEE Transactions on Mutimedia and edited by Dr. Carl James Debono. It proposes a solution that generates high-resolution depth maps having clean boundaries even when a large up-sampling rate is applied.

The second paper is published in IEEE Transactions on Multimedia and edited by Dr. Marek Domański. It designs an algorithm for balancing the view synthesis workload on multicore platforms. The algorithm is able to predict the workload by the data available for previous video frames.

The third paper, published in IEEE Transactions on Wireless Communications and edited by Dr. Qin Wang, investigates how execution efficiency and computing capability at BBU as well as delay constraint of tasks can affect the network power minimization problem in C-RANs.

All the authors, nominators, reviewers, editors, and others who contribute to the release of this issue deserve appreciation with thanks.

IEEE ComSoc MMTC Communications – Review Directors

Qing Yang
University of North Texas, USA
Email: qing.yang@unt.edu

Roger Zimmermann
National University of Singapore, Singapore
Email: rogerz@comp.nus.edu.sg

Wei Wang
San Diego State University, USA
Email: wwang@mail.sdsu.edu

Zhou Su
Shanghai University, China
Email: zhousu@ieee.org

Obtaining Accurate High-Resolution Depth Maps from Low-Resolution Ones

A short review for "A Novel Segmentation based Depth Map Up-Sampling"

Edited by Carl James Debono

Y. Qiao, L. Jiao, S. Yang and B. Hou, "A Novel Segmentation based Depth Map Up-Sampling," IEEE Transactions on Multimedia, vol. 21, vo. 1, Jan. 2019.

To provide high quality 3D immersive applications, such as free-viewpoint television (FTV) and 3D games, there is the need for high-resolution and accurate depth maps. These high resolution depth maps can be acquired using passive methods where they can be found through geometry from a multi-view camera setup [1]. This can produce errors when occlusions are present. Another option is to use active methods where the depth maps are captured using specific equipment like a depth sensor [2]. These sensors might have lower resolutions compared to the color imagery demanding up-sampling to be used in conjunction with the color information for the rendering of virtual views [3,4].

The original paper proposes a solution that generates high-resolution depth maps having clean boundaries even when a large up-sampling rate is applied. The method uses the support of the color image and is divided in two parts. In the first part the color image is segmented to identify the objects in the image and then depth map interpolation is applied.

The color segmentation algorithm used by the authors consists of four steps. A rough segmentation of the high-resolution color image is done using simple linear iterative clustering (SLIC) [5], dividing the segmented superpixels into a number of connected regions. A splitting threshold is then used to determine whether the connected regions obtained are correct. The ones that are determined to be incorrect are split into subregions using a depth guided region-growing strategy. The regions that have no depth seed are merged with their nearest neighbors. This ensures that a seed pixel exists in every independent region. Finally the adjacent regions that have similar depth values are combined together forming larger regions. The last process is repeated until no further region can be merged.

The target pixels are interpolated using the neighboring seeds in the region using the joint trilateral filter proposed by the authors. The joint trilateral filter is weighted according to a color value, a distance value and a region term determined through the segmentation process described above.

The main contributions of the original authors include the introduction of SLIC as a pre-segmentation stage to generate compact and uniformly distributed superpixels. The depth is used to guide the region growing adaptively to split incorrectly clustered regions that have large depth discrepancies. This is followed by region merging to combine the neighboring segments that have similar depths. This retains the depth edges. The developed joint trilateral filter is used to interpolate the target pixels.

The solution was evaluated using depth maps with no missing depth information and others that have missing structural parts and random missing depth values as captured by typical time-of-flight sensors. The assessment was done using the bad pixel rate and the root-mean-square error (RMSE). The bad pixels in the up-sampled image are defined by the authors as those that have a depth value which is different by more than one disparity value from the ground truth. The authors evaluate the results in different regions. These regions are the entire image, the disocclusion regions, where no depth date is available, and the depth discontinuous regions. Depth discontinuous regions are found using a sliding window technique and a threshold. It occurs when the depth gap inside the window is larger than a pre-defined threshold. This gives a measure of the performance at the boundary of different depth areas. The results reported by the authors show good performance both in subjective and objective evaluations.

The quality of depth maps is very important to achieve pleasant 3D viewing. Generation of virtual views depends on this data especially on the accuracy of the edges. Moreover, transmission of multi-view data demands a large bandwidth which can be reduced using the multi-view plus depth representation. Further reductions in bandwidth requirements can be achieved by down-sampling the depth maps and then up-sampling them at the receiver [6]. Thus, such up-sampling algorithms allow down-sampling with a larger factor giving more gains in compression efficiency.

References:

[1] K. Yoon and I. Kweon, "Adaptive support-weight approach for correspondence search," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 650–656, Apr. 2006.

[2] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (TOF) cameras: A survey," *IEEE Sensors Journal*, vol. 11, no. 9, pp. 1917–1926, Sep. 2011.

[3] J. Park, H. Kim, Y.-W. Tai, M. S. Brown, and I. Kweon, "High quality depth map upsampling for 3D-TOF cameras," in *Proceedings of the International Conference on Computer Vision*, Nov. 2011, pp. 1623–1630.

[4] O. Choi and S. W. Jung, "A consensus-driven approach for structure and texture aware depth map up-sampling," *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3321–3335, Aug. 2014.

[5] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC Superpixels," *EPFL Technical Report 149300*, June 2010.

[6] L. Zammit, C.J. Debono, "Improved reconstruction of downsampled MV-HEVC depth video," in *Proceedings of IEEE EUROCON 2017*, July 2017, pp. 176–180.

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 Head of the Department of Communications and Computer Engineering 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 was the IEEE Region 8 Vice-Chair of Technical Activities between 2013 and 2014. He has served on various technical program committees of international conferences and as a reviewer in journals and conferences. His research interests are in multi-view video coding, resilient multimedia transmission, computer vision, and modeling of communication systems.

Free Viewpoint Television: Parallelization of View Synthesis on Multicore Platforms

A short review for "Depth Assisted Adaptive Workload Balancing for Parallel View Synthesis"

Edited by Marek Domański

Xin Jin , Zhanqi Liu , Qian Li , Qionghai Dai, "Depth Assisted Adaptive Workload Balancing for Parallel View Synthesis," IEEE Transactions on Multimedia, vol. 20, no. 11, Nov. 2018.

Virtual view synthesis is aimed at computer generation of video that would be visible from a virtual viewpoint, i.e. video that would be acquired from this viewpoint if a video camera was installed there. For virtual view synthesis, the input data are video streams from real cameras installed in the positions in the neighborhood of the virtual camera, the depth maps and camera parameters. Virtual view synthesis has become an essential technique for virtual navigation, free-viewpoint television, autostereoscopic 3D displays and omnidirectional video.

In the paper, widely-used efficient depth-image-based rendering (DIBR) [1]-[3] is considered as the basic view synthesis technique that provides high-quality virtual views, but also demands significant computational effort, especially for high-resolution video and real-time applications. Therefore, the paper is aimed at parallelization of DIBR. A widely used implementation of DIBR is considered in the paper and used as the reference, i.e. the method implemented in View Synthesis Reference Software (VSRS), and promoted by MPEG until the end of 2018. A drawback of these considerations is related to the use of an outdated version of VSRS [4].

The efficiency of multicore parallel systems is strongly constrained by the core with the heaviest workload. In order to reduce this workload, workload balancing algorithms are used. Such algorithms are designed to predict the workload and to allocate the workload among the cores evenly. The workload for a frame is predicted based on the number of hole pixels in the previous frame, and the numbers of hole pixels are estimated from the properties of the respective depth maps. The estimated workload is allocated among the cores that process parts of the frame that are related to similar workload. Moreover, the synchronization delay in the multicore system has to be kept low.

This paper extends the authors' results from [5], where a high-efficiency workload prediction model was proposed based on the relationship between the amount of holes in the warped images and the depth differences between the neighboring frames. The abovementioned relationship between the number of hole pixels and the summation of depth value is analyzed in detail in the paper reviewed. Cost function relying on recursive prediction and refinement is proposed to partition the current frame dynamically to balance the synthesis workload among the cores.

The effectiveness of the proposed workload prediction is demonstrated experimentally for 11 multiview test sequences from MPEG and JCT-3V. The performance is analyzed in terms of the accuracy of workload prediction, speedup ratio in parallelism, the performance of workload balancing, the objective and subjective view synthesis quality. The methods from [6] and [7] are used for comparisons that are extensively discussed in the paper using the experimental results for the proposed method as well as for the approaches from [6] and [7]. Parallelism and workload balancing performance is demonstrated both for multi-thread PC and embedded platform with 8 processor cores. Quality of the virtual views is compared using both objective quality assessment metrics and subjective tests. Two objective quality assessment metrics are used: multi-scale structural similarity (MS-SSIM) [8] and 3D synthesized view image quality metric (3DSwIM) [9]. Summarizing the objective and subjective testing results together, the authors

claim that the proposed algorithm can efficiently improve the parallelism of view synthesis without sacrificing the visual quality.
In summary, this work provides an algorithm for balancing the view synthesis workload on multicore platforms. For view synthesis, the proposed algorithm is able to predict the workload from the data available for previous video frames.

References:

[1] L. McMillan, "An image-based approach to three-dimensional computer graphics," *Ph.D. dissertation*,Dept. Comput. Sci.,Univ. of NorthCarolina, Chapel Hill, CA, USA, 1997.

[2] C. Fehn, "Depth-image-based rendering (DIBR), compression and transmission for a new approach on 3D-TV," in *Proc. SPIE*, May 2004, vol. 5291, pp. 93–104.

[3] Y. Mori, N. Fukushima, T. Yendo, T. Fujii, and M. Tanimoto, ''View generation with 3D warping using depth information for FTV," *Signal Process, Image Commun.*, vol. 24, no. 1, pp. 265–272, 2009.

[4] View Synthesis Reference Software 3.5. [Online]. Available: http://wg11.sc29.org/svn/repos/MPEG-4/test/trunk/3D/view_synthesis/VSRS.

[5] Z. Liu, X. Jin, and Q. Dai, "Motion and depth assisted workload prediction for parallel view synthesis," in *Proc. Pacific-Rim Conf. Multimedia*, Korea, Sep. 16–19, 2015, LNCS 9315, pp. 3–13.

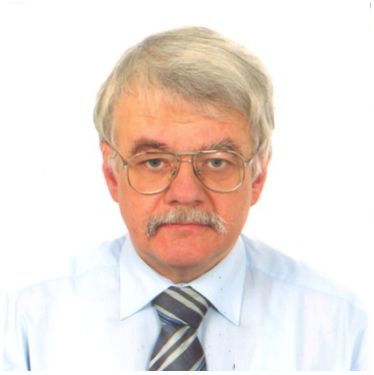
[6] K. Govil, E. Chan, and H. Wasserman, "Comparing algorithm for dynamic speed-setting of a low-power CPU," in *Proc. 1st Annu. Int. Conf. Mobile Comput. Netw. ACM*, 1995, pp. 13–25.

[7] Z. Liu, X. Jin, and Q. Dai, "Region adaptive workload prediction for parallel view synthesis," in *Proc. IEEE Visual Commun. Image Process*., 2015, pp. 1–4.

[8] W. Zhou, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. IEEE 37th Asilomar Conf.Signals, Syst. Comput.*, 2003, vol. 2, pp. 1398–1402.

[9] F. Battisti, E. Bosc,M. Carli, P. L. Callet, and S. Perugia "Objective image quality assessment of 3 D synthesized views," *Signal Process., Image Commun.*, vol. 30, pp. 78–88, 2015.

.



Marek Domański received M.S. (1978), Ph.D. (1983) and Habilitation (1990) degrees from Poznań University of Technology, Poland, where he is a Professor (since 1993) and leads the Chair (Department) of Multimedia Telecommunications and Microelectronics. He co-authored highly ranked technology proposals submitted in response to MPEG calls for scalable video compression (2004) and 3D video coding (2011). He also leaded the team that developed one of the very first AVC decoders for TV set-top boxes (2004) and various AVC, HEVC and AAC HE codec implementations and improvements. He is an author of 3 books, 14 patents and over 300 research papers, mostly on image, video and audio compression, virtual navigation, free-viewpoint television, image processing, multimedia systems, 3D video and color image technology, digital filters and multidimensional signal processing. He served as a member of various steering, program and editorial committees of international journals and international conferences.

Joint Task Scheduling and Resource Allocation in Downlink C-RAN

A short review for "Power Minimization Based Joint Task Scheduling and Resource Allocation in Downlink C-RAN"

Edited by Qin Wang

W. Xia, J. Zhang, T. Q.S. Quek, S. Jin and H. Zhu, "Power Minimization Based Joint Task Scheduling and Resource Allocation in Downlink C-RAN," IEEE Transactions on Wireless Communications, vol. 17, no. 11, Nov. 2018.

As the dramatic increase in the quantity of subscribers and the number of devices, the evolution of information and communication technology is causing energy consumption levels to reach a distressing rate. The massive connectivity also leads to tremendous carbon dioxide emissions into the environment [1]. To reduce energy consumption, a new system architecture called cloud radio access networks (C-RANs) is proposed where conventional base stations are replaced with low-cost remote radio heads (RRHs) and these RRHs are deployed close to user equipment terminals (UEs), so the transmission power is significantly reduced. Furthermore, computational resource is aggregated into a central baseband unit (BBU) pool to improve hardware unitization and centralized signal processing can achieve cooperation gain.

However, new challenges also arise in C-RANs. The dense deployment of RRHs in C-RANs leads to the increase of power consumption. In addition, the increased traffic causes a heavy burden on fronthaul in terms of capacity demand and power consumption. Finally, the power consumption of baseband processing is also considerable, which is determined by the allocation of computational resource. Overall, all the three challenges have a great effect on the network power consumption in C-RANs [2].

This paper aims to minimize the network power consumption under delay constraint where the aforementioned three challenges are considered simultaneously. Specifically, a downlink C-RAN is considered, which is composed of many RRHs which are connected to a BBU pool via fronthaul. In the BBU pool, there is a data center with a set of physical servers. Each UE has one task which is first scheduled on a certain server and a virtual machine (VM) is created by the server to execute this task. Then, the output data is transmitted using RRHs via fronthaul to the UEs. Due to limited fronthaul capacity, the precoded signals are first compressed and then the corresponding compression descriptions are forwarded through the fronthaul [3]. In this paper, the authors formulate a joint network power minimization problem of task scheduling and resource allocation, which includes not only computational resource allocation but also power allocation for transmission. Note that the power minimization problem for transmission is a fast time-scale issue because it depends on small-scale fading which varies in the order of milliseconds. However, the power consumption problem for computation is a slow time-scale issue since the task scheduling and computation resource allocation are usually executed much slower than milliseconds. Therefore, the joint network power minimization problem is a mixed time-scale issue.

Thus, the authors' major contribution is to translate the fast/mixed time-scale problem into a slow time-scale one. Different from reference [4], where the sample averaging was used to approximate the time averaging of the power consumption of transmission, the authors introduce the large system analysis to convert our problem into one that only depends on statistical channel information (i.e., large-scale fading) instead of small-scale fading [5]. Therefore, the power minimization problem for transmission, as well as the joint network power minimization problem, is turned into a slow time-scale one.

For the power minimization problem for computation, the authors propose a bound improving branch and bound (BnB) algorithm to determine the optimal solutions [6]. To reduce the computational complexity and time, the authors also propose a suboptimal combinational algorithm. For the power minimization problem for transmission, an iterative coordinate descent algorithm is proposed to determine solutions. Finally, a distributed algorithm based on

hierarchical decomposition is proposed to solve the joint network power minimization problem.

Massive experiments demonstrate the performance of the proposed solutions. Simulation results show that for the power minimization problem for computation, the combinational algorithm achieves the suboptimal solutions with much less computational complexity and time, compared to the BnB algorithm. In addition, as the delay constraint increased, suggesting the decrease of the QoS demand, the joint network power consumption is also reduced.

In summary, this work provides a framework to investigate how execution efficiency and computing capability at BBU as well as delay constraint of tasks can affect the network power minimization problem in C-RANs.

References:

[1] T. Q. S. Quek, M. Peng, O. Simeone, and W. Yu, Cloud Radio Access Networks: Principles, Technologies, and Applications. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[2] S. Luo, R. Zhang, and T. J. Lim, "Downlink and uplink energy minimization through user association and beamforming in C-RAN," IEEE Trans. Wireless Commun., vol. 14, no. 1, pp. 494–508, Jan. 2015.

[3] Y. Zhou and W. Yu, "Optimized backhaul compression for uplink cloud radio access network," IEEE J. Sel. Areas Commun., vol. 32, no. 6, pp. 1295–1307, Jun. 2014.

[4] J. Tang, T. Q. S. Quek, C. Tsung-Hui, and S. Byonghyo, "Systematic resource allocation in cloud RAN with caching as a service under two time-scale," IEEE J. Sel. Areas Commun., submitted.

[5] W. Xia, J. Zhang, S. Jin, C. K. Wen, F. Gao, and H. Zhu, "Large system analysis of resource allocation in heterogeneous networks with wireless backhaul," IEEE Trans. Commun., vol. 65, no. 11, pp. 5040–5053, Nov. 2017.

[6] Z. Yu, K. Wang, L. Chen, and H. Ji, "Joint multiuser downlink beamforming and admission control for green cloud-RANs with limited fronthaul based on mixed integer semi-definite program," in arXiv preprint arXiv:1610.04851, Oct. 2016.

Qin Wang, Ph.D, is an assistant professor with the College of Engineering and Computing Sciences at New York Institute of Technology. She received B.S. and Ph.D degrees in Communication Engineering from Nanjing University of Posts and Telecommunications, China, in 2011 and 2016, respectively. She was a visiting Ph.D. student in the Department of Computer Science, San Diego State University. Her research interests include multimedia pricing, network resource allocation, and Internet of Things. She has published papers in prestigious journals such as IEEE Transactions on Vehicular Technology and IEEE Communications Magazine, in prestigious conferences such as IEEE INFOCOM SDP Workshop.

Paper Nomination Policy

Following the direction of MMTC, the Communications – Review platform aims at providing research exchange, which includes examining systems, applications, services and techniques where multiple media are used to deliver results. Multimedia includes, but is 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 Review Board Directors: Qing Yang (qing.yang@unt.edu), Roger Zimmermann (rogerz@comp.nus.edu.sg), Wei Wang (wwang@mail.sdsu.edu), and Zhou Su (zhousu@ieee.org). 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

Members of the IEEE MMTC Review Board will review each nominated paper. In order to avoid potential conflict of interest, guest editors external to the Board will review nominated papers co-authored by a Review Board member. The reviewers' names will be kept confidential. If two reviewers agree that the paper is of Review quality, a board editor will be assigned to complete the review (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.

Best Paper Award

Accepted papers in the Communications – Review 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://mmc.committees.comsoc.org/.

MMTC Communications – Review Editorial Board

Multimedia Communications Technical Committee Officers

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.