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

E-LETTER

Vol. 9, No. 1, January 2014

IEEE COMMUNICATIONS SOCIETY

CONTENTS

Message from MMTC Chair

Dear MMTC fellow members,

Welcome to the January issue of 2014 IEEE MMTC E-letter. Time really flies and it has been another fruitful year for MMTC. Take this occasion, I would like to thank the E-letter and R-letter editorial teams as well as chairs and vice–chairs of special interest groups (IG), and all contributors over the past year. Their contribution and dedication have made both E-letter and R-letter valuable platforms for researchers in the multimedia communications society to share and exchange ideas.

In order to recognize outstanding contributions of MMTC E-letter and R-letter editors, we have established the MMTC Excellent Editor Award. Based on the recommendations of E-letter and R-letter directors/co-directors, we have chosen two award winners this year. It is with my great pleasure to announce the awardees for year 2013:

- Dr. Koichi Adachi (Institute for Infocomm Research, A*STAR, Singapore), R-letter Editor
- Dr. Zhu Liu (AT&T Labs, USA), E-letter Editor

Congratulations Dr. Adachi and Dr. Liu, and thank you for your great contributions.

As you may already know, MMTC has been actively sponsoring conferences and workshops in the area of multimedia communications via providing technical support to these conferences. IEEE GLOBECOM 2014 will be held in Austin, Texas. The Symposium of Communications Software, Services and Multimedia Application (CSSMA) is fully endorsed by IEEE MMTC. Representing MMTC, I will serve as the symposium co-chair of CSSMA. I invite all MMTC members to contribute to the symposium via either joining the technical program committee or serving as reviewers. If you are interested in it, please kindly send me an email (ky2123@caa.columbia.edu).

As always, I'd like to thank all IG chairs, board directors for the work you have done for IEEE MMTC. I wish all of you will continue your strong support to our prosperous community in 2014.

Kai Yang

Vice-Chair of Multimedia Communication TC of IEEE ComSoc

SPECIAL ISSUE ON MULTIMEDIA COMMUNICATIONS AND 3D TECHNOLOGY

Guest Editor: Fen Hou

Department of Electrical and Computer Engineering, University of Macau, Macau, China

fenhou@umac.mo

With the advances in the telecommunications and coding technique, the multimedia service is becoming the killing application. In addition to the 2D multimedia application, 3D techniques and applications such as 3DTV, 3D games, 3D printing, 3D cinema and 3D museum attract more and more attentions at both industry and academia. However, how to provide efficient 3D services, especially mobile 3D services, is still of challenge.

This special issue of E-Letter focuses on the different research areas in 3D multimedia service provision including high-efficient video coding, signal process, immersive sensory experience and robust mobile 3D video communications. It is the great honor of the editorial team to have six leading research groups to share their latest research work and solutions to these challenges.

In the first article titled, "*Toward Internet 3D Visual Media Era: a Brief Overview",* Jianfei Cai from Nanyang Technological University in Singapore presented a high-level overview on various types of 3D visual media. The author classified 3D visual data into two categories: synthesized 3D visual data and reconstructed 3D visual data, and discussed the 3D technology from three aspects: 3D content creation, 3D content representation, and 3D content delivery.

In the second article, "*Distributed Video Coding-based Robust Mobile Multi-view 3D Video Communication*", Wei Xiang and Pan Gao from the University of Southern Queensland in Australia proposed error-resilient mobile multi-view 3D video communication system, where computation intensive temporal and inter-view correlation exploration is shifted from the sender side to the receiver side. With the proposed method, the encoder and decoder do not have to have deterministic predictors and the distributed 3D video coding method can provide inherent robust and flexibility. With the low encoding complexity in encoding and the distributed feature, the proposed method is highly suitable for wireless 3D video communications.

The third article is contributed by Shiming Ge *et al.* from State Key Laboratory of Information Security, Institute of Information Engineering at Chinese Academic of Science, and the title is *"Image Missing Block Recovery with Edge Adaptive Projection Method"*. An adaptive projection method is proposed to achieve the efficient recovery of missing image blocks. The traditional error control methods such as Forward Error Correction (FEC) and Automatic Repeat Query (ARQ) need the transmission of additional data and may lead to extra delay. The proposed method overcomes the overhead operations and delay by using the adaptive combination of the frequency and spatial domain information to recover the missing image blocks. Experimental evaluation shows the better recovery quality and robustness comparing with some existing methods.

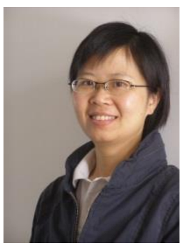
In the fourth paper, *"Compressive Sensing Based Image Processing",* Chanzi Liu, Qingchun Chen and Hongbin Liang presented compressive sensing, which is a useful tool in image processing due to the sparse structure characteristics of most images and signals. The paper introduced the main framework of the compressive sensing and its applications in image processing such as sampling, image de-noising, in-painting, super-resolution reconstruction and image fusion.

The fifth article is *"3D Extensions for High-Efficient Video Coding"*, from Karsten Muller at the Fraunhofer Institute for Telecommunications in Berlin. The authors introduced two extensions for the High-efficiency video coding (HEVC) to offer compression of 3D formats to support different displays such as 3D cinema, 3D home entertainment and mobile devices. The first 3D extension is multi-view HEVC, which provides a simple and easy extension of 2D coding to support stereoscopic video-only applications. The second extension is 3D-HEVC, which can provide the high compression efficiency for multi-view video.

The last article of this special issue is from Christian Timmerer *et al.* at Alpen-Adria-Universität Klagenfurt in Austria and the title is *"Enhancing 3D Video to Enable a Fully Immersive Sensory Experience"*. This article provided discussion on sensory effects such as ambient light, wind, vibration, scent and water spaying.

The paper presented the concept and system architecture of receiving sensory effects in addition to audio and visual contents. The provisioning of these sensory effects can produce a more realistic, immersive user experiences and the visual quality of 3D video can be highly enhanced. The Quality of Sensory Experience (QoE) and the Quality of Sensory Experience (QusSE) are further defined and discussed in the paper.

The six articles cover different aspects of Multimedia and 3D technique and we hope they can inspire the audiences with more research and interests in this area. Finally, we would like to thank all the authors for their great contribution and the E-Letter Board for making this special issue possible.



Fen Hou has been an assistant professor in the Department of Electrical and Computer Engineering at the University of Macau since 2013. Dr. Fen Houreceived the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Canada, in 2008. She worked as a postdoctoral fellow in the Electrical and Computer Engineering at the University of Waterloo and in the Department of Information Engineering at the Chinese University of Hong Kong from 2008 to 2009 and from 2009 to 2011, respectively. She worked as a lecture at the Macao Polytechnic Institute from 2011 to 2012. Her research interests include resource allocation and scheduling in broadband wireless networks, protocol design and QoS provisioning in cognitive radio network, and mechanism design in participatory sensor networks. Dr. Hou is the recipient of IEEE GLOBECOM Best Paper Award in 2010, as well as the Distinguished Service Award in IEEE ComSoc Multimedia Communications Technical Committee (MMTC) in 2011.

Dr. Hou has served as the chair of an Interest Group in IEEE MMTC, as well as a technical program committee member for IEEE WiOpt 2012, IEEE ICC 2011, IEEE WCNC 2011, IEEE GLOBECOM 2010, etc. Dr. Hou has also served as a technical reviewer for IEEE Transactions on Wireless Communications, IEEE Transactions on Vehicular Technology, IEEE Journal on Selected Areas in Communications, IEEE INFOCOM, etc.

Image Missing Block Recovery with Edge Adaptive Projection Method

Shiming Ge, Limin Sun, Shuixian Chen, and Hui Wen,
State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academic of Science, Beijing, China
{geshiming, sunlimin, chenshuixian, wenhui}@iie.ac.cn

1. Introduction

In Discrete Cosine Transformation (DCT)-coded image transmission, channel interferences often result in the data missing of image blocks. To recover the lost data in coded images becomes even more important in wireless video communications. Traditional error control strategies applied for wireless data transmissions, such as Forward Error Correction (FEC) and Automatic Repeat Query (ARQ), are reported to have some practical limitations [1]. These schemes need to transmit additional data and may cause excessive delays. As such, error concealment, which takes advantage of the spatial or temporal correlation of images and videos in post-process stage, is proposed to overcome the overhead operations and delay. The task of error concealment is to estimate the original data from the damaged data that received, without the knowledge of the noise or interference model and the degradation level of transmission channel. Therefore, the recovery of missing data should be based on the available image information and channel prior effectively. The objective is to make the restored video image natural for human visual perception, such as structure continuity, light consistency, texture or detail clarity, and so on.

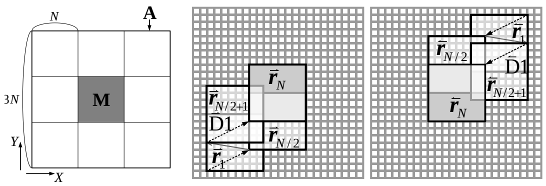
Multiple spatial interpolation methods have been proposed in the past to recover the missing blocks, including Hsia's edge-oriented weighted linear interpolation method [2], Li's best neighborhood matching approach [3], and so on. A typical problem of existing proposals is that they lose the high frequency information of the video images, such as edge or complex texture. To overcome this issue, Hirani and Totsuka proposed to combine the frequency and spatial domain information for interactive recovery [4]. Rane etc. proposed to first decompose the image into structure and texture components and then to perform filling-in of missing blocks with spatial interpolation and texture synthesis respectively [5]. The method, however, requires high computation cost to derive the solution of partial differential equations.

Another popular approach of missing block recovery is by approximating the original data with robust estimation [6-9]. Such proposals usually represent the original image by a high dimension vector including available pixels and missing pixels, and then form an estimation of original image under some constraint conditions through finding an orthonormal transformation. Under energy preservation constraints, Alkachouh and Bellanger utilized DCT based interpolation to estimate the missing blocks [6]. Under sparsity constraints, Guleryuz presented nonlinear approximation method in [7] to recover the missing pixels according to available pixels. In [7], the orthonormal transformation from DCT, Wavelet and Wavelet Package is applied adaptively. Projections onto convex sets (POCS) based methods [8, 9] can be viewed as a generalized estimation-based recovery method. This method can employ either standard orthonormal transformations or nonlinear transformations. In [9], Park etc. introduced alterative projection recovery method called RIBMAP, which attained impressive improvement compared with other recovery methods.

In this paper, we propose an adaptive projection method for the recovery of missing image blocks. Specifically, in our proposal the missing block is classified through edge orientation information of the surrounding neighborhood. We then apply the POCS algorithm [8, 9] to recover the missing image blocks by adaptively combining the frequency and spatial domain information according to the class. Simulation results using adaptive projections demonstrate that our proposed method can restore image edges and complex textures faithfully. Simulation comparison with RIBMAP shows the superiority of our proposed method in restored image quality and robustness.

2. Recovery of missing blocks with projections

Fig.1(a) illustrates image condition for missing pixel recovery, which is similar to [9]. A NxN (N=2^n) missing block **M** with its surrounding neighborhood **A** is shown. The recovery of the missing block is by interpolating the missing pixels according to its neighborhood pixels. Our recovery method classifies **M** by first estimating the edge orientation across it according to the pixels in **A**. We then define a bidirectional pair of slicing recovery windows with the estimated edge orientation, and perform missing pixel recovery in a progressive manner.



(a) (b)

Fig.1. (a) Missing block **M** and its surrounding neighborhood **A.** (b) A bidirectional pair of recovery windows in orientation D1.

Suppose that the orientation of edges in the adjacent surrounding neighborhood **A** expand the structure into the missing block **M**. The structure in the missing block is dictated by the orientation of edges in the surrounding pixels. To avoid the instability in directly performing the edge detection, edge orientation is calculated by voting the gradient magnitudes of all surrounding pixels to one of K quantized gradient orientations {Dk: k=0,1,…,K-1}. The votes are accumulated and edge orientation of the missing block is determined as the orientation with maximal votes. According to the edge orientation, the missing block is classified into one of K classes. In this paper, we set $K = 8$. The quantization step of gradient orientation is $180^o/K$. As such, the slicing manner of a bidirectional pair of recovery windows can be determined. Fig 1(b) shows an example of a slicing recovery window pair in direction D1 where the arrows represent the moving directions. Each window including correct pixels and the estimated missing pixels are represented by a NxN-dimensional recovery vector $\boldsymbol{r}$. When recovery window moves across the whole missing block along edge orientation, the recovery vector is updated progressively and the missing pixels in the vector are restored.

In addition, the surrounding neighborhood **A** can form a group of windows which can be converted into a set of sample vectors with the same size as recovery vectors. The total number of sample vectors is $8N$. The sample vectors are indicted with $\mathbf{S}=\{\mathbf{s}_j{:}j=0,1,\ldots,8N\text{-}1\}$. Our recovery method uses the sample vectors to infer the missing pixels in the recovery vectors in a POCS framework.

A recovery vector is represented as an element of Hilbert space $\boldsymbol{f}$. Given the specified constraint convex sets $\{\mathbf{C}_i\}$ and the related projection operator $\{\boldsymbol{P}_j\}$, the solution can be found by p iterative projections with $\boldsymbol{f}^{(t+1)}=\boldsymbol{P}_p\ldots\boldsymbol{P}_1\boldsymbol{f}^{(t)}$ where t is iteration step index. Four projections are defined in our recovery method as follows.

Projection $\boldsymbol{P}_1$ is defined in DCT domain for capturing the global information of missing blocks, which is similar to [9]. The $8N$ sample vectors form a convex hull in a NxN-dimensional space, in which each vector $\mathbf{s}_j$ is a vertex of the convex hull. First, we identify a best matching block for a recovery window from **A**. Different from [9], the block matching only uses the available pixels in recovery window. Then projection operator $\boldsymbol{P}_1$ is performed in the same manner as [9]. In addition, we use the similar gray range constraint projection $\boldsymbol{P}_2$ and adjacent pixel difference constraint projection $\boldsymbol{P}_3$. Projection $\boldsymbol{P}_3$ can be used to preserve the continuity of image structure. Projection $\boldsymbol{P}_4$ is used to force the smooth image transition across missing blocks. It is performed by weighting the recovery results of two moving directions. The convex set and projection operator are defined as:$C_4 = \{\boldsymbol{f}\!: f_p=w_p f_p^{(1)} +(1-w_p)f_p^{(2)}, p\in\mathbf{M}\}$ and $\boldsymbol{P}_4 f_p=w_p f_p^{(1)} +(1-w_p)f_p^{(2)}, p\in\mathbf{M}$. Here $f_p^{(1)}$ and $f_p^{(2)}$ are the recovery results of two slice directions, and w_p is the weight value which is computed with $w_p=\exp(-4i^2/N^2)$, where i is the slicing order and $i=0,1,\ldots,N\text{-}1$.

Missing pixels are restored by iteratively projecting onto specified convex sets. Algorithm 1 shows the process of restoring a missing block:

Algorithm 1

1) Initialization: Build sample vector group **S**; Set iterative time T; Calculate edge orientation
2) Loop

For i=1 to N

Make recovery vectors in two slicing directions and find best matching sample vectors from **S**.

For t=1 to T

Perform projection $\boldsymbol{P}_1$;

Perform projection $\boldsymbol{P}_2$;

Perform projection $\boldsymbol{P}_3$;

End

End

3) Perform projection $\boldsymbol{P}_4$

3. Performance analysis

We test our proposed scheme on several 256 gray-level images from [12], and compare the results with RIBMAP [9]. RIBMAP has been compared with several block-recovery algorithms proposed in [6, 8, 10, 11], and is proved to be one of the most effective approach for error concealment. Metrics, including the peak signal-to-noise ratio (PSNR), mean of absolute error (MAE) and variance of absolute error (VAE), are used to measure the restored image quality. PSNR reflects the power of recovery algorithm. MAE and VAE reflect the recovery robustness of recovery algorithm. Table.1 summarizes the values of PSNRs, MAEs and VAEs of restored images in the case of 8x8 pixel block. The restored results of RIBMAP are from the original paper [9]. The average PSNR of our method is about 2dB higher than RIBMAP's while the average MAE and VAE are lower than RIBMAP's. The result indicates that our method has better recovery quality and robustness than RIBMAP. Fig.2 gives a visual example of recovery 16x16 blocks. We can see that our proposed method can restore the image structure and texture more effective.

Table 1. Three measures for six images

measure / image	PSNR (dB)		MAE		VAE	
	RIBMAP	Our	RIBMAP	Our	RIBMAP	Our
Lena	34.65	36.43	5.4	4.5	64.8	43.2
Masquerade	29.87	32.74	10.0	7.5	184.3	91.7
Peppers	34.20	35.75	5.9	5.4	69.3	44.7
Boat	30.78	33.86	9.1	8.3	147.4	119.1
Elaine	34.63	36.15	6.8	6.0	47.1	31.4
Couple	31.49	32.41	7.2	7.6	143.6	101.2
Average	**32.60**	**34.56**	**7.4**	**6.6**	**109.4**	**71.9**

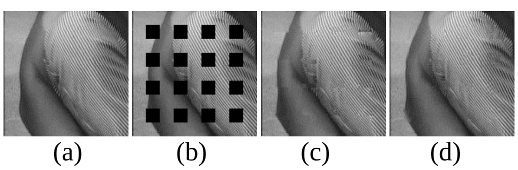


Fig.2. Experiment on a lost block size of 16×16 pixels in the "Barbara" image. (a) original image; (b) damaged image; (c) RIBMAP' recovery result; (d) our recovery result.

References

[1] Yao Wang, Qin-Fan Zhu. Error control and concealment for video communication: a review. In Proceedings of IEEE, 86(5), pages 974-997, 1998. IEEE

[2] Shih-Chang Hsia. An edge-oriented spatial interpolation for consecutive block error concealment. In IEEE Signal Processing Letters, 11(6), pages 577-580, 2004. IEEE.

[3] Wen Li, David Zhang, Zhiyong Liu, and Xiangzhen Qiao. Fast block-based image restoration employing the improved best neighborhood matching approach. In IEEE Transaction on Systems, Man and Cybernetics, 35(4), pages 546-555, 2005. IEEE.

[4] Anil N. Hirani, and Takashi Totsuka. Combining frequency and spatial domain information for fast interactive image noise removal. In Proceedings of the 23rd ACM SIGGRAPH conference on Computer graphics and interactive techniques, pages 269-276, New Orleans, USA, 1996. ACM.

[5] Shantanu D. Rane, Guillermo Sapiro, and Marcelo Bertalmio. Structure and texture filling-in of missing image blocks in wireless transmission and compression applications. In IEEE Transactions on image processing,12(3), pages 296-303,2003. IEEE.

[6] Ziad Alkachouh and Maurice G. Bellanger. Fast DCT-based spatial domain interpolation of blocks in images. In IEEE Transactions on Image Processing, 9(4), pages 729-732,2000. IEEE.

[7] Onur G. Guleryuz. Nonlinear approximation based image recovery using adaptive sparse reconstructions and iterated denoising: part I – theory. In IEEE Transactions on Image Processing, 15(3), pages 539-554, 2006. IEEE.

[8] Huifang Sun and Wilson Kwok. Concealment of damaged blocks transform coded images using projections onto convex sets. In IEEE Transactions on Image Processing, 4(4), pages 470-477, 1995. IEEE.

[9]Jiho Park, Dong-Chul Park, Robert J. Marks, and Mohamed A El-Sharkawi. Recovery of image blocks using the method of alternating projections. In IEEE Transactions on Image Processing, 14(4), pages 461-474, 2005. IEEE.

[10]Sheila S. Hemami and Teresa H.-Y. Meng. Transform coded image reconstruction exploiting interblock correlation. In IEEE Transaction on Image Processing, 4(7), pages 1023-1027, 1995. IEEE.

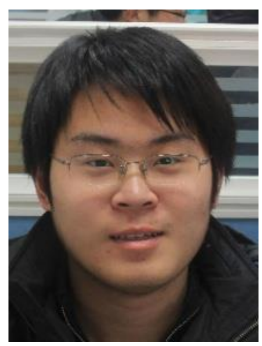
[11]Shahram Shirani, Faouzi Kossentini, Samir Kallel, and Rabab Ward. Reconstruction of baseline JPEG coded images in error prone environments. In IEEE Transaction on Image Processing 9(7), pages 1292-1299, 2000. IEEE.

[12] http://decsai.ugr.es/cvg/CG/base.htm

Shiming Ge is an Associate Professor at State Key Laboratory of Information Security (SKLOIS), Institute of Information Engineering at Chinese Academy of Sciences (IIECAS). Prior to that he was a senior researcher and project manager in ShanDa Interactive Entertainment Ltd. Former, he was a researcher in Samsung Research Center China and a research fellow in Nokia Research Center. He received Bachelor and PhD degrees in Electronic Engineering, both from the University of Science and Technology of China, in 2003 and 2008 respectively. His work has ranged in intelligent multimedia analysis, human computer interaction, image processing and computer vision.

Limin Sun received the BS, MS, and DSc degrees in College of Computer, National University of Defense Technology, in China, in 1988, 1995, and 1998, respectively. Currently, he is working as a professor in the Institute of Information Engineering at Chinese Academy of Sciences (IIECAS). He is an editor of Journal of Computer Science and Journal computer Applications (Chinese journals). He is also a guest editor of special issues in EURASIP, Journal of Wireless Communications and Networking, and Journal of Networks, respectively. His research interests include mobile vehicle networks, delay-tolerant networks, wireless sensor networks, mobile IP, and technologies for next generation of Internet. He is also a member of the IEEE and a senior member of the China Computer Federation (CCF).

Shuixian Chen is an Assistant Professor at State Key Laboratory of Information Security (SKLOIS), Institute of Information Engineering at Chinese Academy of Sciences (IIECAS). She received the PhD degree from the Wuhan University in Computer Science in 2010. Her research interests include image processing, signal analysis and pattern recognition. She has published 16 papers in international journals and conferences in last five years, and applied 16 invention patents.



Hui Wen is currently a PhD candidate at the State Key Laboratory of Information Security (SKLOIS), Institute of Information Engineering at Chinese Academy of Sciences (IIECAS). He received Bachelor degree in Computer Science from the East China University of Technology and Master degree from the Jiang'xi Normal University. His research interests include computer vision, data mining and machine learning.

Compressive Sensing Based Image Processing

Chanzi Liu, Qingchun Chen, and Hongbin Liang
Southwest Jiaotong University, P.R.China

1. Main Framework of Compressive Sensing

The Nyquist–Shannon sampling theorem maintains that a signal must be sampled at a rate at least twice its highest frequency in order to be represented without error [1]. At the same time, we often compress the data soon after sampling, trading off signal representation complexity (bits) for some tolerable error. Obviously this is wasteful of valuable sensing resources. Over the past few years, a new theory of "compressive sensing" (CS) has begun to emerge, in which the signal is sampled (and simultaneously compressed) at a greatly reduced rate. CS provides a new data acquisition approach based on signal structure [2] [3].

The fundamental principle of CS is to project a signal from a high dimensional space to another significantly smaller dimension one. To this end, the signal x must be sparse or compressible over some basis or redundant dictionaryΨ, i.e., $x = \Psi s$. In order to allow for exact or approximate reconstruction from the reduced dimensionality measurement $y = \Phi x = \Phi\Psi s$, the measurement matrix $\Phi \in \mathbb{R}^{M\times N}, M \ll N$ should satisfy the restricted isometry property (RIP) conditions[4], which requires that all submatrices composed by certain columns of the measurement matrix are well-conditioned. And the signal x can be restored by solving the following problem

$$min\|s\|_0 \;\; s.t. \;\; y = \Phi\Psi s \qquad (1)$$

The convex relaxation of l_1-minimization ($min\|s\|_1 \; s.t. y=\Phi\Psi s$ instead of $min\|s\|_0$, also known as basis pursuit) and greedy algorithms are two methods utilized to cope with the NP-hard problem in (1). Rather than minimizing an objective function globally, greedy algorithms make an approximation at each step. Typical greedy algorithms include matching pursuit (MP), orthogonal matching pursuit (OMP) [5], subspace pursuit (SP) [6], compressive sampling matching pursuit (CoSaMP)[7]. Recently, it is shown that the adaptive filter algorithm such as the least mean square (LMS) can be employed as well [8].

CS provides entirely new perspectives on fundamental principles and doctrines in signal processing, such as the sampling bounds, the choices of bases for signal representation and reconstruction. These new insights re-innovate a wide range of applications like image/video processing tasks (denoising, deblurring, inpaint-ing, compression, and superresolution), feature recognition, multimedia data mining, bio-informatic data decoding, channel estimation, spectrum sensing, and applications with sensors and computers network. One appealing feature of the CS framework is that it encourages the use of dictionaries that are adaptive to specific classes of signals or data associated with applications. Such dictionaries can be effectively derived from exemplars with sparse representation property. In many applications mentioned above, impressive performance can be achieved.

2. CS Application in Image Processing

More and more CS based image/video application are put forward and here we only highlight some important progress.

1) Sampling and Imaging
CS suggests that analog image source can be sampled by using flexible measurement matrix to significantly save the number of expensive sensor devices, thus avoiding the resources waste of the traditional "first sample then recompression". Recently, CS based sampling and imaging received a great number of research efforts. For example, it is shown that, single-pixel CS camera can be devised to collect a certain number of pixels for a complete imaging task [9]. In seismic and magnetic resonance imaging (MRI), it is expected to rely on a smaller amount of random observations to obtain high-precision reconstruction of the target signal. It is believed in that, the idea of "using computing power to exchange the expensive hardware devices" will bring us more and more CS based high-performance real-time imaging systems and devices in the near future.

2) Enhanced Image Processing
In addition to the sampling and imaging, CS based techniques also find great potential in almost every aspect of the enhanced image processing, such as image denoising, inpainting, super-resolution reconstruction and image fusion.

The basic idea of CS based image denoising lies in the distinguishability between the image and the noise and contamination. Since every compressible signal possess its peculiar sparse structure over the well-defined over-complete dictionary, the projection of the noisy image over the dictionary provides an effective method to suppress the noise and contamination by removing those sparse ingredients which are not related to the a-priori image sparse structure. The adaptive dictionary derived from patches of noisy images can provide better denoising performance. At the same time, it should be addressed that, more efforts are still needed to figure out a proper metric to assess accurately all distinguishability characteristics (including the sparse structure characteristics) between noise and image signal and its relationship with the image denoising quality.

Image can be regarded as sparse combinations of atoms of predetermined dictionaries. If the sparsity of the original image is kept in the spoiled one, hopefully the original image can be restored by solving a similar problem in (1). It is shown that superior image inpainting quality can be achieved by using non-orthogonal redundant dictionary instead of the orthogonal dictionary [10]. CS based image inpainting provides us a promising image inpainting routine. Similar to the CS-based image denoising, more efforts are still needed to assess more accurately the influence of the sparse structure characteris-tics on the achieved image inpainting performance. Hopefully, the problem such as how the amount of missing pixels impacts the quality of image recovery can be explicated as well through the above analysis.

Seeking a consistent sparse representation for each patch of a high-resolution (HR) image and its low-resolution(LR) copy over the correspond-ing HR dictionary and the LR dictionary provides an effective routine to restore a HR image from its LR copy. In fact, the super-resolution(SR) reconstruction from a single LR image is similar to solve the reconstruction problem in (1), where y denotes the LR image, $x = \Psi s$ corresponds to the HR image to be restored, Φ denotes image degraded model (such as point spread function). Although the degraded matrix Φ does not necessarily satisfy the RIP property, additional fuzzy operator matrix can be employed to improve the solvability for a better SR reconstruction quality. Compared to the joint dictionary-based SR processing, the CS-based one can save the tedious dictionary training step. The second CS-based SR techniques can be employed to save the consistent sparse decomposition of the LR image over two dictionaries. When multiple image fusion is concerned, CS-based image fusion provides another useful transform domian image fusion choice. It is expected that multiple calibrated images of the identical scenarios share almost the same sparse properties, thus those highlighted sparse structures through sparse decomposition can be utilized to enhance the image fusion with less storage requirements.

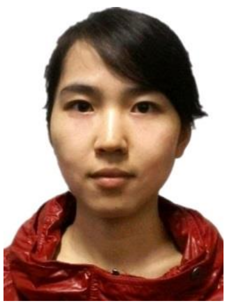
3. Conclusions

Sparse structure characteristics of most image and signal makes CS a useful tool for a broad spectrum of natural science and engineering applications. The relationship between the sparse characteristics and features of the images, the distinguishability among different signal in terms of the sparse characteristics, as well as the involved quantitative assessment method are still far from well understood. However, undoubtedly CS will play an important role in the advanced multimedia processing problem. All the CS-based technique deserve in-depth investigation in the multimedia technique society.

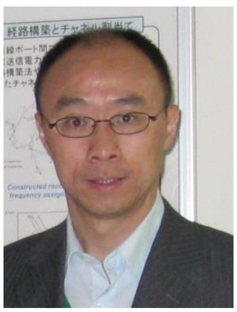
References

[1] D. Hajela. On computing the minimum distance for faster-than-Nyquist signaling. IEEE Trans. Inform. Theory, 1990, 36(2): pp. 289-295.
[2] E. J. Candès, J. K. Romberg, T. Tao. Robust uncertainty principles: Exact signal recon-struction from highly incomplete frequency information. IEEE Trans. Inform. Theory, 2006, 52(2): pp. 489-509.
[3] D. L. Donoho. Compressed sensing. IEEE Trans. Inform. Theory, 2006, 52(4):pp.1289-1306.
[4] Emmanuel J. Candes and Terence Tao. Decoding by Linear Programming. IEEE Trans. Inform. Theory, 2005, 51(12): pp. 4203-4215.
[5] J. Tropp and A. Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. IEEE Trans. Inform. Theory, vol. 53, no.12, pp. 4655-4666, Dec. 2007.
[6] W. Dai and O. Milenkovic. Subspace pursuit for compressive sensing: Closing the gap between performance and complexity. arXiv 0803. 0811v3 [CS.NA], Jan. 2009.
[7] D. Needell and J. A. Tropp. Cosamp: Iterative Signal Recovery From Incomplete and Inaccurate Samples. Appl. Comput. Harmonic Anal., vol. 26, no. 3, pp. 301-321, 2009.
[8] Jian Jin, Yuantao Gu, and Shunliang Mei. A Stochastic Gradient Approach on Compressive Sensing Signal Reconstruction Based on Adaptive Filtering Framework. IEEE J. Selected Topics in Signal Processing, Vol.4, No.2, 2010.
[9] M. F. DUARTE, M. A. DAVENPORT et al. Single-pixel imaging via compressive sampling. IEEE Signal Processing Magazine, 2008, 25(2): pp.83-91.
[10] Jianwei Ma. Compressed sensing by inverse scale space and curvelet thresholding. Applied Mathematics and Computation, 2008, 206(2): 980-

988.
[11] J.Yang, J.Wright, T.Huang, and Y.Ma. Image super-resolution as sparse representation of raw image patches. In IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2008, pp.1~8.



Chanzi Liu received her B.S. degrees in 2005, and is now pursuing her PhD degree in the key lab of information coding & transmission, Southwest Jiaotong University, China. Her research interests include: compressive sensing, signal processing technologies, image processing, mass datas processing.



Qingchen Chen (IEEE Member) received his B.S. and M.S. degrees in 1994 and 1997, respectively. He received his Ph.D. degree with honor from Southwest Jiaotong University in 2004. He is currently a professor of the key lab of information coding & transmission, Southwest Jiaotong University, China. Dr. Chen was the recipient of the NCET program in 2006 and the Fok-Ying Tung Education Foundation in 2005. Dr. Chen visited Harvard University from 2009 to 2010. Dr. Chen is the inventor of 12 patents, the author of two book chapters and over 70 technical papers. Dr. Chen's research interests include: Information theory and coding, wireless network, signal processing technologies.

Hongbin Liang (S'09) received the B.Sc. degree incommunication engineering from Beijing University of Post and Telecommunication, Beijing, China, in 1995, the M.Sc. and Ph.D. degrees both in electrical engineering from Southwest Jiaotong University, Chengdu, China, in 2001 and 2012, respectively. He is currently working in the School of Transportation and Logistics with Southwest Jiaotong University. From 2001 to 2009, he was a Software Engineer with the Motorola R&D Center of China, where he focused on system requirement analysis and Third Generation Partnership Project protocol analysis. From 2009 to 2011, he was a visiting Ph.D. student with the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada. His current research interests focus on resource allocation, quality of service, security and efficiency in cloud computing, and wireless sensor networks.

Towards Internet 3D Visual Media Era: A Brief Overview

Jianfei Cai
Nanyang Technological University, Singapore
asjfcai@ntu.edu.sg

1. Introduction

3D visual media applications such as 3DTV, 3D telepresence, virtual reality, 3D digital museum and 3D games have become more and more popular because they provide more realistic and immersive user experience, and offer more flexibilities for users to interact with the media. However, up to date, the dominate Internet visual media content are still images and videos. Despite the huge progress in computer vision and computer graphics, 3D visual media are still limited to local environment or the professional community, and it is not able to penetrate to common internet users.

There are a few reasons that prevent 3D visual media from becoming popular in Internet. First, 3D visual media is difficult to create. Unlike image and video, which can be easily generated by consumer cameras, creating 3D visual media is very challenging for common people. Second, 3D visual media is of large data volume and is often associated with interactive applications, which brings up new challenges to the bandwidth-limited and QoS-lacked Internet. Third, 3D display technology is still not so convenient and price-attractive. Most commercial 3D displays are either based on active shutter glasses or passive polarized glasses. While they are capable of rendering stereoscopic views for users, the views generated are the same regardless of the user position.

There are various types of 3D visual media. In general, 3D visual data can be classified into two categories: synthesized 3D visual data and reconstructed 3D visual data. The synthesized 3D visual media is directly created in the cyber space and it does not need to match real objects or scenes, while the reconstructed 3D visual data is created with the goal of digitizing real objects or scenes in a photorealistic way. In this brief overview, we focus on the reconstructed 3D visual media, which extends the traditional 2D visual media of images and videos. Our purpose is to review the pipeline of 3D content creation, delivery and consumption, and highlight the technologies and directions that would help move towards Internet 3D era as well as highlighting the corresponding challenges.

2. 3D Content Creation

Creating or reconstructing 3D from images or videos is not a brand new problem. It has been studied extensively in the past and still received much attention in the research community with the emergence of new sensors and new mathematic tools. Such image based 3D modeling can be generally classified into three categories: multi-view stereo (MVS) technique, photometric 3D reconstruction technique and kinect based 3D reconstruction.

Multi-view stereo (MVS) is a common technology to build 3D object model from multi-view images. Among various MVS algorithms, depth-map merging based method is a very popular one in the recent years due to its flexibility to integrate any available stereo matching algorithms. Depth-map merging based method usually consists of two steps: the first step is to calculate depth-maps from each stereo image pairs and the second step is to merge the depth-maps to generate 3D models [5]. A nice survey on MVS can be found in [11]. One state-of-the-art system is the OSM-Bundler, which is a structure-from-motion (SfM) system from unordered image collections [12]. OSM-Bundler is able to produce sparse point clouds and can be combined with PMVS2 software package [1] to generate dense point cloud.

Photometric 3D reconstruction [14] is another popular way for reconstructing 3D models from a set of images. Unlike multi-view stereo, which does not manipulate the light source but utilizes the images captured at different viewpoints, photometric algorithms typically fix the view point but vary the lighting conditions. By utilizing the known lighting variations and the observed pixel intensity changes, the normal and the shape of the object can be estimated. Recent developments such as [4] are looking into the integration of photometric and multi-view stereo approaches.

In 2010, Microsoft has launched the Kinect sensor for game applications. Kinect is equipped with an infrared camera and a RGB camera. The Infrared camera can generate the depth information easily by capturing the continuously-projected infrared structured light. With the assistance of this additional depth information, many challenging computer vision problems can now be simplified and tackled in an efficient manner. Kinect has been used in 3D reconstruction recently such as in [6, 10]. Among them, KinectFusion [10] represents the state-of-the-art technique, which is able to reconstruct 3D scene in real time. In KinectFusion, basically a kinect is hold and scanning across the scene, where the camera pose is automatically tracked and the corresponding depth data are updated in real time. It has been shown in [10] that KinectFusion is able to achieve 3D reconstruction at decent quality with fast speed but needs to be at lab environments and operated by researchers.

Despite the huge progress in image based modeling, we still have not seen practical solution for common mobile users to easily create 3D object models through capturing images or videos. Recently, we have seen some efforts moving toward this direction. For example, in [9], a user is allowed to casually capture a video by holding his common hand-held camera and moving around an object for which he wants to create a 3D model. By asking the user to interactively segment 3~4 frames, the system can extract a highly accurate and coherent silhouette of the object for the rest of the frames. With the obtained silhouettes, a rough 3D model can be easily created by using the visual hull technique.

Another issue is that most of the existing 3D reconstruction algorithms mainly focus on the reconstruction distortion, i.e. reconstructing the 3D model as close to the real object as possible. Little attention has been paid to the issue of reconstruction rate. Reconstruction models generated by the state-of-the-art multi-view stereo algorithms such as [5] are typically of huge volume, which are not suitable to be transmitted over internet. Simplification methods need to be applied as a post-processing to reduce the size of the reconstructed models so that they can become usable and sharable among Internet users. Such a fine-to-coarse processing, i.e. high quality multi-view reconstruction followed by simplification, is resource-inefficient since it over-creates numerous 3D points at the beginning and then removes many of them in the simplification stage. To solve this problem, rate-distortion concept has been introduced into the 3D reconstruction in [8].

3. 3D Content Representation

Once the 3D visual media data is created, the next question is how to represent the 3D visual information. 3D visual data can be represented in various ways including mesh, point cloud, voxels and depth maps. Different representations have different advantages and disadvantages. It is necessary to investigate the suitable 3D representation for internet users and conversions among different representations.

In general, there are two types of 3D visual media representations: geometry based 3D representation and image-based 3D representation. For geometry based 3D representation, the scene geometry is typically described by 3D polygonal meshes. Real-world objects are reproduced using geometry 3D surfaces mapped with associated attributes such as texture. The main advantages for geometry based 3D representation lie in its great flexibility in 3D rendering and the support from graphics hardware. On the other hand, image based 3D representation has the advantage in avoiding the complex 3D scene reconstruction required in geometry based 3D representation. However, image based 3D representation requires dense sampling of the real world with numerous original videos from different views, which produces huge amount of data. In addition, there are also some methods that make use of both geometry and image based techniques to combine the advantages of the two types of 3D representations [13].

With the increased flexibility and accuracy of acquiring depth information, depth based 3D representation has become more and more popular. Due to the nice features of depth-based rendering (DBR), depth based representation (V+D) has been adopted by MPEG 3DV Ad-Hoc group. We believe V+D will be the dominant 3D representation for Internet application because it is compatible with the standard image and video compression techniques and is convenient to generate virtual views at the display side. Many compression schemes for V+D representations have been proposed [7].

4. 3D Content Delivery

Delivery 2D visual information over Internet has been extensively studied. The key characteristics of 2D visual media being captured by content delivery community include unequal content importance, error-resilience, sensitivity to transmission delay, etc. Thus, most of the existing media delivery solutions are typically based on the common ideas of unequal error/loss protection, joint source-channel coding, prioritized transmission, cross-layer optimization with different customizations.

It would not have much research value by simply treating 3D visual media as one dimension increase of the traditional 2D visual media and directly applying 2D media content delivery techniques. In fact, 3D visual media has some unique characteristics that are worth to be considered for Internet applications. For example, in [2], by noticing that in many 3D applications users only observe 3D objects / scenes from a limited number of views, we propose a segmentation-based method to facilitate view-dependent 3D model transmission in order to save the bandwidth. In [3], we further consider the illumination effects into the view-dependent 3D media delivery framework, where less number of bits is spent to those over-bright or over-dark regions.

In addition to the well know characteristics such as huge data volume, 3D visual media are often associated with highly interactive applications, which bring up more stringent QoS requirements for Internet content delivery. Moreover, 3D quality of experience (QoE) is totally different from the 2D QoE, and is currently a very hot research topic.

5. Summary

In this short letter, we have briefly reviewed the pipeline of 3D content creation, representation and delivery. We are still far way from Internet 3D Visual Media Era. More research and developments need to be done. Particularly, we need to find ways to allow Internet users to easily generate 3D visual media content, efficiently store the content, easily deliver and play the 3D visual media. We hope in one day 3D media could become easily reachable by common internet users in the way that is as simple as creating, playing, sharing and delivering images & videos. Such a goal requires continuous efforts from the academia and research institutes as well as the industry.

References

[1] Y. Furukawa and J. Ponce, "Accurate, Dense, and Robust Multi-View Stereopsis," CVPR, July 2007.

[2] W. Guan, J. Cai, J. Zheng and C. W. Chen, "Segmentation based View-Dependent 3D Graphics Model Transmission", IEEE Trans. on Multimedia, special issue on multimedia applications in mobile/wireless context, vol. 10, no. 5, pp. 724-734, Aug 2008.

[3] W. Guan, J. Cai, J. Zhang and J. Zheng, "Progressive coding and illumination and view dependent transmission of 3D meshes using R-D optimization", IEEE Trans. Circuits and Systems for Video Technology, vol. 20, no. 4, pp. 575-586, Apr 2010.

[4] C. Hernández, G. Vogiatzis and R. Cipolla, "Multi-view Photometric Stereo," IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 2008.

[5] J. Li, E. Li, Y. Chen, L. Xu, and Y. Zhang, "Bundled depth-map merging for multi-view stereo," in CVPR 2010, pages 2769–2776.

[6] A. Maimone and H. Fuchs. Encumbrance-free telepresence system with real-time 3D capture and display using commodity depth cameras. In ISMAR 2011, to appear, 2011.

[7] K. Muller, P. Merkle, and T. Wiegand, "3-D Video Representation Using Depth Maps", Proceedings of the IEEE, Vol. 99, No. 4, pp. 643-656, 2011.

[8] V. Nghiem, J. Cai and J. Zheng, "Rate-distortion Optimized Progressive 3D Reconstruction from Multi-view Image", Pacific Graphics 2010.

[9] T. Nguyen, J. Cai, J. Zheng and J. Li, "Interactive object segmentation from multi-view images", accepted by Elsevier Journal of Visual Communications and Image Representation.

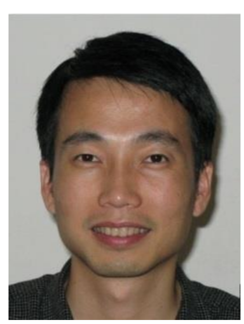
[10] R. A. Newcombe, S. Izadi, O. Hilliges, D. Molyneaux, D. Kim, A. J. Davison, P. Kohli, J. Shotton, S. Hodges, and A. Fitzgibbon, "KinectFusion: Real-Time Dense Surface Mapping and Tracking," in IEEE ISMAR, October 2011.

[11] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms," in CVPR, vol. 1, 2006.

[12] N. Snavely, S. M. Seitz, and R. Szeliski, "Photo tourism: Exploring photo collections in 3D," in SIGGRAPH, pages 835–846, 2006.

[13] A. Smolic, et al., "Free viewpoint video extraction, representation, coding, and rendering", IEEE ICIP, pp. 3287-3290, 2004.

[14] T. Wu, K. Tang, C. Tang, T. Wong, "Dense Photometric Stereo: A Markov Random Field Approach," IEEE Trans. Pattern Anal. Mach. Intell. 28(11): 1830-1846 (2006)



Jianfei Cai Jianfei Cai received his PhD degree from the University of Missouri-Columbia. Currently, he is the Head of Visual & Interactive Computing Division at the School of Computer Engineering, Nanyang Technological University, Singapore. His major research interests include visual information processing and multimedia networking. He has published over 100 technical papers in international conferences and journals. He has been actively participating in program committees of various conferences. He had served as the leading Technical Program Chair for IEEE International Conference on Multimedia & Expo (ICME) 2012 and he currently sits in the steering committee of ICME. He was an invited speaker for the first IEEE Signal Processing Society Summer School on 3D and high definition / high contrast video process systems in 2011. He is also an Associate Editor for IEEE Transactions on Circuits and Systems for Video Technology (T-CSVT), and a senior member of IEEE.

Distributed Video Coding-based Robust Mobile Multi-view 3-D Video Communications

Wei Xiang and Pan Gao
University of Southern Queensland, Australia
{wei.xiang, pan.gao}@usq.edu.au

1. Introduction

With increasing numbers of mobile phone subscribers, the usage of mobile Internet services is becoming more and more popular in our daily life. Meanwhile, mobile video streaming has becoming a natural augmentation to today's thriving Internet video streaming services due to the rapid growth of video compression and wireless communications technologies. On the other hand, as the *Avatar* experience swept the world, three-dimensional (3-D) video has attracted intensive attentions around the global. With the 3-D content provision, distribution, and visualization technologies coming to maturity, 3-D video is gradually becoming popular on mobile platforms. In wireless environments, mobile 3-D video can provide the 3-D visual experience everywhere and anytime. To realize mobile 3-D video applications, multi-view video representation is an emerging interactive video technology with stereoscopic vision. Different from conventional single-view video, multi-view video consists of a series of video clips that are captured simultaneously from a group of video cameras with different viewing angles.

The state-of-the-art standard for multi-view video coding is the MVC extension of H.264/AVC [1, 2]. The MVC standard employs motion and disparity estimation to fully explore both temporal and inter-view correlations to achieve high compression efficiency. However, the complexity of motion and disparity estimation causes a significant burden for mobile 3-D video platforms with low-complexity computation and small power consumption. Furthermore, when the compressed multi-view video signal is transmitted over an error-prone channel, it is extremely sensitive to transmission errors owing to the highly sophisticate compression techniques. If an error occurs in a frame of one view, it can propagate along both the temporal and inter-view directions, severely degrading the reconstructed video quality. This problem is exacerbated in error-prone wireless channels, where packet losses are far more frequent and bursty than in wire-line networks. Therefore, MVC is inappropriate for delivering multi-view 3-D video over mobile Internet.

Recently, distributed video coding (DVC) has received increased research attention due to its low complexity and inherent robustness in comparison to traditional predictive video coding [3]. In the following, we will briefly introduce distributed multi-view video coding, and then present a robust mobile multi-view 3-D video communications system.

2. Distributed multi-view 3-D video coding

The essential idea behind DVC is to exploit the temporal correlation of the video signal in the decoding phase rather than in the encoding one. In this way, classic motion-compensated prediction is no longer performed at the encoder, with a subsequent significant reduction in computational complexity of the encoder. The encoding rate is reduced by transmitting only the parity bits of a suitable systematic channel code. At the decoder, the redundancy of the video sequence is taken into account by performing a motion-compensated prediction based on the already received data, and the received parity bits are then used to recover as much as the original information from the side information. Therefore, DVC offers a number of potential advantages in contrast with conventional predictive coding. Firstly, it shifts the computationally intensive task of motion estimation from the encoder to its decoder counterpart. The second advantage of DVC is attributed to its intrinsic ability to cope with transmission errors since no prediction loop exists in the encoding phase.

DVC can be naturally applied to multi-view 3-D video, as it can exploit the redundancies already present in mono-view video and the inter-camera correlation in multi-view video. In a simple scheme with two views, one possibility is to encode one view using DVC and use the second view as the side information. A second possibility consists of using DVC for both views and using the implicit correlation between the two views to extract their time-dependent difference as the side information. In more general terms, the multi-view video structure can be exploited by generating the side information from the combination of inter-view correlation and time-dependent temporal correlation. Prior work on distributed compression of multi-view video has largely been focused on compression performance by removing redundancies present in overlapping camera views. A variety of tools and approaches have been proposed over the years to improve the performance of distributed multi-view video system, e.g., inter-view side information generation [4], fusion of temporal and inter-view side information [5], modeling correlation noise statistics [6], and reconstruction optimization [7], and so on. However, analyses and exploration on the error resiliency of distributed multi-view video compression systems are still at its infancy stage [8, 9].

3. Proposed error-resilient mobile multi-view 3-D video communications system

A robust mobile multi-view 3-D video communications system illustrated in Fig. 1 is proposed. Unlike standard MVC, each view is encoded separately by Wyner-Ziv (WZ) encoding at the frame level, and inter-view redundancy is exploited to generate the side information at the WZ decoder. The encoder and decoder do not have to have deterministic predictors that are identical. Therefore, this leads to inherent robustness and flexibility in the WZ based video codec, as well as low encoding complexity in encoding that are highly suitable for wireless 3-D video communications.

For the purpose of brevity, the proposed WZ-based multi-view video system is composed of three cameras, which is illustrated in Fig. 1. For the two side views, the odd and even frames are encoded using the H.264/AVC intra encoder and WZ encoding, respectively. Conversely, for the central view, the odd and even frames are WZ coded and H.264/AVC intra encoded, respectively. Therefore, the encoder works independently without the knowledge of the view signals of the other cameras. At the decoder, for the side view, the side information of each WZ frame is generated by motion-compensated temporal interpolation as the mono-view DVC scheme. However, for the central view, each WZ frame (except the first WZ frame) can be decoded using either the temporal side information generated from its temporal neighboring intra frames, or the inter-view side information generated from the intra frames of its neighboring cameras. The best predictor of different potential predictors is determined by the flexible prediction algorithm proposed in [4] according to the motion information of adjacent views.

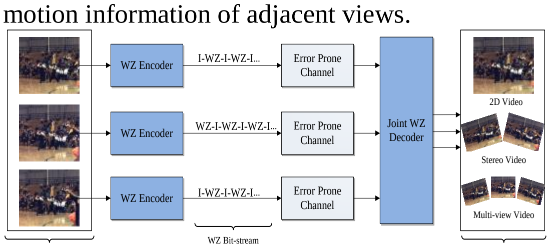


Fig. 1. Robust mobile multi-view 3-D video communications system.

In our experiments, forward error correction (FEC) by the means of Reed-Solomon codes is utilized for comparative evaluation, in which each camera is independently encoded using H.264/AVC in a simulcast manner. As can be seen from Fig. 2, the proposed robust multi-view mobile 3-D video communications framework is able to result in more visually pleasing reconstruction than the FEC approach.

(a) Reed-Solomon scheme (PSNR = 28.47 dB)

(b) Proposed system (PSNR = 30.13 dB)

Fig. 2. Visual results comparison for “Race1” sequence at the packet loss of 10%.

4. Conclusion

A multi-view distributed video coding framework has been proposed for mobile 3-D video transmission over the wireless Internet, in which computation-intensive temporal and inter-view correlation exploration is shifted from the encoder to the decoder side. Since there is no prediction loop in the encoding phase, it is more resilient to channel errors.

In future work, we will explore distributed coding of multi-view video plus depth (MVD) based 3-D video to achieve better packet error resilience. This would require further research into side information generation considering the unique characteristics of the texture video and depth map. In addition, the unequal importance of the texture video and depth map can also be considered in the distributed MVD-based 3-D video communication system.

References

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

[2] P. Merkle, A. Smolic, K. Mller, and T. Wiegand, “Efficient prediction structures for multiview video coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 11, pp. 1461-1473, Nov. 2007.

[3] L. Ciobanu and L. s Corte-Real, “Multi-view codec with low-complexity encoding for distributed video coding,” *Multimedia Tools and Applications*, vol. 64, no. 3, pp. 731-755, June 2013.

[4] X. Guo, F. Wu, D. Zhao, and W. Gao, “Wyner-Ziv-based multiview video coding,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 6, pp. 713-724, June 2008.

[5] M. Fan, F. Ye, J. Di and A. Men, “A novel fusion method in distributed multi-view video coding over wireless video sensor network,” in *Proc. IEEE Wireless Communications and Networking Conference* (WCNC), 2012, pp. 3161-3164.

[6] Y. Li, S. Ma, D. Zhao, and W. Gao, “Modeling correlation noise statistics at decoder for multi-view distributed video coding,” in *Proc. IEEE Symp. Circuits and Systems*, May 2009, pp. 2597-2600.

[7] N. Cheung, A. Ortega, and G. Cheung, “Rate distortion based reconstruction optimization in distributed source coding for interactive multiview video stream,” in *Proc. IEEE Int. Conf. Image Processing*, Sep. 2010, pp. 3721-3724.

[8] C. Guillemot, F. Pereira, L. Torres, T. Ebranhimi, R. Leonardi, and J. Ostermann, “Distributed monoview and multiview video coding,” *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 67-76, Sept. 2007.

[9] G.-M., Su, Y.-C. Lai, A. Kwasinski, H. Wang, “3-D video communication: challenges and opportunities,” *International Journal Communication Systems*, vol. 24, no. 10, pp. 1261-1281, Oct. 2011.

Wei Xiang (S’00–M’04–SM’10) received the B.Eng. and M.Eng. degrees, both in electronic engineering, from the University of Electronic Science and Technology of China, Chengdu, China, in 1997 and 2000, respectively, and the Ph.D. degree in telecommunications engineering from the University of South Australia, Adelaide, Australia, in 2004.
Since January 2004, he has been with the Faculty of Engineering and Surveying, University of Southern Queensland, Toowoomba, Australia, where he currently holds a faculty post of PPP. In 2008, he was a visiting scholar at the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. During Oct. 2010 and Mar. 2011, he was a visiting scholar at the University of Mississippi, Oxford, MS, USA. He received the Best Paper Award at 2011 IEEE WCNC. He was named a Queensland International Fellow (2010-2011) by the Queensland Government of Australia, an Endeavour Research Fellow (2012-2013) by the Commonwealth Government of Australia, and a Smart Futures Fellow (2012-2015) by the Queensland Government of Australia. His research interests are in the broad area of communications and information theory, particularly coding and signal processing for multimedia communications systems.

Pan Gao received the B.Eng. degree in computer science and technology from Sichuan Normal University, Chengdu, China, in 2009. He is currently pursuing his Ph.D. degree in the Faculty of Engineering and Surveying at the University of Southern Queensland, Toowoomba, Australia.
His research interests include multi-view video coding, distributed video coding, and error control coding.

3D Extensions for High-Efficiency Video Coding

Karsten Müller (IEEE Senior Member)
karsten.mueller@hhi.fraunhofer.de

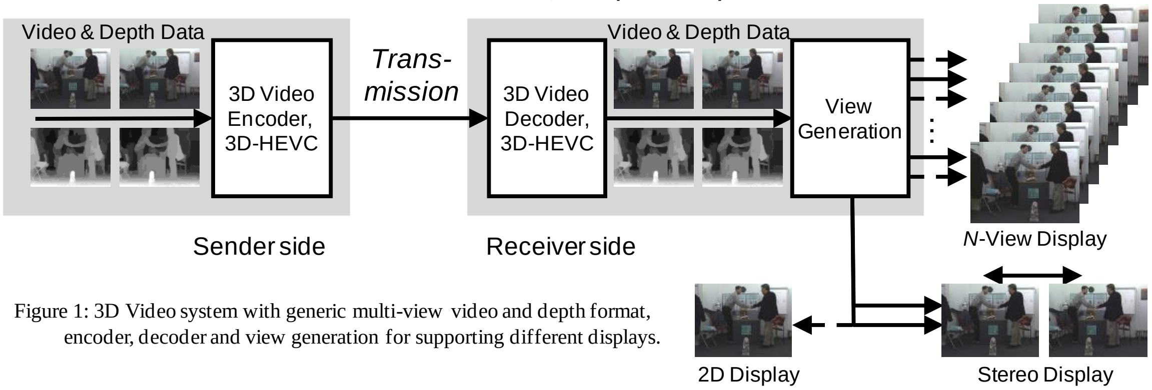


Figure 1: 3D Video system with generic multi-view video and depth format, encoder, decoder and view generation for supporting different displays.

3D applications have been commercialized in different areas, such as 3D cinema, 3D home entertainment and mobile applications. The first types of 3D displays were stereoscopic displays, which require glasses for multi-user audiences, as used in 3D cinema and home entertainment. These displays show the originally recorded two views at exactly the same position, such that no further content adaptation is required. Accordingly, a conventional stereoscopic video format (CSV) is used for compression and transmission. Newer types of 3D displays allow glasses-free 3D viewing by showing a larger number of *N* views [1]. These auto-stereoscopic multi-view displays have different number and spatial positions of views. Therefore, a more generic 3D video format, such as multi-view video and depth (MVD) is used for compression and transmission, as shown in Figure 1. At the 3D display, the required number of views is generated from the decoded MVD data through intermediate view synthesis. Besides supporting multi-view displays, the 3D video system, as shown in Figure 1, also supports different stereoscopic displays with fixed or variable base line, as well as legacy 2D displays.

For efficient 3D video transmission, coding methods have been developed for CSV and MVD formats, and are currently standardized as extensions to existing 2D video codecs to provide backward compatibility. The most recent extensions, currently under development, are introduced in the following sections.

2D Base View Coding

Each video and depth component within a CSV or MVD format is first of all compressed, applying a classical 2D video coding structure. Here, the most recent method is the high-efficiency video coding (HEVC) standard, which was officially approved in April 2013 [2]. This standard was jointly developed by the ITU-T Visual Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) under the labels “ITU-T Recommendation H.265” and “ISO/IEC 23008-2 (MPEG H Part 2)” respectively. The codec uses a number of tools for reducing the correlation with video sequences, including spatial prediction within a picture, temporal motion-compensated prediction between pictures at different time instances, transform coding of the prediction residual, and entropy coding. In HEVC, these tools are highly optimized, such that it achieves the same subjective video quality at only about 50% of the bit rate on average [3], in comparison to its predecessor (H.264/ MPEG-4 AVC High Profile).

Both CSV and MVD contain one video component that is independently coded of other components with an unaltered HEVC structure. This independent component or base view provides compatibility with the decoding process in classical 2D displays. For coding the other components in CSV and MVD, additional tools are added to HEVC, as described in the following section.

Multi-View HEVC

The first 3D extension is Multi-View HEVC (MV-HEVC), which mainly targets CSV coding and shall support stereoscopic video-only applications in a simple way. Here, only coding methods are used, which do not change the HEVC base codec below block level.

One of the most important aspects for efficiently coding multiple views is the redundancy reduction among these views at the same time instance, for which the content is usually rather similar and only varies by a slightly different viewing position. As already described, the same concepts and coding tools of the 2D base view are also used for the dependent view. In addition, disparity-compensated prediction (DCP) has been added to the HEVC design as an alternative to motion-compensated prediction (MCP). Here, MCP refers to inter-picture prediction that uses already coded pictures of the *same* view at *different* time instances, while DCP refers to inter-picture prediction that uses already coded pictures of *other* views at the *same* time instance. Here, the coding tree block syntax and decoding process of HEVC remain unchanged when adding DCP to the MV-HEVC codec. Thus, only the high-level syntax is modified so that already coded video pictures of the same access unit can be inserted into the reference picture lists.

3D-HEVC

The second 3D-extension is 3D-HEVC, which targets MVD coding with the best coding performance in order to achieve higher coding gains, e.g. for 3D applications with 4K display resolution. Here, new tools are added for coding the dependent views and depth data [4], [5]. These new tools can be clustered, according to their redundancy reduction principles:

Inter-View Prediction: Similar to the compression of dependent views in MV-HEVC, the redundancy reduction across different views is one of the most important aspects for efficient coding. In addition to disparity compensated prediction, 3D-HEVC uses further tools for inter-view prediction: The first tool is view synthesis prediction, which uses depth-based rendering to warp pixels from a reference view to a dependent view, while DCP uses one linear vector for a block. The second tool is inter-view motion parameter prediction. Also, motion vectors for the same content in the different views can be similar, such that they can be predicted across views, using again the depth/disparity information. Third, inter-view residual prediction is used. Again, also the residual data in different views is similar for a certain amount of blocks, such that prediction across views can gain coding efficiency.

Depth Intra Prediction: As 3D-HEVC also encodes depth information, special coding tools have been developed, which consider the specific characteristics of depth data, i.e. sharp edges at object boundaries and constant or slowly varying areas within objects: Accordingly, depth blocks within objects are coded by special depth modes, e.g. only coding the average value or predicting a planar function from already coded neighboring blocks without residual data. For depth blocks at object edges, special bipartition modes have been developed, which identify two regions within a block and assign each pixel to one region. The separation between the regions is signaled in different ways: On one hand, the separation is approximated by a straight line for regular cases. The line is taken from a lookup table or can be predicted from neighboring blocks. On the other hand, the separation information can be derived from a co-located texture block in cases of irregular partitioning and disconnected regions. Besides the separation information, the mean values of both partitions are also coded.

Inter-Component Prediction: Coding tools for reducing redundancies between the video and co-located depth component of each view were also developed for 3D-HEVC. One depth coding mode, that uses texture information for depth coding, has already been described above. Next, the motion parameter inheritance checks the partitioning and motion data from the texture information, whether it can be used for efficient coding of the current depth block. Also, tools for block partitioning prediction can be applied, e.g. quad-tree prediction, where subdivision information of the texture is used to restrict the subdivision of a co-located depth block. This assumes, that the texture is finer structured, than depth, such that a depth block is never subdivided further, than the texture.

Encoder Optimization: Finally, 3D-HEVC uses a special encoder control for the depth data. For video data, the classical rate-distortion optimization is used, when the optimal coding mode is sought. Here, the Lagrangian cost function is used, a weighted sum of video rate and video distortion in terms of mean squared error (MSE) between original and reconstructed video data. In contrast, reconstructed depth maps are only used for the synthesis of intermediate views and not directly viewed. This is also shown in Figure 1, where the view synthesis process for the respective 3D display follows the decoding process of an MVD format with two views. Therefore, the coding efficiency in 3D-HEVC is improved by applying a cost function with depth rate, but video distortion of synthesized intermediate views. This view synthesis optimization thus considers the special purpose of depth maps for the view synthesis, and guarantees a good quality for a dense range of views for any multi-view, as well as stereoscopic display.

Conclusions

Two extensions for HEVC were introduced, which offer compression of 3D formats in order to support different displays. The first extension, MV-HEVC, provides a simple and easy extension of 2D video coding for conventional stereo systems without changing the HEVC base codec structure. The second extension, 3D-HEVC, gives the highest compression efficiency for multi-view video and depth data. This 3D codec includes a number of additional tools for coding dependent views, as well as depth data, and introduces a special encoder optimization. It provides compatibility with stereoscopic video-only systems, as well as 2D video. Thus, the 3D-HEVC codec also includes MV-HEVC as a subset.

In summary, 3D-HEVC will enable application that require a high compression efficiency, such as transmitting 3D 4K content for stereoscopic as well as auto-stereoscopic multi-view displays.

References

[1] J. Konrad and M. Halle, "3-D Displays and Signal Processing", *IEEE Signal Processing Magazine*, vol. 24, no. 6, pp. 97-111, Nov. 2007.

[2] G. J. Sullivan *et al*, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, Dec. 2012.

[3] J.-R Ohm *et al*, "Comparison of the Coding Efficiency of Video Coding Standards – Including High Efficiency Video Coding (HEVC)," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1669-1684, Dec. 2012.

[4] K. Müller *et al*: "3D High Efficiency Video Coding for Multi-View Video and Depth Data", *IEEE Transactions on Image Processing, Special Issue on 3D Video Representation, Compression and Rendering*, vol. 22, no. 9, pp. 3366-3378, Sept. 2013.

[5] A. Vetro, K. Müller, "Depth-Based 3D Video Formats and Coding Technology", in "Emerging Technologies for 3D Video: Creation, Coding, Transmission and Rendering", pp. 139-161, Editors: F. Dufaux, B. Pesquet-Popesu and M. Cagnazzo, *John Wiley & Sons, Chichester, UK*, ISBN 978-1-1183-5511-4, 2013.

Karsten Müller (M'98-SM'07) received the Dr.-Ing. degree in Electrical Engineering and Dipl.-Ing. degree from the Technical University of Berlin, Germany, in 2006 and 1997 respectively. In 1993 he spent one year studying Electronics and Communication Engineering at Napier University of Edinburgh/Scotland, including a half-year working period at Integrated Communication Systems Inc. in Westwick/England.

He has been with the Fraunhofer Institute for Telecommunications, Heinrich-Hertz-Institut, Berlin, since 1997, where he is currently head of the 3D Coding Group in the Image Processing Department. His research interests include representation, coding and reconstruction of 3D scenes in Free Viewpoint Video scenarios and Coding, multi-view applications and combined 2D/3D similarity analysis. He has been involved in international standardization activities, successfully contributing to the ISO/IEC Moving Picture Experts Group for work items on multi-view, multi-texture and 3D Video Coding. Here, he co-chaired the adHoc group on 3D Video Coding.

Enhancing 3D Video to Enable a Fully Immersive Sensory Experience

Christian Timmerer, Markus Waltl, Benjamin Rainer, Stefan Lederer, and Hermann Hellwagner
Alpen-Adria-Universität Klagenfurt, Austria
{firstname.lastname}@itec.aau.at

1. Introduction

Since the appearance of the movie "Avatar", 3D video gained more and more commercial importance both in cinema and home theaters. Nowadays, most of the current television sets support 3D through stereoscopic or multi-view displays and also Web platforms such as YouTube provide 3D content [1]. 3D video truly enhances the visual quality by providing a more realistic, immersive user experience but still focuses on visual stimulus only.

In our research we target yet another Quality of Experience (QoE) dimension by addressing human senses that go beyond hearing and seeing as the consumption of multimedia content may or shall stimulate potentially all human senses. Therefore, the multimedia content is annotated with sensory information describing sensory effects (e.g., additional ambient light effects, wind, vibration, scent, water spraying) which are synchronized with the actual multimedia content and rendered on appropriate devices (e.g., ambient lights, fans, motion chairs, scent vaporizer, water sprayer, etc.). The ultimate goal of this approach is that the user will also perceive these additional sensory effects giving her/him the sensation of being part of the particular multimedia content and resulting in a worthwhile, informative user experience which is referred to as sensory experience [2].

Related work in this domain comprises, for example, [3] where authors provide vibration and light effects for ringtones on mobile devices. The influence of multimedia content augmented with scents on the viewer is evaluated in [4] including perspectives and challenges. In [5], we found that additional light, wind, and vibration effects increase the QoE and have a positive influence on the perceived emotions. In previous work, we focused on 2D content for enhancing the QoE which is extended here with 3D content. The concept and tools enabling a fully immersive sensory experience is described in Section 2. Section 3 introduces the quality of sensory experience and Section 3 concludes this paper.

2. Sensory Experience: Concept and Tools

The concept and system architecture of receiving sensory effects in addition to audio/visual content is depicted in Figure 1. The media and the corresponding Sensory Effect Metadata (SEM) may be obtained from a Digital Versatile Disc (DVD), Blu-ray Disc (BD), or any kind of online service (e.g., download/play or streaming portal). The media processing engine acts as the mediation device and is responsible for playing the actual media resource and accompanying sensory effects in a synchronized way based on the users' setup in terms of both media and sensory effect rendering. The users' digital living room is extended with additional rendering devices enabling the (increased) stimulation of senses other than hearing and seeing. For example, a motion chair, fan/ventilator, heater/cooler, etc. may be used to address the somatosensory (human sensory) sub-system, whereas scent vaporizer device stimulates the olfactory sub-system. The visual sub-system may be further stimulated using (additional) ambient light devices. Note that the term sub-system refers to the human sensory system comprising the sub-systems visual, auditory, somatosensory, gustatory, and olfactory.

The data format for describing such sensory effects is defined by ISO/MPEG in the context of MPEG-V Media Context and Control. In particular, Sensory Information (Part 3) [6] defines the Sensory Effect Description Language (SEDL), an XML Schema-based language, which enables one to describe sensory effects. The actual sensory effects are not part of SEDL but defined within the Sensory Effect Vocabulary (SEV) for extensibility and flexibility, allowing each application domain to define its own sensory effects. A description conforming to SEDL is referred to as Sensory Effect Metadata (SEM) and may be associated with any kind of multimedia content (e.g., movies, music, Web sites, games).

Software tools for creating and consuming multimedia content enriched with sensory effects are provided within our Sensory Experience Lab under open source licenses and briefly introduced in the following (http://selab.itec.aau.at):

- *Off-the-shelf amBX system & SDK* (http://www.ambx.com/) comprising left and right 2.1 sound speaker lights with a sub-woofer, a wall washer, a set of fans, and a wrist rumbler including an appropriate SDK in order to control these devices.
- *Sensory Effect Video Annotation (SEVino)* tool allows for describing sensory effects of a video sequence. It is based on Java and provides means

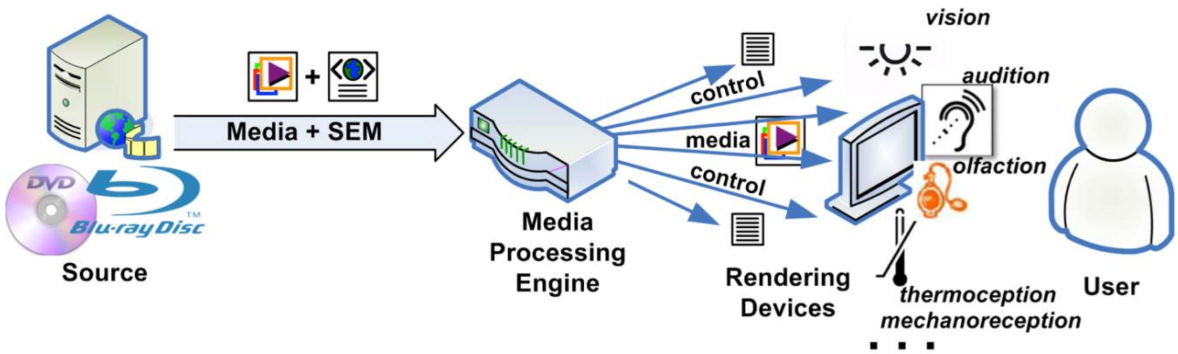


Figure 1. Concept of Sensory Experience.

for simply entering and editing of sensory effects.

- Sensory Effect Metadata Player (SEMP) is a DirectShow-based media player with support for sensory effects and the amBX system.
- The *AmbientLib* provides plug-ins for various Web browsers (e.g., Firefox, Opera, IE) which then are able to render sensory effects within the browser from HTML5 or Flash videos.

In order to support stereoscopic video, SEMP has been extended as follows. MPEG-V allows the specification of automatic extraction of sensory effects and, thus, SEMP includes an automatic average color calculation algorithm. The color calculation algorithm uses a so-called SampleGrabber for retrieving the currently displayed frame and splits it into three parts (left, middle, and right). From each part the average color is calculated and used for rendering light effects on available devices. In the case of stereoscopic video, only the left image of a stereoscopic video frame is used to calculate the average color for the three parts. For the actual rendering of stereoscopic videos, SEMP incorporates the DirectShow filters of the commercial Stereoscopic Player (http://www.3dtv.at).

An example deployment is shown in Figure 2.

Figure 2. Example Deployment of 3D Video enhanced with Sensory Effects [7].

3. Quality of Sensory Experience

The QoE is defined as is defined as "the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and / or enjoyment of the application or service in the light of the users' personality and current state" [8].

However, different application domains may have different requirements in terms of QoE. Thus, there is

a need to provide specializations of this generally agreed definition of QoE pertaining to the respective application domain taking into account its requirements formulated by means of influence factors and features of QoE (cf. Sections 5 and 6 of [8]). Consequently, an application-specific QoE definition is provided by selecting the influence factors and features of QoE reflecting the requirements of the application domain and incorporating them into the generally agreed definition of QoE. In the context of this research the QoE is referred to as the Quality of Sensory Experience (QuaSE).

QoE evaluations are typically done through subjective quality assessments following predefined methods (e.g., ITU-T P.910, P.911) and QoE models are derived from that. Based on the results of various subjective quality assessments [2][5], we have defined a general utility model for evaluating QuaSE [9] which suggests a linear relationship between the QoE of the multimedia content without sensory effects (QoE_{wo}) and the QoE of the multimedia content which is enriched with sensory effects (QoE_w). Equation (1) shows our proposed utility model for sensory experience.

$$\mathbf{QoE_w = QoE_{wo} * (\delta + \sum w_i b_i)} \qquad (1)$$

In our utility model, w_i represents the weighting factor for a sensory effect of type i, e.g., in our setup $i \in \{light(l), wind(w), vibration(v)\}$. Further sensory effect types (e.g., scent) may be incorporated easily, e.g., as soon as appropriate devices become available. The binary variables b_i ($b_i \in \{0,1\}$) are used to identify whether effect i is present for a given setup and δ is used for fine-tuning. The QoE_{wo} may be assessed through any existing model such as those given in [10] or by an appropriate QoS to QoE mapping [11].

Our utility model has been validated against previous subjective quality assessments [9] showing that the highest variability is provided by the variable denoting whether vibration is available or not. Hence, vibration effects have the biggest influence on the QoE, whereas wind and light have a lower impact.

4. Conclusions

In todays' efforts towards defining the QoE, the user is becoming the center of the research focus. However, for new multimedia applications such as 3D or ultra high-definition, QoE models are still in its infancy.

Our work focused on yet another quality dimension aiming to come up with a holistic view and providing tools to create, consume, and capture the Quality of Sensory Experience. The tools are available as open source and a general utility model has been proposed including preliminary evaluations. Future work items include the validation (and refinement, if needed) of the proposed model by applying it within emerging multimedia application such as 3D and ultra high-definition.

Acknowledgments: This work was supported in part by the EC in the context of the ALICANTE (FP7-ICT-248652), SocialSensor (FP7-ICT-287975), and QUALINET (COST IC 1003) projects and partly performed in the Lakeside Labs research cluster at AAU.

References

[1] YouTube Blog, "Now in 3D: Join the Experiment With Us!", 2009. Online: http://bit.ly/lZ6pd (last access: Nov. 2013).

[2] C. Timmerer, et al., "Assessing the Quality of Sensory Experience for Multimedia Presentations", *Elsevier Signal Processing: Image Communication*, vol. 27, no. 8, pp. 909-916, 2012.

[3] A. Chang and C. O'Sullivan, "Audio-Haptic Feedback in Mobile Phones", *ACM Conference on Human Factors in Computing Systems (CHI '05)*, Portland, OR, USA, pp. 1264-1267, 2005.

[4] G. Ghinea and O. A. Ademoye, "Olfaction-Enhanced Multimedia: Perspectives and Challenges", *Multimedia Tools and Applications*, vol. 55, no. 3, pp. 1-26, 2011.

[5] B. Rainer et al., "Investigating the Impact of Sensory Effects on the Quality of Experience and Emotional Response in Web Videos", *Proc. of the 4th Int'l. Workshop on Quality of Multimedia Experience (QoMEX'12)*, Yarra Valley, Australia, pp. 278-283, 2012.

[6] ISO/IEC 23005-3, Information Technology — Media Context and Control — Sensory Information, ISO/IEC JTC 1/SC 29/WG 11/N11425, Geneva, Switzerland.

[7] M. Waltl, et al., "A 4D Multimedia Player enabling Sensory Experience", *Proc. of the 5th Int'l Workshop on Quality of Multimedia Experience (QoMEX'13)*, Klagenfurt, Austria, pp. 126–127, 22013.

[8] P. Le Callet, S. Möller, A. Perkis, (eds.) "Qualinet White Paper on Definitions of Quality of Experience", *European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003)*, Lausanne, Switzerland, Version 1.2, 2013.

[9] C. Timmerer, et al., "A Utility Model for Sensory Experience", *In Proceedings of the 5th International Workshop on Quality of Multimedia Experience (QoMEX'13)*, Klagenfurt, Austria, pp. 224-229, 2013.

[10] F. Pereira, "A Triple User Characterization Model for Video Adaptation and Quality of Experience Evaluation", *Proc. of the 7th Workshop on Multimedia Signal Processing (MMSP)*, Shanghai, China, pp. 1-4, 2005.

[11] Fiedler, M., Hossfeld, T., Tran-Gia, P., "A generic quantitative relationship between quality of experience and quality of service", *IEEE Network*, vol.24, no.2, pp.36-41, March-April 2010.

Christian Timmerer is an assistant professor at Alpen-Adria-Universität Klagenfurt, Austria and his research focus is on immersive multimedia communication, streaming, adaptation, and quality of experience. He was general chair of WIAMIS 2008 and QoMEX 2013 and has participated in several EC-funded projects, notably DANAE, ENTHRONE, P2P-Next, ALICANTE, SocialSensor, and the COST Action IC1003 QUALINET. Dr. Timmerer also participated in ISO/MPEG work for several years – notably, in the area of MPEG-21, MPEG-M, MPEG-V, and MPEG-DASH. Follow him on http://www.twitter.com/timse7 and subscribe to his blog http://blog.timmerer.com.

Markus Waltl was a Ph.D. student at Alpen-Adria-Universität Klagenfurt, Austria. He finished his Ph.D. thesis in 2013 on the impact of sensory effects on the quality of multimedia experience. He is now with SKIDATA AG, Klagenfurt, Austria.

Benjamin Rainer is a Ph.D. student at Alpen-Adria-Universität Klagenfurt, Austria, where he received his M.Sc. (Dipl.-Ing.) degree in Computer Science. His research interests include sensory effects, audio/video encoding with emphasis on video encoding, parallel computing, and Quality of Experience.

Stefan Lederer is a university assistant at Alpen-Adria-Universität Klagenfurt, Austria. He received his M.Sc. (Dipl.-Ing.) in Computer Science in Mar'12 from the Alpen-Adria-Universität (AAU) Klagenfurt. His research topics include transport of modern/rich media, multimedia adaptation, QoS/QoE as well as future internet architectures. He participated in several EC-funded Projects (ALICANTE, SocialSensor) and in the MPEG-DASH standardization.

Hermann Hellwagner is a full professor of Informatics in the Institute of Information Technology, Alpen-Adria-Universität Klagenfurt, Austria, leading the Multimedia Communications group. His research areas are distributed multimedia systems, multimedia communications, and quality of service. He is a senior member of the IEEE, member of the ACM, GI, OCG, and Vice President of the Austrian Science Fund.

SPECIAL ISSUE ON DEVICE-TO-DEVICE COMMUNICATIONS

Guest Editor: Yik-Chung Wu, The University of Hong Kong, HK

ycwu@eee.hku.hk

As smart phones become part of our lives, the demand of wireless data is ever-increasing; the traditional infrastructure based cellular system will fail to catch up sooner or later. Then, the idea of device-to-device (D2D) emerges. With two nearby devices directly communicating without going through the basestation, it is obvious that this D2D communication mode would save energy, increase spectral efficiency, extend coverage in cellular system, and help creating new services such as D2D-assisted social networking, and location-based mobile advisement. Due to the potential of benefits brought by D2D communications, D2D communications has attracted a lot of attention in the past few years from both academia and industry. It is believed that D2D communications will be incorporated into future wireless system standard. This special issue gathered six articles covering different aspects of this exciting topic.

The first one "Device-to-Device Communications in Cellular Networks" by Chen et al, is an overview of the D2D concept in cellular networks. It provides basic descriptions on different elements of D2D communications such as mode selection, device discovery, spectrum selection, resource allocation, and integration with LTE. Potential research directions are also discussed. The second article "Device-to-Device Communication and Network Coding: Friends or Foes?" by Fodor et al, discusses the advantages and disadvantages of combining D2D and network coding in cellular system. The performance gains that D2D plus network coding brings in terms of end-to-end SINR, spectral efficiency, and average invested power, are presented.

While D2D communications improve spectral efficiency by reusing the frequency in the cellular system, it introduces interference to the cellular users. Thus interference management is an important topic in D2D research. And we have two contributions in this special issue. The letter "Energy Efficient Scheduling for LTE-A D2D Communication" by Mumtaz et al, discusses a scheduling scheme that allocates physical resource block aiming at minimizing the energy per bit of the system. Performance advantages compared to an existing scheduler is demonstrated. Another letter on interference management is "Resource Allocation for Device-to-device Communications Underlaying Cellular Networks" by Feng et al, discusses the resource allocation algorithms under perfect CSI and partial CSI. For partial CSI case, this letter covers methods on user selection based limited feedback, probability-based resource allocation, and a combination of the above two methods. Potential research topics are also given.

The fifth letter is "Infrastructureless Synchronization for Device-to-Device Discovery" by Cheng et al, introduces how recent advances on clock synchronization in wireless sensor network might help solving the synchronization issue in D2D communications, especially in device discovery phase. The last article of this special issue "Research Progress on Device-to-Device Communication in China" by Song et al, summarizes different D2D research activities currently underway in China, covering both academic research and industrial projects. This article gives us a peek on how China engages and helps shaping this important topic.

I hope you would enjoy reading these six selected pieces, and obtain a quick general idea on what D2D communications is and why it is important in the future. It is also hope that this special issue would stimulate more ideas and collaborations, further refining D2D communications from different perspectives.

Yik-Chung Wu received the B.Eng. (EEE) degree in 1998 and the M.Phil. degree in 2001 from the University of Hong Kong (HKU). He received the Croucher Foundation scholarship in 2002 to study Ph.D. degree at Texas A&M University, College Station, and graduated in 2005. From August 2005 to August 2006, he was with the Thomson Corporate Research, Princeton, NJ, as a Member of Technical Staff. Since September 2006, he has been with HKU, currently as an Associate Professor. His research interests are in general area of signal processing and communication systems, and in particular distributed signal processing and communications; optimization theories for communication systems; estimation and detection theories in transceiver designs; and smart grid. Dr. Wu served as an Editor for IEEE Communications Letters, is currently an Editor for IEEE Transactions on Communications and Journal of Communications and Networks.

Device-to-Device Communications in Cellular Networks

Hao Chen[1], Lingjia Liu[1], Ying Li[2], Bin Yu[3], Lei Shi[4], and Yang Yi[5]
[1]Department of Electrical Engineering and Computer Science, University of Kansas, USA
[2]Samsung Telecommunications, Dallas, TX, America
[3]Samsung Telecommunications, Beijing, China
[4]Department of Automation, Tsinghua University, Beijing, China
[5]Department of Computer Science and Electrical Engineering, University of Missouri – Kansas City, USA
{hao001,lingjialiu}@ittc.ku.edu, ying.li@ieee.org, bin82.yu@samsung.com, shil11@mails.tsinghua.edu.cn, cindyyi@umkc.edu

1. Introduction

Mobile wireless communications have profoundly changed everyday life on a global scale over the last few decades. According to the International Telecommunication Union (ITU) [1], in 2013, the number of global mobile wireless subscriptions surpassed 6.8 billion. In May of 2013, Cisco Systems [2] predicted a staggering 66% compound annual growth rate for global mobile wireless data traffic from 2012 to 2017, with all indications pointing towards an explosive growth of 13 times by year 2017. Under this situation, "how to response to the expected expansion of wireless connectivity" becomes a pressing issue. Device to device (D2D) communication, an evolving product of cellular networks, has attracted attention from both industry and academia in the past decades. This new interest is motivated by several factors: the first factor is that current cellular networks could not support the huge number of connected devices in future application scenarios, such as machine-to-machine (M2M) communication and Internet of Things (IoT); the second factor is that ad hoc networks could not provide enough throughput as the number of connected devices increases, according to the results on wireless network scaling laws. Combing the advantages of infrastructure-based cellular communications and ad-hoc communications, D2D communication brings a promising communication paradigm and will likely be adopted in 3GPP Release 12 LTE-Advanced systems.

D2D communication in cellular networks is defined as direct communications between two mobile users without traversing the Base Station (BS) or core networks [3]. By exploiting direct communications between nearby mobile devices, D2D communication in cellular networks has advantages such as improved spectrum reuse, improved energy efficiency, reduced end-to-end delay, and extended coverage etc. D2D communication can be widely used in proximity based services (ProSe), such as local advertisement, social networking applications, public safety and local data transfer. Various D2D application cases have been identified in [4].

In the following letter, we provide an overview of current development in D2D communication in cellular networks and also present some potential developing directions for this technique.

2. D2D in Cellular Networks

In this section, we will introduce some important issues related to conducting D2D communications in cellular networks.

Mode Selection (D2D link vs. Cellular link)

The first issue in cellular D2D communications is to decide whether a device should work in the D2D mode or the cellular mode. This mode selection scheme can be done by the infrastructure or by the device itself. Based on comparing the path loss between the device and BS to a pre-defined threshold, [5] gives a mode selection scheme that enables a device to choose communication mode individually. This scheme is computing efficient, but it may cause interference to other D2D users. Therefore, the majority of the literature [3][5][6][9] focus on the scenario where the BS assigns devices communication mode because BS can coordinate the interference of the whole network. Furthermore, the BS can jointly optimize the mode selection and resource allocation under various criteria such as spectral efficiency, power efficiency, and system throughput depending on different application scenarios.

Device Discovery

Device Discovery refers to the process of finding proximity devices for mobile devices who want to set up D2D links. This function could be assigned to terminals (direct discovery) or the infrastructure (Evolved Packet Core (EPC) level discovery) [7]. In direct discovery case, it is mobile device itself that searches nearby neighbors autonomously by broadcasting discovery signals. The broadcasting signal may include mobile device identity, location and application-related information. An idea about reusing existing LTE/LTE-Advanced physical layer signals such as PSS/SSS and PRACH to do device discovery has been introduced in [8]. Although the broadcasting will reduce mobile device's battery life, direct discovery can work no matter whether infrastructure is available or not. Therefore, direct discovery can be used in public safety scenarios.

If the infrastructure (BS or EPC) is assigned to do device discovery, the infrastructure can choose the best neighbor for the device among various candidates. Furthermore, the handover overhead for device discovery will be dramatically reduced if the overall network information in the infrastructure is exploited. However, EPC-level discovery will not be available when the mobile device is out of the coverage or the infrastructure fails to work in a disaster.

Spectrum Selection (In-band vs. Out-of-band)

Based on the spectrum sharing strategy, D2D communication can be classified into two types: in-band D2D, where D2D links utilize cellular spectrum and out-of-band D2D, where D2D links exploit unlicensed spectrum (e.g. 2.4GHz ISM band). The key problem in in-band D2D is how to control the interference between D2D users and cellular users, while the problem in out-of-band D2D is that D2D users have to compete with the various users in the unlicensed band which leads to low quality of service (QoS) guarantee. If the cellular link and D2D link use orthogonal cellular spectrum, this in-band paradigm is called overlay; otherwise, if the D2D link can access to the same spectrum concurrently with cellular links, this paradigm is called underlay. This is much similar with the category method in cognitive radio networks.

The advantage of the overlay case is that both the D2D communications and the cellular communications can achieve high link quality since there is no interference between D2D links and cellular links. However, the network has to decide how much spectrum should be assigned for D2D communication efficiently. In the underlay case, on the contrary, the total throughput of the network can be very high; however, complicated interference management schemes have to be utilized to guarantee the QoS of the communication links.

By integrating the idea of sensing based spectrum sharing in cognitive radio networks [12], a new spectrum access paradigm, sensing based spectrum sharing D2D communications could be used to improve the overall system performance. In this new mode, D2D mobile devices have the ability to perform spectrum sensing. If the base station is available, the mobile device will feed back the sensing results to the base station as a part of the D2D communication request. Otherwise, depending on the sensing results, mobile device will act differently in this mode: if the sensing result shows that the spectrum is not occupied by current cellular mobile devices, the D2D device will access the spectrum with a relatively high power; on the other hand, if the sensing result shows that the spectrum is occupied, the D2D device will transmit concurrently with cellular mobile devices but at a lower power on this occupied spectrum. In order to guarantee the QoS of cellular mobile devices, interference constraints or rate loss constraints of cellular links can be added to D2D links.

By combing overlay and underlay paradigms, the advantages for this new paradigm can be summarized as follows. Firstly, this reduces the computing burden of base station by giving more freedom to mobile devices. One of the application scenarios of D2D communications is M2M, where BS needs to do spectrum allocation for hundreds of D2D links. It is a heavy burden at the base station which also requires a lot of control overhead. Secondly, this mode can achieve higher throughput than the traditional overlay or underlay paradigm. In underlay paradigm, D2D link may waste the spectrum resource if they still take the interference to cellular devices into consideration when cellular spectrum is not occupied. Thirdly, through coordinating with the base station, this paradigm could take advantage of both infrastructure-based cellular network and cognitive radio based ad hoc networks.

Integration with LTE/LTE-A PHY

The LTE/LTE-Advanced standard supports two modes of duplex operation, time division duplex (TDD) and frequency division duplex (FDD). The difference in duplex operation will lead to different system design for D2D communication. In FDD mode, the downlink and uplink of cellular communication will use different access schemes (uplink using SC-FDMA, downlink using OFDMA) and spectrum band. If D2D links reuse the downlink resource, it will cause severe interference to its proximity receiving mobile devices. On the other hand, if the D2D link reuses uplink resources, the interference caused to BS may be tolerable since BS is relatively far away. Furthermore, multi-antennas equipped at BS can be used to mitigate the interference. Considering the above mentioned issues together with the additional transmitter/receiver complexity, SC-FDMA is more suitable for mobile devices to conduct D2D communications in cellular networks. In TDD mode, the uplink and downlink use the same spectrum band and the same modulation (OFDMA), there is no difference on reusing uplink or downlink resources. Therefore, reusing uplink cellular resource is recommended for the future D2D communication and it is much easier to conduct D2D communication in TDD LTE/LTE-Advanced cellular networks.

Resource Allocation

Resource allocation and interference management is one of the key challenges in the development of D2D communications and has attracted lots of research interest. Different from the same issue in ad hoc networks, the infrastructure can be exploited to jointly optimize power allocation, time and frequency resource block assignment, and mode selection in order to improve the whole network performance (end to end delay, throughput, spectral efficiency and energy efficiency, coverage etc). Taking the interference into account, the power allocation problem will probably be non-convex optimization problem. Furthermore, to model the mode selection scheme, integer variables have to be introduced. Therefore, the joint optimization problem in D2D communications will be likely in the form of mixed integer nonlinear programming as indicated by [9]. The various simplified versions of the problem and their corresponding solutions have been surveyed in [3]. However, these approaches are based on different systematic assumptions and preconditions. System designers may find it hard to choose from these strategies since there have not been a generally accepted theoretical framework or systematic simulation platform to compare these results.

3. Potential Future Work

In this section, we elaborate on some of the possible research directions on and open problems in D2D communications in cellular networks:

1) **Cross-layer design**. In the proposed SBSS-D2D communication, the system performances (throughput, energy efficiency etc.) are jointly decided by the spectrum sensing time, sensing threshold, power allocation, the mode selection strategy, transmission strategy etc. Therefore, the cross layer optimization mechanism will definitely improve the overall system performance.
2) **New medium access mechanism for huge number of D2D links**. Current infrastructure-mode cellular links are centrally scheduled by the BS. However the computing burden will become heavier as the number of D2D links scales. The CSMA based ad hoc mode causes no computing burden to BS, but achieve lower spectral-efficiency. In D2D communication, new medium access mechanism should be developed which fully exploits the resource at BSs to reduce the overhead as well as decrease the computing complexity.
3) **Systematic performance analyze of D2D communication**. The majority of the available literature on the performance of D2D networks is based on numerical simulations or home grown simulations. Those evaluations are suitable for justifying specific algorithms and scenarios, but may not be convincing or suitable for providing detail guidance for system design. Therefore, a systematic framework is needed to evaluate the effects of various D2D communication parameters on system performance. Recently, stochastic geometry, which models the nodes distribution as Poisson point process (PPP), has been applied to analyze the performance of D2D communication [10]. The authors provide a tractable hybrid network model and unified performance analysis approach which can be extended to analyze the performance of different schemes in D2D networks. Furthermore, it has been shown that PPP model is about as accurate in terms of both SINR and handover rate as the hexagonal grid for a representative urban cellular network [11].

4. Conclusion

In this letter, we describe D2D communications in cellular networks. Some important topics related to system design for D2D in cellular networks are discussed. A new D2D in-band spectrum sharing paradigm is introduced to improve the overall system performance. We also elaborate some potential future directions in D2D cellular networks.

References

[1] ITU-D, “Measuring the information society 2011,” in *International Telecommunications Union (ITU)*, 2011.
[2] Cisco Systems, Cisco Visual Networking Index: Forecast and Methodology, 2012-2017, May 2013.
[3] A. Asadi, Q. Wang, V. Mancuso, “A survey on Device-to-Device Communication in Cellular Networks”, submitted to *IEEE Communications Sruveys and Tutorials*, Oct. 2013. arXiv:1310.0720
[4] 3GPP, “3rd generation partnership project; technical specification group sa; feasibility study for proximity services (prose) (release 12),” August 2012, Tech. Report.
[5] H. Xie, P. E. Omiyi and H. Haas, “Maximizing spectral efficiency in 3G with hybrid ad-hoc UTRA TDD/UTRA FDD cellular mobile communications,” in Proceedings of the *IEEE International Symposium on Spread Spectrum Techniques and Applications, Sydney, Australia*, Sep. 2004, pp. 613–617.
[6] E. Frlan, “Direct communication wireless radio system,” United States Patent 6047178, April 2000.
[7] 3GPP, “3rd generation partnership project; technical specification group SA; study on architecture enhancements to support proximity services (prose) (release 12),” June 2013, TR23.703 V0.4.1.
[8] X. Lin, J. G. Andrews, A. Ghosh and R. Ratasuk, “An Overview on 3GPP Device-to-Device Proximity Services”, arXiv:1310.0116, October, 2013.
[9] M. Belleschi, G. Fodor, A. Abrardo, "Performance analysis of a distributed resource allocation scheme for D2D communications," *IEEE* GLOBECOM Workshops (GC Wkshps), vol., no., pp.358,362, 5-9 Dec. 2011
[10] X. Lin, J. G. Andrews, and A. Ghosh, “A comprehensive framework for device-to-device communications in cellular networks,” submitted to *IEEE Journal on Selected Areas in Communications*, May 2013. Available at arXiv:1305.4219.

[11] J. Andrews, F. Baccelli, and R. Ganti, “A tractable approach to coverage and rate in cellular networks,” *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122-3134, 2011.

[12] Kang X; L. Ying-Chang; H.K. Garg, Z. Lan, "Sensing-Based Spectrum Sharing in Cognitive Radio Networks," *IEEE Transactions on Vehicular Technology*, vol.58, no.8, pp.4649,4654, Oct. 2009

Hao Chen received his B.S. and M.S. both in Information Engineering from Xi’an Jiao Tong University, China. Now, He is pursuing his Ph.D. in the University of Kansas, USA. His research interests include cross-layer design in cognitive radio network and D2D communications.

Lingjia Liu (S’03–M’08) received the B.S. degree in Electronic engineering from Shanghai Jiao Tong University, China, and the Ph.D. degree in Electrical Engineering from Texas A&M University, USA. He is currently working as an Assistant Professor in the EECS Department at the University of Kansas (KU), USA. Prior to that he spent more than three years in Samsung Research America leading Samsung’s work on multi-input-multiple-output (MIMO), coordinated multi-point (CoMP) transmission, and heterogeneous networks for 3GPP LTE and LTE-Advanced standards. His general research interests lie in the areas of wireless communication systems, statistical signal processing, queuing theory, information theory, with emphasis on delay-sensitive communication over wireless networks, and smart energy networks.

Dr. Liu served as Technical Program Committee (TPC) member/chairs of various international conferences including *IEEE* ICC, *IEEE* GLOBECOM, *IEEE* ICNC, WCSP, etc. He currently serves as an Editor for *IEEE Trans. on Wireless Comm.*, and Associate Editors for *EURASIP Journal on Wireless Comm. and Networking* and *International Journal on Comm. Systems*.

Ying Li (S’07–M’09) received the B.E. degree (with honors) and the M.E. degree in Electrical Engineering from Xi’an Jiao Tong University China, and the M.A. and Ph.D. degrees in Electrical Engineering from Princeton University, USA. She has been with Samsung Research America since October 2008, where she is involved in research on smart energy networks and wireless communications. She was a visiting Ph.D. student in Swiss Federal Institute of Technology (EPFL), Switzerland and in Motorola Multimedia Research Lab respectively. She worked as a faculty member in the Department of Information and Communication Engineering at Xi’an Jiao Tong University from 2000 to 2003 and as a visiting scholar in Fuji Xerox Co. Ltd., Japan, from 2000 to 2001. Her research areas include optimization, smart energy networks, energy monitoring and management, communication networks, next generation wireless communications, heterogeneous networks, cross-layer design, content distribution, information theory, and signal processing.

Bin Yu is a senior engineer in Beijing Samsung Telecom R&D Center. He received his M.S. in Radio Frequency Communications from University of Southampton in Feb. 2006. He has over 7 years’ experience of conducting leading-edge research on several wireless communication systems. He has been participating in standardization of both TD-SCDMA HSPA and LTE/LTE-A in 3GPP since 2006, where he has contributed approximately 100 technical proposals in RAN working group 1 and 2.

Lei Shi received her B.S. in Communication Engineering from Beijing University of Post and Telecommunication in 2011. Since then, she has been a Master student in the department of automation in Tsinghua University, China. Currently, she is a visiting scholar at the University of Kansas, USA. Her research interests are D2D communications and wireless communications.

Yang (Cindy) Yi received the B.S. and M.S. degrees with honors in Electronic Engineering from Shanghai Jiao Tong University, Shanghai, China. She completed the Ph.D. degree in Electrical and Computer Engineering at Texas A&M University, USA. She is currently working as an Assistant Professor in CSEE Department at the University of Missouri – Kansas City (UMKC). Her research interests lie in the general area of VLSI-CAD, high performance computing, and energy-efficient circuit design.

Device-to-Device Communication and Network Coding: Friends or Foes?

Gabor Fodor, Ather Gattami and Aidilla Pradini
Ericsson Research
{*gabor.fodor|ather.gattami|aidilla.pradini*}*@ericsson.com*

1. Introduction

Device-to-device (D2D) communication in cellular spectrum supported by a cellular network enables direct communication between pieces of user equipment (UE) [1]–[2]. The main reason for incorporating D2D communication in cellular networks is to exploit the proximity of UEs when engaged in local communication sessions such as social networking, media sharing, or proximity-based services [3].

In the presence of such proximate communication opportunities, D2D has been shown to harvest not only the proximity gain in terms of improved link budget, but also the so called reuse and hop gains [2], [5]–[8]. A key technology component of D2D is *mode selection (MS),* which selects the cellular or direct communication mode for a D2D pair based on factors such as the current resource situation, traffic load, and interference level [8]. Recognizing the potential of D2D, the research community has proposed efficient scheduling, resource allocation, and power control algorithms that help realize the gains of local communications, while at the same time protecting the cellular layer from interference caused by local traffic [9]. These promising results have triggered standards bodies such as the 3rd Generation Partnership Project (3GPP) to study the possibilities of introducing D2D in future releases of Long-Term Evolution (LTE) networks [4].

Along another line of research, it has been observed that physical layer network coding (NWC) improves the spectrum efficiency by facilitating resource reuse by multiple transmissions and taking advantage of advanced signal processing techniques [10]–[11]. Despite the obvious differences between cellular network-integrated D2D and NWC technologies, both aim to improve spectral efficiency and increase network capacity by enabling tighter reuse of resources. As a related study has noted, NWC can be used, under some assumptions, to further enhance the efficiency of D2D communication by combining cellular and direct transmission in integrated D2D and cellular networks [12]. On the other hand, the joint application of D2D and NWC may be costly in terms of UE capabilities, measurement reports, and signaling, while it is not clear whether an integrated D2D-NWC-based solution for dealing with local traffic results in additional gains over a system that uses *either* D2D or NWC alone. We aim to answer the following questions: (1) Does NWC provide gains in integrated D2D-cellular networks? (2) Does D2D provide gains in a cellular network employing NWC?

The work of the first author has been partially performed in the framework of the FP7 project ITC 317669 METIS.

Therefore, the purpose of the present letter is to discuss some of the advantages and disadvantages of employing D2D and NWC in cellular networks and to explore the potential of integrating these two technology components in such a way that their joint usage benefits both the cellular and the D2D (local) traffic.

2. Employing D2D and NWC to Support Local Traffic

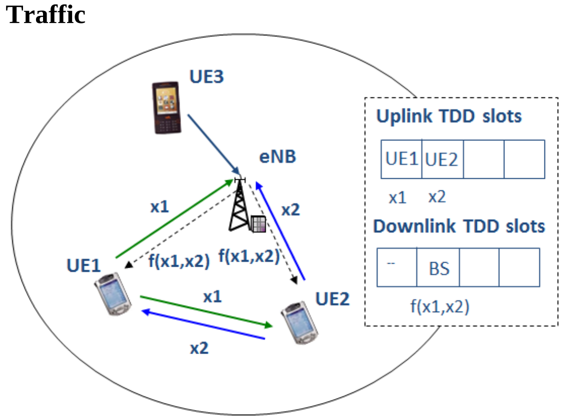


Figure 1: D2D and NWC technologies integrated in a cellular network.

See Figure 1 for a comparison of the operation of D2D and NWC in the presence of local traffic. In this scenario, UE1 and UE2 are served by the same base station (eNB) and engage in a local communication session. D2D technology offers the possibility of direct communication, in which case a bidirectional exchange of signals x_1 and x_2 require two orthogonal resources (because we do not consider the application of full-duplex radio in this letter). For example, assuming time division duplexing (TDD), using D2D, x_1 and x_2 can be sent in subsequent time slots (TS).

If UE1 and UE2 are served by the same eNB in a cellular network employing only physical layer NWC (that is, without D2D communication), two time slots can support the exchange of x_1 and x_2. In this case, UE1 and UE2 transmit on the same resource (TS-1), while the eNB uses TS-2 to transmit the NWC data $\boldsymbol{f(x_1, x_2)}$ to UE1 and UE2 simultaneously [10]-[12]. UE1 and UE2 receive $\boldsymbol{f(x_1, x_2)}$ and decode x_2 and x_1, respectively.

As an alternative to the proposed physical layer (also called 2 time slot, 2-TS) network coding scheme, the 3-TS scheme uses different resources for transmitting x_1 and x_2 to the eNB, while the eNB uses TS-3 to transmit $\boldsymbol{f(x_1, x_2)}$. In this 3-TS scheme, the joint application of D2D and NWC becomes possible, as depicted in Figure 1. In the joint mode, UE2 receives *both* the direct transmission from UE1 (in uplink TS-1) *and* the network coded transmission from the eNB (in DL TS-2). UE2 can then employ signal processing (for example, the maximum ratio combined with maximum likelihood detection, as in [12]) to combine the received signals such that the bit error rate is improved over the 2-TS scheme without D2D transmission and reception.

Figure 2 summarizes the possible transmission modes enabled for local traffic by cellular, D2D, and NWC technologies. Note that, in Figure 2, the traditional cellular transmissions (without D2D and NWC) of x_1 and x_2 correspond to a 4-TS scheme [10], since both x_1 and x_2 must be transmitted through the eNB using an uplink *and* a downlink resource.

3. Performance Aspects

In order to gain insights into the advantages of the available local communication schemes, including D2D or NWC alone or in combination, we consider a simplified signal model and use a realistic system simulator to analyze system performance. We evaluate the end-to-end signal-to-interference-and-noise ratio (SINR), total transmit power in the bidirectional transmission ($P_A + P_B + P_R$), and spectral efficiency (a logarithmic function of the SINR divided by the number of required time slots).

Signal Model

The signal models applicable for NWC- and D2D-based transmissions are shown by Figure 3, where h denotes the complex channel coefficients and n denotes additive Gaussian noise.

Mode Selection

Based on Figure 2 and the signal model, we expect a tradeoff between the number of used resources, invested transmission energy, and the resulting SINR levels, and thereby the achieved spectral and energy efficiencies. Therefore, for an integrated D2D-NWC-cellular network, we developed two MS algorithms that aim to maximize the achieved SINR (MS-NWC 1) and the spectral efficiency (MS-NWC 2), respectively.

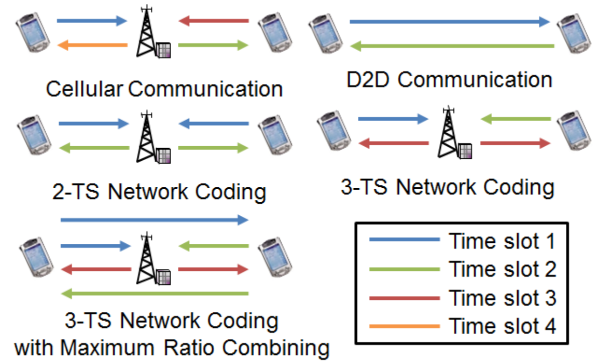


Figure 2: An overview of the available transmission modes for local communications.

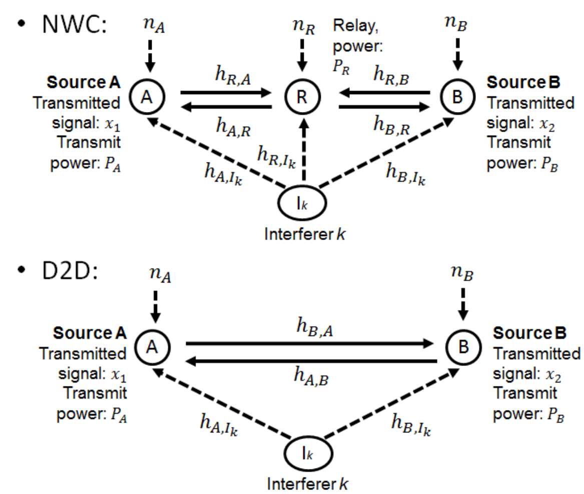


Figure 3: Signal models for NWC and D2D transmissions.

4. Numerical Results

Figure 4 compares the SINR performance of the transmission schemes of Figure 2, along with MS algorithms in a cellular network that supports both NWC and D2D (MS-NWC 1 and MS-NWC 2) and MS in an integrated D2D-cellular network (MS without NWC). The SINR is maximized with proper MS and the gain of employing NWC in an integrated cellular and D2D network in terms of SINR is negligible.

As Figure 5 shows, however, NWC can lead to significant spectral efficiency increase if proper mode selection is employed; for example, MS-NWC 2. This shows that spectral efficiency is the main benefit of introducing NWC into an integrated D2D-cellular network.

Furthermore, as indicated by Figure 6, this high spectral efficiency can be realized at low power consumption (see Figure 6, MS-NWC 2 column); in other words, a network that supports both D2D and NWC is more energy-efficient than a network that supports only NWC. This is the gain that D2D brings in a cellular network employing NWC.

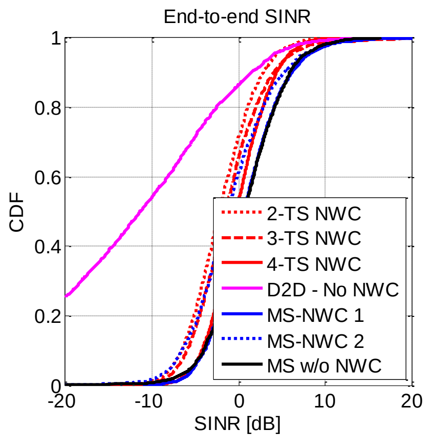


Figure 4: End-to-end SINR performance of the transmission schemes under study.

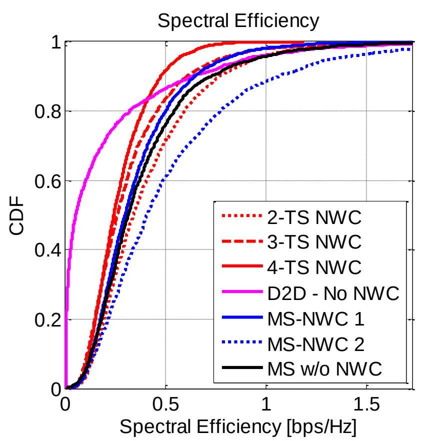


Figure 5: Spectral efficiency of the transmission schemes under study.

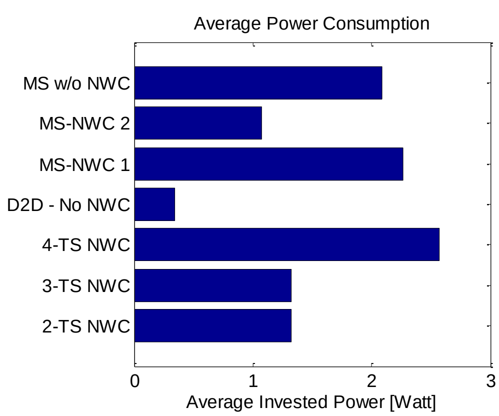


Figure 6: Average power consumption of the transmission schemes under study.

4. Conclusion

In this electronic letter, we have raised the question of how D2D and NWC technologies can be integrated in cellular networks in scenarios in which local (proximate) communication opportunities exist. Initial investigations suggest that D2D and NWC can complement one another and be friends provided a proper mode selection algorithm is applied by the network.

References

[1] K. Doppler, M. Rinne, C. Wijting, C. Riberio and K. Hugl, "Device-to-Device Communication as an Underlay to LTE-Advanced Networks", *IEEE Comm. Mag.*, Vol.7, No. 12, pp. 42–49, 2009.

[2] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider, G. Miklos and Z. Turanyi, "Design Aspects of Network Assisted Device-to-Device Communications", *IEEE Comm. Mag.*, Vol. 50, No. 3, pp. 170–177, March 2012.

[3] L. Lei, Z. Zhong, C. Lin, X. Shen, "Operator Controlled Device-to-Device Communications in LTE-Advanced Networks", *IEEE Wireless Comm.*, pp. 96–104, June 2012.

[4] 3GPP *Tech. Report* 22.803, "Scenarios and Requirements for General Use Cases and National Security and Public Safety", May 2013.

[5] N. Lee, X. Lin, J. Andrews and R. Heath, "Power Control for D2D Underlaid Cellular Networks: Modeling, Algorithms and Analysis", submitted to IEEE JSAC. arXiv:1305.6161 [cs.IT].

[6] G. Fodor and N. Reider, "A Distributed Power Control and Mode Selection Algorithm for D2D Communications", *EURASIP J. on Wireless Comm. and Networking*, 2012.

[7] G. Fodor, M. Belleschi, D. D. Penda, M. Johansson, "A Comparative Study of Power Control Approaches to D2D Power Control", *IEEE ICC*, Budapest, Hungary, June 2013.

[8] M. Belleschi, G. Fodor, A. Abrardo, "Performance Analysis of a Distributed Resource Allocation Scheme for D2D Communications", *IEEE WS on M2M Comm*, Houston, TX, USA, December 2011.

[9] A. Asadi, Q. Wang, V. Mancuso, "A Survey of Device-to-Device Communication in Cellular Networks", submitted to the *IEEE Surveys on Communications*, October 2013.

[10] R. H. Louie, Y. Li, B. Vucetic, "Practical Physical Layer Network Coding for Two-Way Relay Channels: Performance Analysis and Comparison", *IEEE Trans. on Wireless Comm.*, Vol. 9, No. 2, pp. 764–777, Feb. 2010.

[11] F. Yang, M. Huang, S. Zhang, W. Zhou,"Performance Analysis on Two-Way Relay System with Co-Channel Interference", *Wireless Pers. Comm.* (Springer) 72:415–434, Feb. 2013.

[12] M. Rodziewicz, "Network Coding Aided Device-to-Device Communication", 18th *European Wireless Conference*, April 2012.

Energy Efficient Scheduling for LTE-A D2D Communication

Shahid Mumtaz, Kazi M. S. Huq, Ayman Radwan and Jonathan Rodriguez
Instituto de Telecomunica çõ̃es, Aveiro, Portugal
dr.shahid.mumtaz@ieee.org

1. Introduction

In D2D communication, devices are communicating with each other without intermediate nodes (infrastructure less). D2D communication uses cellular spectrum (license band) supported by a cellular infrastructure and promises three types of gain: a) the proximity of user equipments (UE) may allow for extremely high bit rates, low delays and low energy consumption [1-2]; b) the reuse gain implies that radio resources may be simultaneously used by cellular as well as D2D links, tightening the reuse factor even of a resuse-1 system [2]; c) finally, the hop gain refers to using both an uplink and a downlink resource when communicating via the access point in the cellular mode. Moreover, D2D communication may extend the cellular coverage and facilitate new types of wireless peer-to-peer services. D2D is also economical communication because it uses the same pre-existing cellular infrastructure. Moreover, reusing radio resource between D2D users and cellular user (CU) can significantly improve spectrum efficiency. By default, in LTE-A there is no intra-cell interference due to the orthogonality of the subcarriers but this orthogonality will be lost when D2D communication takes place under CU. One of the key aspects of D2D communication is the set of spectrum bands in which D2D communication takes place. Reusing radio resources between a D2D link and cellular mode link can significantly improve spectrum efficiency and minimize energy consumption. However, D2D users may generate high interference to cellular user located in their communication area or in nearby cells if they use the same spectrum with the eNB users for data transmission [1]. It is important to efficiently allocate frequency resources to D2D UEs in order to reduce interference between D2D and cellular mode users when D2D and cellular users' communications utilize the same frequency band. Therefore, in this paper we will present the novel scheduler which assigns the resources in an energy efficient way between D2D and CU users.

2. Energy Efficient Scheduler for D2D

We propose a scheduling policy which reduces the energy cost per bit of the frequency resource block. Most of the conventional schedulers, including Maximum Throughput (MT) and Proportional Fairness (PF), make a decision based on the throughput/QoS and instantaneous channel condition as part of a cross-layer scheduling approach. However, a new factor that should be considered to enhance the system performance is the energy per bit, which until now has been treated scarcely. The transmit energy is insufficient when the radio resources are fully utilized, huge amounts of data are required to be transmitted and most D2D/CU have poor channel conditions [3]. In this case, if the scheduling metric of a packet scheduling scheme considers the ratio of the transmit energy to the number of transmission bits, more improvement in the system performance should be expected [3]. For this reason, in a system with limited transmit energy, it is more efficient to allocate Physical Resource Block (PRBs) to the D2D/CU that require the least ratio of the transmit energy to the number of transmission bits. Thus, in the proposed packet scheduling scheme, the scheduling metric selects the UEs to be allocated in ascending order of the ratio of the to transmit energy, E_u^m, to the number of transmission bits B_u^m, of the PRB m of the UE u as follows:

$$\Psi(u,m) = \arg\min_{u,m} \frac{E_u^m}{B_u^m} = \arg\min_{u,m} \frac{P_u^m gT}{B_u^m} \qquad (1)$$

where $\Psi(u,m)$ is the scheduling metric which denotes the index of selected UE u and PRB m respectively; energy is the multiple of power and time. i.e. $P_u^m gT$, since $P_u^m = \frac{\xi(B_u^m)}{h_u^m}$ [4], we can redefine the metrics given below [28]

$$\arg\min_{u,m} \frac{P_u^m gT}{B_u^m} = \arg\min_{u,m} \frac{\xi(B_u^m) gT}{h_u^m \cdot B_u^m} \qquad (2)$$

where h_u^m is the channel power gain and $\xi(B_u^m)$ is the minimum transmit power required for B_u^m transmitted bits on the PRB m of the UE u. The $\xi(B_u^m)$ required for transmission of B_u^m bits with the target BER of Probability of Error (POE) is given by [3] and [4]

$$\xi(B_u^m) = \frac{\left(\sigma_u^m\right)^2}{3}\left[Q^{-1}\left(\frac{POE}{4}\right)\right]^2\left(2^{B_u^m} - 1\right) \qquad (3)$$

where $\sigma_{u,m}^2$ is the noise variance for the subcarriers in the PRB m at the UE u, and $Q(x) = 1/\sqrt{2\pi} \cdot \int_x^\infty e^{-t/2} dt$.

Let $\hat{P}_u^m$ denote the maximum transmit power at the transmitter that can be assigned for the UE u and the PRB m. Then, the minimum channel gain required for successful transmission of B_u^m bits through the PRB m is given by $h_{\min}\left(B_u^m\right)=\xi(B_u^m)/\hat{P}_u^m$, where $\xi(B_u^m)$ is expressed in equation (3). Since we have $h_u^m\left(B_u^m\right)=\xi(B_u^m)/P_u^m$, the excess channel gain, Ω_u^m (B_u^m) is the maximum positive integer that satisfies $(\Omega_u^m \geq 0)$ is written as

$$\Omega_u^m = h_u^m - h_{\min}(B_u^m) = \xi(B_u^m)\left(\frac{1}{P_u^m}-\frac{1}{\hat{P}_u^m}\right) \tag{4}$$

From equation (4), we get

$$\frac{1}{P_u^m}=\frac{\Omega_u^m}{\xi(B_u^m)}+\frac{1}{\hat{P}_u^m} \Rightarrow P_u^m=\frac{1}{\dfrac{\Omega_u^m}{\xi(B_u^m)}+\dfrac{1}{\hat{P}_u^m}} \tag{5}$$

Using equation (5) in (2), we get

$$\Psi(u,m)=\arg\min_{u,m}\left(\frac{T}{\left(\dfrac{\Omega_u^m}{\xi(B_u^m)}+\dfrac{1}{\hat{P}_u^m}\right) gB_u^m}\right) \tag{6}$$

When the $\hat{P}_u^m$ of the transmitter is too large, (6) can be rewritten as

$$\Psi(u,m)=\arg\min_{u,m}\left(\frac{T}{\left(\dfrac{\Omega_u^m}{\xi(B_u^m)}+0\right) gB_u^m}\right)=\arg\min_{u,m}\left(\frac{T}{\dfrac{\Omega_u^m}{\xi(B_u^m)} gB_u^m}\right) \tag{7}$$

Afterwards,

$$\Psi(u,m)=\arg\min_{u,m}\left(\frac{T}{\dfrac{\Omega_u^m}{\xi(B_u^m)} gB_u^m}\right)=\arg\min_{u,m}\left(\frac{T}{\dfrac{\Omega_u^m}{\left(\xi(B_u^m)/B_u^m\right)}}\right)$$

$$\Psi(u,m)=\arg\min_{u,m}\left(\frac{\left(\xi(B_u^m)/B_u^m\right) gT}{\Omega_u^m}\right)=\arg\min_{u,m}\left(\frac{\left(\bar{E}(B_u^m)/B_u^m\right)}{\Omega_u^m}\right) \tag{8}$$

where, $\bar{E}(B_u^m)=\xi(B_u^m) gT$ → is minimum received energy. Equation (8) can be rewritten as

$$\arg\min_{u,m}\left(\frac{\left(\bar{E}(B_u^m)/B_u^m\right)}{\Omega_u^m}\right)=\arg\max_{u,m}\left(\frac{\Omega_u^m}{\left(\bar{E}(B_u^m)/B_u^m\right)}\right) \tag{9}$$

Because $\arg\min(x)=\arg\max\left(\frac{1}{x}\right)$. Eventually, the scheduling metric can be expressed as

$$\Psi(u,m)=\arg\max_{u.m}\frac{\Omega_u^m}{\left(\bar{E}(B_u^m)/B_u^m\right)} \tag{10}$$

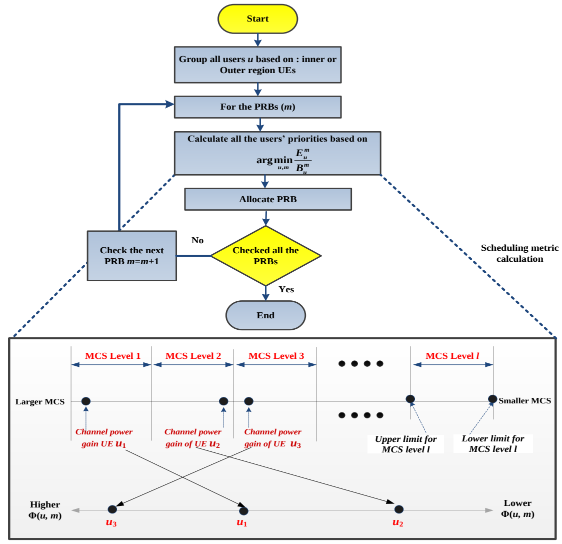


Figure 1: Energy Efficient Scheduling

The proposed energy efficient scheduler allocates the PRB m to the UE with larger excess channel gain which is distant to the required received energy per bit as in equation (10). For the UEs with equal value of excess channel gain, the proposed scheduler allocates the PRB to the UE with the smallest received energy per bit. For example, consider UE u_1, u_2, and u_3 in Figure 1 in an order of MCS level 1, 2, and 3, respectively. According to the MCS levels, the MCS level 1 sends the highest data rate while the MCS level l transmits the lowest data rate. As per 3GPP LTE Adaptive Modulation Coding (AMC) scheme, UE u_1 is able to transmit more bits than UE u_2 but UE u_1 requires smaller transmit energy per bit than UE u_2. It is because the channel power gain of UE u_1 is much larger than the minimum required channel power gain for the MCS level 1 which may require small transmit energy, while the channel power gain for UE u_2 is close to the minimum value for the MCS level 2 which requires larger transmit energy than the other cases. Meanwhile, UE u_3 has almost the same excess channel gain as UE u1, but it requires less received energy per bit, $\left(\bar{E}(B_u^m)/B_u^m\right)$ than UE u_1 because $\left(\bar{E}(B_u^m)/B_u^m\right)$, in (10) increases exponentially with B_u^m; hence, the value of $\left(\bar{E}(B_u^m)/B_u^m\right)$, for UE u_3 is smaller than UE u_1 having higher MCS level than UE u_3. Consequently, the EES scheduler selects the UEs to be allocated in order of UE u_3, UE u_1, and UE u_2.

3. Performance evaluation

Figure 2 and Figure 3 show the performance of proposed scheduling scheme in term of average cell throughput vs. transmit power, where the maximum allowable transmit power for the *outer region* is 46dbm in proposed scheme. The proposed scheme can sustain more than 13Mpbs average cell throughput with 30dbm.

In addition, the proposed scheme can save transmit power in the region around 6dbm more than MT algorithm, while sustaining the same cell throughput.
Similarly, Figure 3 shows the average cell throughput against transmit power, where the maximum allowable transmit power for the *inner region* is 41dbm in our proposed scheme. The proposed scheme can sustain more than 17Mpbs average cell throughput with 20dbm. In addition, the proposed scheme can save transmit power about 8dbm more than the MT algorithm, while sustaining same cell throughput.

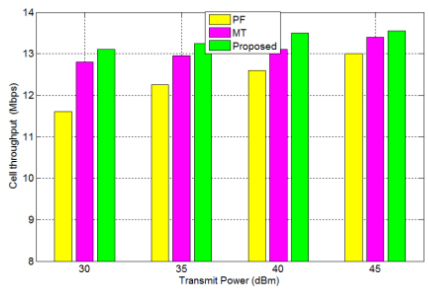


Figure 2: Cell edge Throughput vs. transmit power

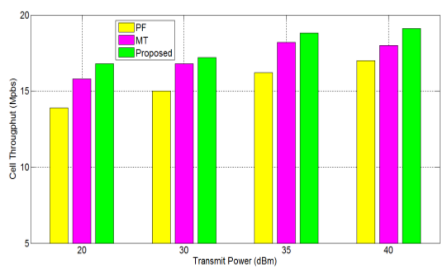


Figure 3: Cell center Throughput vs. transmit power

4. Conclusion

In this paper we presented the novel energy efficient resource allocation scheme for D2D.This scheme assigned the resource in an energy efficient way which reduces the energy cost per bit of the system. Our scheme improved the energy efficiency of the system when compared to SOA scheduler.

References

[1] K. Doppler, M. P. Rinne, P. Janis, C. Ribeiro, and K. Hugl, “Device-to-Device Communications; Functional Prospects for LTE-Advanced Networks”. IEEE International Conference on Communications Workshops, 2009, Jun. 2009, pp.1– 6.

[2] Gábor et. al. “Design Aspects of Network Assisted Device to Device Communication”, IEEE communication Magazine March 2012.

[3] Jungsup Song at.al "Packet-scheduling algorithm by the ratio of transmit power to the transmission bits in 3gpp LTE Downlink" Journal EURASIP Journal on Wireless Communications and Networking - Volume 2010, January 2010

[4] W. Cheong Yui, R. S. Cheng, K. B. Lataief, and R. D. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE Journal on Selected Areas in Communications,* vol. 17, pp. 1747-1758, October, 1999.

SHAHID MUMTAZ received his MSc. degree from the Blekinge Institute of Technology, Sweden and his Ph.D. degree from University of Aveiro, Portugal. He is now a senior research engineer at the Instituto de Telecomunicações - Pólo de Aveiro, Portugal, working in EU funded projects. His research interests include MIMO techniques, multi-hop relaying communication, cooperative techniques, cognitive radios, game theory, energy efficient framework for 4G, position information assisted communication, joint PHY and MAC layer optimization in LTE standard. He is author of several conference, journal and book chapter publications.

KAZI M. S. HUQ received B.Sc. in CSE from Ahsanullah University of Science & Technology, Bangladesh in 2003. He obtained his M.Sc. in EE - specialization Telecommunications from Blekinge Institute of Technology, Sweden in 2006. Since April 2008, he started working at Instituto de Telecomunicações, Pólo de Aveiro, Portugal.

Ayman Radwan received his Master of Applied Science (MASc) from Carleton University (Canada) and his Ph.D. from Queen’s University (Canada), in 2003 and 2009, respectively. He is currently a Senior Research Engineer with the Instituto de Telecomunicações, Portugal. Ayman is mainly managing EU funded projects, targeting next generation sustainable mobile networks.

JONATHAN RODRIGUEZ received his Masters degree in Electronic and Electrical Engineering and Ph.D from the University of Surrey (UK), in 1998 and 2004 respectively. In 2002, he became a Research Fellow at the Centre for Communication Systems Research and was responsible for coordinating Surrey involvement in European research projects under framework 5 and 6. Since 2005, he is a Senior Researcher at the Instituto de Telecomunicações (Portugal), and founded the 4TELL Wireless Communication Research Group in 2008.

Resource Allocation for Device-to-Device Communications Underlying Cellular Networks

Daquan Feng[1], Lu Lu[2], Yi Yuan-Wu[3], Geoffrey Ye Li[2] , Gang Feng[1], Shaoqian Li[1],
[1]National Key Lab on Communications, UESTC, Chengdu, China
[2]School of ECE, Georgia Institute of Technology, Atlanta, GA, USA
[3]Orange Lab Network, Department of Wireless Technology Evolution, Paris, France
{fdquan,lulu0528}@gmail.com, liye@ece.gatech.edu, yi.yuan@orange.com, {fenggang, lsq}@uestc.edu.cn

1. Introduction

As the popularity of smart mobile devices, there is a growing trend for proximity-based applications such as *peer-to-peer* (P2P) file sharing and local multicasting and advertising. To satisfy the increasing demand of local traffic load and provide better user experience, *device-to-device* (D2D) communications have being considered as one of the key techniques in the *Long Term Evolution Advanced* (LTE-Advanced) networks [1-2].

As illustrated in Figure 1, with D2D communications, proximity users in cellular networks can transmit data directly to each other without going through the *base station* (BS). Taking the advantages of the proximity gain, reuse gain, and hop gain [3], D2D communications can significantly improve user experience and resource efficiency. It has been demonstrated in [1] that D2D communications can increase the overall network throughput by up to 65% compared to the typical cellular communications. From [2], with D2D communications, the power consumption of *user equipment* (UE) can be saved by around 70%.

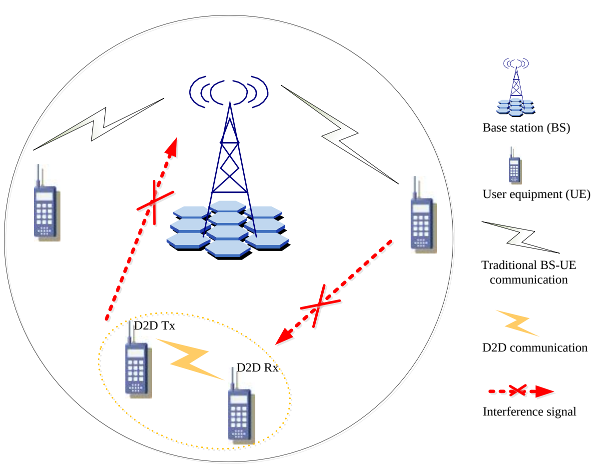


Figure 1. D2D communication underlying cellular networks

D2D communications may bring harmful interference to the existing *cellular users* (CUs) if not designed properly [4-5]. Thus, effective interference management becomes one of the most critical challenges for D2D communications to realize the potential benefits. Various schemes have been proposed to deal with the interference issue, including reusing channel selection [4], distributed resource allocation [5], and transmission mode switching [6].

The aforementioned works have either aimed at increasing the network throughput or guaranteeing the reliability of D2D communications. To consider both metrics simultaneously, a fixed power margin scheme [7] has been developed to coordinate interference between D2D users and regular CUs. However, finding a suitable power margin is not trivial. Resource allocation in D2D communications can be formulated as a *mixed-integer-nonlinear-programming* (MINLP) problem, which can be solved using the heuristic algorithm in [8]. But the algorithm in [8] has not considered the cooperation between the CUs and D2D pairs.

In this article, we will focus on the joint power and channel allocation for D2D communications underlying cellular networks.

2. Optimal Resource Allocation with Perfect CSI

Inspired by the works in [7, 8], we have developed a framework of resource allocation for D2D communications underlying cellular networks in [9], where the overall network throughput of existing CUs and admissible D2D pairs are maximized and at the same time the QoS requirements for both CUs and D2D pairs are guaranteed. Specifically, our algorithm includes three parts. First, a minimum distance metric is proposed for the BS to decide whether a D2D pair can be accessed or not under the minimum SINR requirements for both CUs and D2D pairs. Then, an optimal power control scheme is investigated for each D2D pair and its possible CU partners to maximize the overall throughput. At last, a maximum weight bipartite matching based scheme is developed to determine a specific CU partner for each admissible D2D pair.

Fig. 2 compares the performance of the proposed scheme with the fixed margin scheme in [7] and the heuristic scheme in [8]. From the figure, even in a fully loaded cellular network, up to 70% existing users can access the network as D2D users.

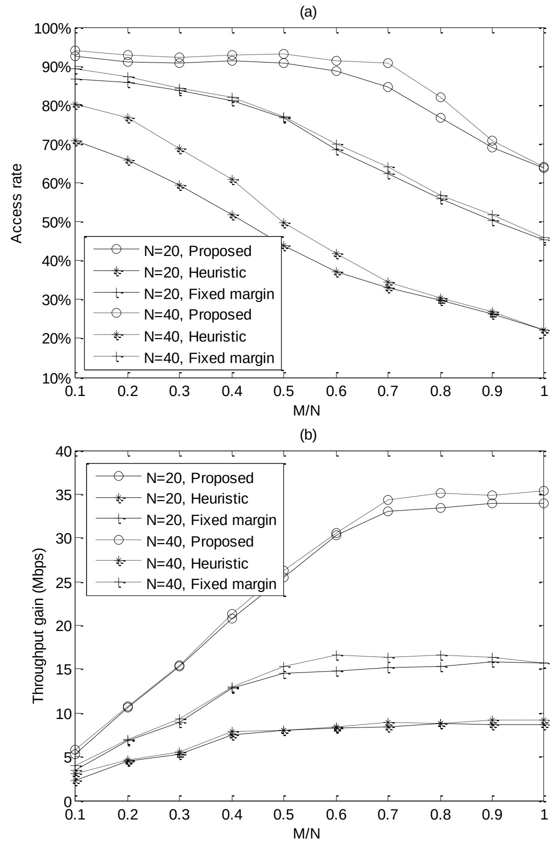


Figure 2. D2D access rate and throughput gain

3. Resource Allocation with Partial CSI

In [9], we have assumed that the BS has perfect *channel-state-information* (CSI) of all links. In general, BS can collect the CSI of users by the classic training and channel estimation. However, obtaining the instantaneous CSI between CUs and D2D pairs is difficult and causes high overhead [10], especially when the number of CUs and D2D pairs are large. Thus, it is more reasonable to assume the BS only has the partial CSI including distance based pathloss and shadowing, which can be obtained by empirical measurements. We have developed resource allocation algorithms in this situation in [11-13].

User selection based limited feedback: In [11], we have proposed an efficient and simple user selection scheme based limited feedback, KMDR, to deal with channel uncertainty. For the KMDR scheme, each D2D receiver only feeds back the CSI of the CUs with the largest *maximum distance ratio* (MDR) metric, which is the ratio of the actual distance between CU and receiver of D2D pair to the required minimum distance to satisfy all the access constraints without considering the fading effect. This approach can save up to 70% CSI feedback while it can still provide a near optimal performance as in [9].

Probability-based resource allocation: In [12], we have developed a probability-based resource allocation scheme by utilizing channel statistical characteristics as an alternative to deal with channel uncertainty. In this scheme, a probabilistic D2D access metric is first used to check whether a D2D pair can be accessed or not with the condition that both its own outage requirement and the partner CU's minimum SINR requirement are satisfied. Then, a joint power and channel allocation strategy is derived to maximize the overall throughput of the admissible D2D pairs and CUs. The proposed probabilistic algorithm can effectively protect the QoS of users and provide a near optimal performance in terms of D2D access rate and throughput gain, especially when the radius of D2D cluster is small.

Combined algorithm: To take the advantage of both the probabilistic and limited feedback strategies, a combined resource allocation algorithm is developed in [13]. In the combined algorithm, the KMDR scheme is first executed to find the *K* most potential partner CUs and to trace and report the CSI of CU-D2D receiver links. And then, the probabilistic strategy is adopted to find the potential partner CUs from the unselected CUs in the KMDR scheme. In this way, the KMDR scheme helps the probabilistic scheme to improve the reliability since more accurate CSI is obtained. Furthermore, the probabilistic scheme can improve the diversity of the KMDR since more potential partner CUs are found. As a result, the combined algorithm outperforms the probabilistic scheme and the limited feedback scheme.

4. Potential Research Topics

Existing research has shown significant benefits of D2D communications. To further exploit the potential gains of D2D communications, more research in the area is desired, including, D2D transmission design for *multi-input and multi-output* (MIMO) systems, D2D assisted *mobile social networks* (MSNs), opportunistic D2D routing in *heterogonous networks* (HetNet)**,** D2D communications for *multiple radio access* (multi-RAT) and full-duplex communications.

5. Conclusions

In this article, we have discussed the resource allocation strategies for D2D communications underlying cellular networks to optimize the overall network performance. Our preliminary research has shown substantial benefits of D2D communications. To further exploit the potential gains, more research still needs to be performed.

References

[1] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42–49, Dec. 2009.

[2] G. Fodor, E. Dahlman, G. Mildh, S. Parkvall, N. Reider et al., "Design aspects of network assisted device-to-device communications," *IEEE Commun. Mag.*, vol. 50, no. 3, pp. 170–177, 2012.

[3] G. Fodor and N. Reider, "A distributed power control scheme for cellular network assisted d2d communications," in *Proc. GLOBECOM'11*, 2011, pp.1–6.

[4] Tao Han, Rui Yin, Yanfang Xu, Guanding Yu, "Uplink channel reusing selection optimization for Device-to-Device communication underlaying cellular networks", in *Proc. IEEE PIMRC'12*, Sept., 2012, pp. 559-564.

[5] Rui Yin, Guanding Yu, Caijun Zhong, Zhaoyang Zhang, "Distributed resource allocation for D2D communication underlaying cellular networks", in *Proc. IEEE ICC Workshop* 2013, June 2013, pp. 138-143.

[6] K. Doppler, C.-H. Yu, C. Ribeiro, and P. J¨anis, "Mode selection for device-to-device communication underlaying an LTE-advanced network," in *Proc. WCNC' 10*, 2010, pp. 1–6.

[7] B. Kaufman and B. Aazhang, "Cellular networks with an overlaid device to device network," in *Proc. IEEE 42nd Asilomar Conf. on Signals, Syst. and Comput.*, 2008, pp. 1537–1541.

[8] M. Zulhasnine, C. Huang, and A. Srinivasan, "Efficient resource allocation for device-to-device communication underlaying LTE network," in *Proc. IEEE WiMob' 10*, 2010, pp. 368–375.

[9] D.-Q. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng and S.-Q. Li, "Device-to-Device communications underlaying cellular networks " , *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3541-3551, 2013.

[10] H. Min, J. Lee, S. Park, and D. Hong, "Capacity enhancement using an interference limited area for device-to-device uplink underlaying cellular networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 3995 –4000, 2011.

[11] D.-Q. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng and S.-Q. Li, "User selection based on limited feedback in Device-to-Device Communications", *Proc. IEEE PIMRC'13*, London, UK, Sept., 2013.

[12] D.-Q. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng and S.-Q. Li, "Optimal resource allocation for Device-to-Device communications in fading channels", *Proc. IEEE GLOBECOM'13*, Atlanta, Dec., 2013.

[13] D.-Q. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, G. Feng and S.-Q. Li, "QoS-aware resource allocation for Device-to-Device communications with Partial CSI", submitted to *IEEE Trans. Wireless Commun.*, 2013.

Daquan Feng is currently pursuing his Ph.D. degree at National key Laboratory on Communications, University of Electronic Science and Technology of China (UESTC), Chengdu, China. His research interests include device-to-device communications, energy-efficient wireless network design, and heterogeneous network.

Lu Lu received her B.S.E degree and M.S.E degree from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 2007 and 2010, respectively. She then got her licentiate degree from Royal Institute of Technology (KTH) in 2011. She is currently working toward the Ph.D. degree with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. Her research interests include MIMO, cooperative communications, and cognitive radio networks.

Yi Yuan-Wu received B.S.E. degreee from Huazhong University, Wuhan, China, in 1982, the Master of Sciences degree in 1983 and Ph.D. in 1987 in signal processing and telecommunication from Rennes University, Rennes, France. Between 1989 and 1991, she worked in Thomson-LGT on video diffusion networks. She joined Orange Labs, Issy-les-Moulineaux France in 1992 and then is senior research expert of Orange Labs since 2004.

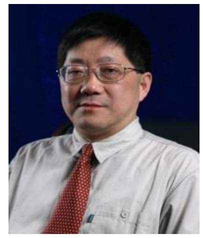
Geoffrey Ye Li received his B.S.E. and M.S.E. degrees in 1983 and 1986, respectively, from the Department of Wireless Engineering, Nanjing Institute of Technology, Nanjing, China, and his Ph.D. degree in 1994 from the Department of Electrical

Engineering, Auburn University, Alabama. He is currently a Professor with the School of ECE at Georgia Institute of Technology.

His general research interests include statistical signal processing and communications, with emphasis on cross-layer optimization for spectral- and energy-efficient networks, cognitive radios and opportunistic spectrum access, and practical issues in LTE systems. He has been awarded an IEEE Fellow for his contributions to signal processing for wireless communications since 2006, and won Stephen O. Rice Prize Paper Award in 2010 and WTC Wireless Recognition Award in 2013 from the IEEE Communications Society and James Evans Avant Garde Award in 2013 from the IEEE Vehicular Technology Society.

Gang Feng received his B.S.E. and M.S.E. from the University of Electronic Science and Technology of China (UESTC), in 1986 and 1989, respectively, and the Ph.D. degrees in Information Engineering from The Chinese University of Hong Kong in 1998. At present he is a professor with National key Laboratory on Communications, UESTC.

Dr. Feng has extensive research experience and has published widely in computer networking and wireless networking research. His research interests include resource management in wireless networks, wireless network coding, energy efficient wireless networking, etc. Dr. Feng is a senior member of IEEE.



Shaoqian Li received his M.S.E. degree in Information and Communication Systems from University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1984 and his B. S.E from Northwest Institute of Telecommunication Engineering (current Xidian University), Xi'an, China, in 1981. He joined UESTC in 1984. Currently, he is Professor of UESTC and Director of National key Laboratory on Communications of UESTC.

Since 1990s, he has served for many domestic academic conferences as Senior member of China Institute of Communications (CIC) and Chinese Institute of Electronics (CIE). And since 2001, he has served for Intl. Conf. on Commun., Circuits, and Systems (ICCCAS), in 2007, in 2008, in 2009, in 2010, in 2012, as Chair of Steering Committee, and/or General (co-) Chair, respectively. He also served as consultant Member of Board of Communications and Information Systems of Academic Degrees committee of the State Council, P. R. China, and Member of expert group of Key Project on Next-Generation Mobile Broadband Wireless Communications Systems towards 2020, Ministry of Industry and Information Technology of P. R. China. He is now IEEE Senior Member.

Infrastructureless Synchronization for Device-to-Device Discovery

Lei Cheng and Yik-Chung Wu
Department of Electrical and Electronic Engineering, The University of Hong Kong
{*leicheng, ycwu*}*@eee.hku.hk*

1. Introduction

Device-to-Device (D2D) communications in cellular network have been studied to provide higher network capacity [1],[2]. As a promising energy efficient proximity-based communication paradigm, D2D communications will likely be adopted in 3GPP Release 12 LTE-Advanced systems. As a result, research on D2D communications has been highly motivated recently.

In addition to the resource management which is extensively studied [3] (and references therein), D2D discovery is another important topic but less studied. Good D2D discovery protocols are supposed to discover the neighborhood devices timely to improve the user experience, work well in highly dense environment and meet energy constraints of mobile devices. However, it is not a trivial task to meet these requirements simultaneously.

Device discovery protocols with base station assistance are proposed in [4],[5]. These approaches mainly utilize multiple accesses techniques, for example, OFDMA, to organize beacon transmissions. However, these methods will occupy additional cellular radio resources and fails to work when the devices are out of the coverage.

These problems in infrastructure-based approaches would vanish if we explore infrastructureless device discovery protocols, which have been investigated in ad-hoc wireless networks [6]-[8]. However, these approaches [6]-[8] work in randomized fashions where a device randomly choose to be active or idle at any one time, which result in high duty cycle and correspondingly, large energy consumptions.

More recently, [9] proposes a promising and easily implementable protocol for device discovery. The key idea of this protocol is to let devices in a neighborhood form a group, provided that they are synchronized beforehand. Devices in a group transmit the beacons in sequence, and the beacons contain the group member list. As the number of devices in a group increases, the devices can transmit the beacon less frequently and sleep for a longer time to save energy. Since the beacons are transmitted in sequence, there is no packet collision, the discovery time can be fast and the protocol is scalable. Simulation results in [9] have verified its good performance and meet requirements mentioned above simultaneously.

However, the prime assumption of [9] is that all the devices are synchronized, which are not easily satisfied in practice. Moreover, the clock of each device may drift from each other due to many factors, such as imperfection of oscillators and environmental changes. Lack of synchronization will lead to failure of the protocol operation. Thus, discovery algorithm without synchronization will be of more interest in practice.

In the following, the key operations of recently proposed synchronized distributed protocol [9] are introduced using an example in Section 2. Then, drawing the ideas from recent advances in clock synchronization in wireless sensor network [10]-[12], possible ways of modification to [9] are introduced in Section 3. Finally, conclusions are drawn in Section 4.

2. Distributed Protocol for Device Discovery

Assume devices B and A have formed a group and accordingly generate a group list (BA). Beacon will be transmitted in turn by different device every period T, and follow the order defined by the group list, as seen in Fig.1. For each device, before beacon transmission, there is a "join window", which accepts join request from new devices. Also, each device will listen to the beacon from the previous device defined in the group list, and is called "update window".

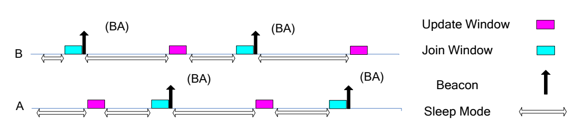


Fig.1 A Beacon transmission pattern for 2 devices.

Now, if C wants to join the group, it scans for T and assume it receives the beacon from B. Then, it sends out "join request" a little bit before the next T, which is the "join window" of A. Device C and A will perform a three-way handshake as shown in Fig.2.

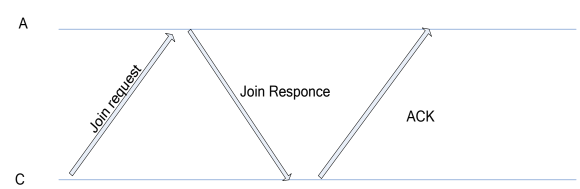


Fig.2 The procedures of a three-way handshake.

If C is accepted by A, device C then schedules its beacon transmission time after (N-1)T, where N is the total number of group members in the new list. On the other hand, device A sends out the beacon with updated group list BCA, and adjust its transmission schedule. B will receive this new beacon, and updates its transmission schedule accordingly. The procedures described above are shown in Fig.3. If a new device D wants to join the group (BCA), the operations are carried out similarly.

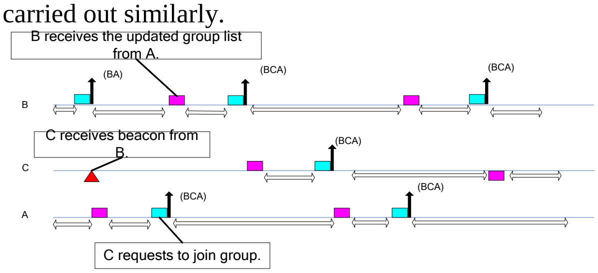


Fig. 3 The procedures of a device joining a group.

However, if a new device is not synchronized to the existing group, it will miss the join request window. Moreover, this protocol is susceptible to clock drift of each device in D2D network, otherwise some devices will miss the update window and the protocol will fail to work in long term.

3. Joint Synchronization and Device Discovery

Firstly we introduce the clock model of each device as

$$C(t) = \theta + f \cdot t \qquad (1)$$

where the parameter θ and f are called clock offset (phase difference) and clock skew (frequency difference), respectively. The global synchronization for a group requires all the devices' clocks are the same as a nominated reference device's clock. From (1), the clock relationship between any device (for example, device B) and the reference device (device A) can be represented by

$$C_B(t) = \theta^{AB} + f^{AB} \cdot C_A(t),$$

where θ^{AB} and f^{AB} stand for the relative clock offset and skew between two devices. Thus, the task of pairwise synchronization is to estimate θ^{AB} and f^{AB} so that device B can adjust its clock accordingly. And the global synchronization is equivalent to estimating the relative clock offset and skew at each device with respect to the reference device A.

In practice, the measurements collected for estimating clock offsets and skews are obtained by timing message delivery among devices. There are three well-known timing message signaling approaches: the two-way message exchange, the one-way message dissemination and the receiver-receiver synchronization. Details of these approaches can be referred in [10]. All of these three message signaling approaches face the same challenge: message delivery will experience random delay, which is modeled by a probabilistic density function (pdf) model such as Gaussian, exponential, and so on.

It is seen that from a signal processing viewpoint, clock synchronization among devices can be interpreted as estimating parameters from measurements corrupted by noise (random delay) with a known or unknown pdf. For pairwise synchronization, the maximum likelihood estimator (MLE) and the corresponding Cramer-Rao bound (CRB) for joint skew and offset estimation under Gaussian random delays were derived in [11]. Besides optimal MLEs, suboptimal but lower complexity algorithms were also reported in [11]. For network-wide synchronization, [12] propose a fully distributed algorithm based on belief propagation (BP) which has low overhead and can achieve scalable synchronization with better accuracy than consensus algorithms.

It is desired to incorporate clock synchronization and maintenance into device discovery protocols such as the one introduced in Section 2. In particular, pairwise synchronization algorithms can be integrated into the join procedure and network-wide synchronization algorithms can be explored to maintain synchronization in the group.

Take the join procedures depicted in Fig.2 and Fig.3 as an example. Assuming the devices in an existing group are already synchronized, and it is the device that wants to join the group is not synchronized. If device C wants to join the group, it is required that device C is synchronized to the group clock for predicting the next join window. If time-stamps can be embedded into the beacons sent from different devices, device C can execute a pairwise synchronization algorithm to adjust its clock after receiving a few beacons from devices in group. Notice that in this case, the message signaling approach is one-way message dissemination. The corresponding estimation algorithms can be obtained by simplifying the existing results in the context of two-way message change mechanism [10] [11].

Notice that in the listening of beacons, device C only requires to collect a few time-stamps such that it can predict the next join window. Then, more accurate synchronization can be carried out during the three-way handshake as shown in Fig. 2. In particular, the three-way handshake can be combined with a two-way message exchange mechanism, by having a few more message exchanges.

For the protocol in Section 2 [9], each device can only hear the previous device in the group list, which results in a simple tree structure of the network, as shown in Fig 4. In this case, pairwise synchronization is known to have rapid error accumulation, and devices far away from the reference device may have inaccurate clocks.

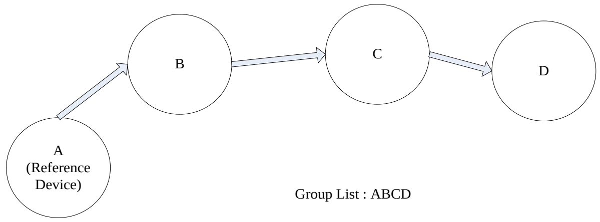


Fig. 4 Tree Structure of the network

In order to further improve the clock accuracies, the beacons from devices can also embed clock offset and skew estimates. Then, BP-based algorithm can be executed to refine the clock estimate in different devices. Furthermore, if we relax the requirement that each device only listen to a single device, the network would become more connected, and the synchronization performance can be enhanced [12]. Notice that the BP-based algorithm works even in a sparely connected networks, which might help to overcome the fundamental limitation of [9] that each device needs to be able to listen to all other devices.

4. Conclusion

In this letter, importance of device-to-device discovery and an example protocol are reviewed. One basic assumption in device discovery is the prior synchronization among devices, which might not be possible in practice. Modifications inspired by recent advances in clock synchronization in wireless sensor network are presented to improve the protocol's practicability and reliability. Other synchronized discovery protocols may also find ingredients from [10]-[12] valuable in supporting their assumptions.

References

[1] K. Doppler et al., "Device-to-Device Communication as an Underlay to LTE-Advanced Networks," IEEE Commun. Mag., vol. 7, no. 12, 2009, pp. 42–49.

[2] X. Lin, J. G. Andrews, and A. Ghosh, "A comprehensive framework for device-to-device communications in cellular networks," submitted to IEEE Journal on Selected Areas in Communications, May 2013. Available at arXiv preprint arXiv:1305.4219

[3] A. Asadi, W. Qing, and M. Vincenzo, "A Survey on Device-to-Device Communication in Cellular Networks." Submitted to IEEE Communications Surveys and Tutorials Aug. 2013. Available at arXiv preprint arXiv:1310.0720 (2013).

[4] K. Doppler, et al. "Advances in d2d communications: energy efficient service and device discovery radio," in IEEE Wireless VITAE, 2011, pp. 1–6.

[5] F. Baccelli, N. Khude, R. Laroia, J. Li, T. Richardson, S. Shakkottai,S. Tavildar, and X. Wu, "On the design of device-to-device autonomous discovery," in IEEE COMSNETS, 2012, pp. 1–9.

[6] M. J. McGlynn and S. A. Borbash, "Birthday protocols for low energy deployment and flexible neighbor discovery in ad hoc wireless networks," in ACM MobiHoc, 2001, pp. 137–145.

[7] S. Vasudevan, J. Kurose, and D. Towsley, "On neighbor discovery in wireless networks with directional antennas," in IEEE INFOCOM, vol. 4, 2005, pp. 2502–2512.

[8] S. A. Borbash, A. Ephremides, and M. J. McGlynn, "An asynchronous neighbor discovery algorithm for wireless sensor networks," Ad Hoc Networks, vol. 5, no. 7, pp. 998–1016, 2007.

[9] P. Huang, et al. "Energy efficient and scalable device-to-device discovery protocol with fast discovery," in Sensor, Mesh and Ad Hoc Communications and Networks (SECON), 2013 10th Annual IEEE Communications Society Conference on. IEEE, 2013.

[10] Y.-C. Wu, Q. Chaudhari, and E. Serpedin, "Clock synchronization of wireless sensor networks," IEEE Signal Processing Magazine, vol. 28, no. 1, pp. 124 –138, 2011.

[11] M. Leng and Y.-C. Wu, ''On clock synchronization algorithms for wireless sensor networks under unknown delay,'' IEEE Trans. Veh. Technol., vol. 59, no. 1, pp. 182–190, Jan. 2010.

[12] J. Du and Y.-C. Wu, "Distributed Clock Skew and Offset Estimation in Wireless Sensor Networks: Asynchronous Algorithm and Convergence Analysis," IEEE Trans. on Wireless Communications, Vol. 12, no. 11, pp. 5908-5917, Nov. 2013.

Lei Cheng received the B.Eng. degree in 2013 from Zhejiang University (ZJU). He was a visiting student at North Carolina State University (NCSU), in summer, 2012. Currently, He is pursuing his Ph.D. degree at the University of Hong Kong (HKU). His research interests are in distributed signal processing with applications to wireless communications.

Yik-Chung Wu received the B.Eng. (EEE) degree in 1998 and the M.Phil. degree in 2001 from the University of Hong Kong (HKU). He received the Croucher Foundation scholarship in 2002 to study Ph.D. degree at Texas A&M University, College Station, and graduated in 2005. From August 2005 to August 2006, he was with the Thomson Corporate Research, Princeton, NJ, as a Member of Technical Staff. Since September 2006, he has been with HKU, currently as an Associate Professor. He was a visiting scholar at Princeton University, in summer 2011. His research interests are in general area of signal processing and communication systems, and in particular distributed signal processing and communications; optimization theories for communication systems; estimation and detection theories in transceiver designs; and smart grid. Dr. Wu served as an Editor for IEEE Communications Letters, is currently an Editor for IEEE Transactions on Communications and Journal of Communications and Network.

Research Progress on Device-to-Device Communication in China

[1]*Lingyang Song*, [2]*Jia Shen, and* [3]*Shaohui Sun*
[1]*School of Electrical Engineering and Computer Science, Peking University, Beijing, China*
[2] *China Academy of Telecommunications Research, Ministry of Industry and Information Technology, China*
[3]*State Key Laboratory of Wireless Mobile Communications, Datang Telecom technology and industry group, Beijing, China*
{lingyang.song@pku.edu.cn, shenjia@mail.ritt.com.cn, sunshaohui@catt.cn}

1. Introduction

As more and more new mobile multimedia rich services are becoming available to mobile users, there is an ever increasing demand for higher data rate wireless access. Consequently, new wireless technologies such as LTE and WiMAX have been introduced, which are capable of providing high speed, large capacity, and guaranteed quality-of-service (QoS) mobile services [1]. However, these centralized network topology based technologies are inherently limited by the capability of base station (BS) and access point, which they could be congested due to a large number of communicating devices. Recently, the concept of Device-to-Device (D2D) communication has been introduced to allow local peer-to-peer transmission among mobile devices without bypassing the base station and access point [2].

The D2D communication can underlay or overlay to a cellular network, using the same resources to improve the system throughput [3]. The mobile station (MS) in D2D connections remains loosely controlled by the BSs in a network controlled manner [4]. The BSs can control the resources used for the cellular and the D2D links. It has been postulated that D2D communication will become a key feature supported by next generation cellular networks, which has the following advantages: extend coverage [2]; offload in cellular networks [3]; improve energy efficiency [6]; increase throughput and spectrum efficiency [4]; create of new services such as mobile social/vehicular ad-hoc networking services, etc.

In China, the D2D related research has also attracted a great deal of attention from numerous researchers and wireless engineers in both academia and industries. In the universities, research topics regarding D2D communication cover a wide range from physical to MAC layer, and upper layer, etc. While for industries, people mainly look at the possibilities of standardization in 3GPP as well as real implementation and prototypes, and their current focus is on neighborhood discovery and public safety applications. Almost in every local wireless communication conferences and seminars, there will be a few presentations about D2D communication.

Meanwhile, many national research projects have been approved to support D2D communication as an underlay to cellular networks. In China, there are three major national research funding agencies: National Nature Science Foundation of China (NSFC), Ministry of Industry and Information Technology (MIIT), and Ministry of Science and Technology (MOST), supporting a number of national and international research projects. These governmental agencies have open grant call on a yearly basis, and make careful selection from a huge number of proposals. For NSFC, they typically call for funding proposals without trying to restrict the research topics or scope, and thus, several D2D projects have been granted on various topics to support university teachers to do basic research and student training. Comparatively, MIIT and MOST typically give specific topics: MIIT mainly supports industrial oriented research, or the ones that have the potential to be standardized; MOST has many programs, e.g. 863 and 973 plans, to support large research projects, and the first one focuses on implementation, and the later one cares more on fundamental research that are critical to the national long-term development.

The aim of this article is present the research progress on D2D communication in China. We will first briefly provide a survey on the national D2D research activities, and later on introduce two main research projects supported by MIIT and MOST.

2. Research Activities on D2D Communication

Academic Research

The researchers in China worked on D2D communication at almost the same time when the term was first invented. Since 2010, they have published nearly 100 research papers in the IEEE Xplore, and these research papers cover a wide range of topics, such as mobile-to-mobile channel modeling, device discovery and association, link management and mode selection, interference management, radio resource management, mobility measurements, modeling and management, multi-hop D2D communications, and game theory and incentive mechanisms.

Industrial Research

Most industries doing D2D communication are of the

3GPP partners, such as the operators (China Mobile, China Telecom, and China Unicom), and the vendors (Huawei, ZTE, and Datang). They are submitting technical documents to 3GPP directly before every radio access network (RAN) meeting. In China, there is also a special interest group on D2D communication for IMT-2020 Promoting Group of MIIT formed by the researchers and engineers in both universities and industries, trying to produce a white paper on this topic, whose target is for summarizing the key technologies and ideas for future standardization and implementation.

China Communications Standards Association has also in the recent few years initialized the standardization of the dedicated short-range communications (DSRC) for vehicular safety applications by using the concept of D2D communication. Meanwhile, a two-year project for vehicular ad-hoc networking services has been kicked off in the middle of 2013 to study the TD-LTE based DSRC technologies.

3. Research Projects on D2D Communication

NSFC, MIIT, and MOST have supported a number of D2D related research projects. In this section, we mainly introduce two projects, device-to-device direct communication, granted by MIIT, and device-to-device local area networks (D2D-LAN) supported by MOST.

Device-to-Device Direct Communication

In 2011, MIIT announced an open call in the framework of National Science and Technology Major Project on D2D communication, and supported three projects on this topic, most of which are led by industries, and the universities can be involved as members. The duration of every project is two year (2012-2013). The aim of these projects is to provide feasible technical solutions for D2D communication in order to extend coverage, realize mobile relay, improve capacity, and decrease network construction cost for IMT-Advanced evolution. Particularly, the network controlled one-hop between devices transmission is the main focus, and flexible mobile relay is another innovative feature. The output requirements include national and international patents, 3GPP technical documents, research articles, key technologies report, simulation tools for performance evaluation, and prototypes for demonstration. This is the final year of these projects.

Device-to-Device Local Area Networks

MOST recently supports a five year fundamental research project belonging to 973 programs (2013-2017), named coordinated heterogeneous device-to-device local area networks underlaying cellular networks (D2D-LAN) [11]. This D2D-LAN project is led by Peking University, and the other two partners are from Shenzhen University, and Research Institute of Chinese Hong Kong University in Shenzhen. As a way forward to current research on D2D communication, this project aims to design architectures and protocols of seamless working for mobile ad-hoc networks operating in the licensed band of cellular networks.

In an effort to achieve performance improvement by combining cellular and mobile ad-hoc networks, D2D-LAN project comprises of three major work packages (WPs) in order to cover physical, MAC, networking, to system layers: dynamic heterogeneous underlaying configuration, spectrum efficient and interference-aware transmission, and cross-layer resource management. As interference is the main limitation in the deployment of underlaying networks, special attention is given to those problems regarding spectrum allocation and interference avoidance. In particular, D2D-LAN uses recent advances in wireless communications, including the wireless network coding, utility optimization, congestion and access control, game theory, physical-layer security, etc, to guarantee the network performance, exhibited vast potential for achieving high capacity and full exploitation of communication resources. The results obtained so far through intensive collaboration among the project partners were rather encouraging in comparison with relevant state-of-the-art approaches and thus pave the way to further study of more composite protocols in the future.

The core objective of D2D-LAN project is to build the gap between theory and practice, from physical, MAC, networking to application layers, through the adoption of network optimization and signal processing for designing architectures and protocols for wireless cellular networks in order to meet the data traffic surge problems. With regards to the system configuration, we primarily focus on the TDD system which allows the D2D users to use the same spectrum in both uplink and downlink. Specifically, the inter-networking between D2D and cellular networks will be carefully considered. Approaches for operators to control the whole networks will be proposed. In addition, specific protocols, such as mode switch, device discovery and synchronization, mobility managements, security, etc, will be discussed.

One of the major challenges by enabling D2D-LAN communication is to realize efficient data spreading in the D2D network without causing severe disturbance of the original cellular networks. Power control, cooperative transmission, and multiple access methods need to be carefully investigated. The use of wireless network coding has been recognized as an efficient way to improve the network performance in term of spectrum efficiency, and thus, how to adopt this technique is worthy of intensive study. Besides, with the introduction of the device freedom, the optimization problem of OFDMA networks becomes more complicated, where multiple D2D users are allowed to coexist with the cellular user as an underlay. Hence, how to effectively coordinate space, time, frequency, power, and device becomes quite interesting. Energy efficiency is another issue that needs to be addressed in the context of D2D communication due to relatively shorter communication range.

Other challenges to be extensively researched include: identification of services for which D2D communication is useful, radio resource allocation and resource management for D2D links, self-organizing D2D links, and capacity and performance evaluation. Finally, many applications, such as mobile social networks, vehicular ad-hoc networks, or even machine-type communications, will be studied by considering specific constraints.

Conclusions

This article mainly reported the recent research progress on D2D communication in China, which indeed has attracted lots of attention from both universities and industries. In particular, two major projects, i.e. D2D direct, and D2D-LAN, supported by MIIT and MOST are introduced for standardization and fundamental research purposes, respectively. All these indicate that Chinese have shown great interests, and taken much effort in this exciting area.

References

[1] L. Song and J. Shen, *Evolved Network Planning and Optimiation for UMTS and LTE*, Auerbach Publications, CRC Press, 2010.

[2] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 42-49, Dec. 2009.

[3] C. Xu, L. Song, Z. Han, D. Li, and B. Jiao, "Resource allocation using a reverse iterative combinatorial auction for device-to-device underlay cellular networks," *IEEE Globe Communication Conference (Globecom)*, Los Angles, USA, Dec. 2012.

[4] P. Janis, V. Koivunen, C. Ribeiro, J. Korhonen, K. Doppler, and K. Hugl, "Interference-aware resource allocation for device-to-device radio underlaying cellular networks," *IEEE Vehicular Technology Conference 2009-Spring*, Apr. 2009.

[5] F. Wang, L. Song, Z. Han, Q. Zhao, and X. Wang, "Joint scheduling and resource allocation for device-to-device underlay communication," *IEEE Wireless Communications and Networking Conference (WCNC)*, Shanghai, Beijing, Apr. 2013.

[6] F. Wang, C. Xu, L. Song, Z. Han, and B. Zhang, "Energy-efficient radio resource and power allocation for device-to-device communication underlaying cellular networks," *IEEE International Conference on Wireless Communications and Signal Processing*, China, Oct. 25 - 27, 2012.

[7] Device-to-device local area networks (D2D-LAN) http://wireless.pku.edu.cn/home/songly/973project/

Lingyang Song (S'03-M'06 -SM'11) received his PhD from the University of York, UK, in 2007, where he received the K. M. Stott Prize for excellent research. He worked as a postdoctoral research fellow at the University of Oslo, Norway, and Harvard University, until rejoining Philips Research UK in March 2008. In May 2009, he joined the School of Electronics Engineering and Computer Science, Peking University, China, as a full professor. His main research interests include cooperative and cognitive communications. He is currently on the editorial board of IEEE Transactions on Wireless Communications. Dr. Song is the recipient of 2012 IEEE Asia Pacific (AP) Young Researcher Award, and 2012 NSFC Outstanding Young Investigator Award.

Jia Shen has been working with China Academy of Telecommunications Research (CATR) of Ministry of Industry and Information Technology (MIIT), China since 2004 when he received his PhD from the University of York, UK. Now he is a director engineer of CATR on 4G, 5G and M2M standardization, tests and trials. He ever joined IEEE 802.16e standardization in 2004, and then has been a standardization delegate in 3GPP RAN for LTE/LTE-Advanced since 2005. During 2007~2009, he acted as the vice chair of Technology WG of IMT-Advanced Promoting Group of MIIT. Since 2008, as the leader of Test/Specification SWG of TD-LTE Working Group of MIIT, he has served as the technical organizer of large-scale TD-LTE tests and trials in China.

Shaohui Sun received his Ph.D. from Xidian University, Xi'an, China, in 2003. Form March 2003 to June 2006, he was a postdoctoral research fellow at the Datang Telecom technology and industry group, Beijing, China. From June 2006 to December 2010, he worked at the Datang mobile communications equipment company Ltd., Beijing, where he has been deeply involved in the development and standardization of the Third-Generation Partnership Project Long-Term Evolution (3GPP LTE).Since January 2011, he has been the Chief Technical Officer with Datang Wireless Mobile innovation Center of the Datang Telecom technology and industry group. Now, his research area of interest includes multiple antenna technology, heterogeneous wireless networks, and relay.

MMTC OFFICERS