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

MMTC Communications – Review

Vol. 8, No. 2, April 2017

IEEE COMMUNICATIONS SOCIETY

TABLE OF CONTENTS

Message from the Review Board Directors

Welcome to the April 2017 issue of the IEEE ComSoc MMTC Communications – Review.

This issue comprises five reviews that cover multiple facets of multimedia communication research including image quality assessment, economical management of quality of experience, etc. These reviews are briefly introduced below.

The first paper, published in IEEE Transactions on Multimedia and edited by Lifeng Sun, presents a blind image quality assessment methodology.

The second paper, published in IEEE International Conference on Quality of Multimedia Experience and edited by Wei Wang, looks into QoE management policies and interactions between content provider and ISPs.

The third paper, published in IEEE Transactions on Multimedia and edited by Qing Yang, discusses a game theoretical solution for social media cloud.

The fourth paper, published in IEEE Transactions on Multimedia and edited by Dongdong Zhang, explores the web-based image dataset reconstruction.

The fifth paper, published in ACM Transactions on Multimedia Computing, Communications and Applications and edited by Carsten Griwodz, investigates a new form of multimedia human tongue taste.

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

IEEE ComSoc MMTC Communications – Review Directors

Pradeep K. Atrey
State University of New York at Albany, USA
Email: patrey@albany.edu

Qing Yang
Montana State University, USA
Email: qing.yang@montana.edu

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

Jun Wu
Tongji University, China
Email: wujun@tongji.edu.cn

Blind Image Quality Assessment: A New Model Based On Free Energy Principle

A short review for "Using Free Energy Principle For Blind Image Quality Assessment"

Edited by Lifeng Sun

Over billions of digital photographs were captured each year, and this number is increasing annually. A natural problem is that visual quality of such a great amount of photographs is hard to guarantee. Hence, the systems to automatically monitor, control and improve the visual quality of digital photographs are highly desirable [1]. Due to its capability of simulating human visual perception to visual quality, image quality assessment (IQA) is usually deployed to solve this problem. Very recently, it was found that IQA-based image/video processing can be widely used in automatic enhancement [2], transmission [3], screen content video compression [4], tone mapping [5], and more.

No-reference (NR) IQA metrics have attracted much more attention since in most application scenarios, lossless reference images are not accessible. Traditional NR IQA metrics are distortion-specific, which assume the distortion types are known and fixed. They were proposed to succeed in evaluating the quality of noisy, compressed, blurred and contrast-distorted images. More concentrations have been recently shifted to general-purpose NR IQA algorithms [6], [7] and [8]. General-purpose NR IQA models were developed mainly by extracting effective features from distorted images followed by training a regression module using those features. Those features were usually extracted based on the natural scene statistics (NSS) model.

In this paper, the authors design an NR Free Energy based Robust Metric (NFERM) using the recently revealed free energy based brain theory and classical human visual system (HVS) inspired features. The used features can be classified into three groups. The first one includes 13 features of the free energy and the structural degradation information. The free energy feature comes from the RR free energy based distortion metric (FEDM) [9], which defines the psychovisual quality as the agreement between an input image and the output of internal generative model, while the structural degradation information is computed by the RR structural degradation model (SDM) [10] that amends SSIM with itself. Although the two RR IQA metrics still require partial reference information, the free energy feature and the structural degradation information of original images are found to be of an approximate linear relationship. When artifacts are injected, the above-mentioned near-linear relationship will be broken, and the associated variations can be measured to predict the visual quality degradation and thus blindly estimate the quality of an input image.

Furthermore, the free energy theory reveals that the human visual system (HVS) always attempts to reduce the uncertainty based on the internal generative model when perceiving and understanding an input visual stimulus. For illustration consider the following example that human brains can automatically restore a noisy image. The authors apply the linear autoregressive (AR) model to approximate the generative model to predict an image that the HVS perceives from an input distorted one. Then, six important HVS inspired low-level features (e.g. structural information and gradient magnitude), which are computed from the distorted and predicted images, constitute the second group of features. The third group of four features arises from the NSS model. The authors estimate the possible losses of "naturalness" in the distorted image by fitting the generalized Gaussian distribution to mean subtracted contrast normalized (MSCN) coefficients. An important note is that, with free energy principle and image scene statistics, this paper links FR, RR and NR IQA together, and proposes a general model for higher performance via a proper fusion of existing full-, reduced and no-reference IQA models.

To summarize, the main contributions of their paper are:

- A new NSS model has been constructed based on the free energy principle. Different from the previous work, the proposed NSS model not only widely occurs on a broad range of image

scenes, but also can be explained by the recently revealed free energy based brain theory.

● Based on the proposed NSS model, a free energy-based no-reference IQA metric has been proposed. Through extensive experiments, the proposed quality metric has been proven to be of superior performance as compared with state-of-the-art blind distortion-specific and general-purpose IQA metrics on six commonly used image quality databases.

Acknowledgement:

The R-Letter Editorial Board thanks the authors of the paper for providing a summary of its contributions.

References:

[1] A. C. Bovik, "Automatic prediction of perceptual image and video quality," Proc. IEEE, Sep. 2013.
[2] K. Gu, G. Zhai, X. Yang, W. Zhang, and C. W. Chen, "Automatic contrast enhancement technology with saliency preservation," IEEE T-CSVT, Sep. 2015.
[3] H. R. Wu, A. Reibman, W. Lin, F. Pereira, and S. S. Hemami, "Perceptual visual signal compression and transmission," Proc. IEEE, Sep. 2013.
[4] S. Wang, K. Gu, K. Zeng, Z. Wang, and W. Lin, "Objective quality assessment and perceptual compression of screen content images," IEEE Computer Graphics and Applications.
[5] K. Gu, S. Wang, G. Zhai, S. Ma, X. Yang, W. Lin, W. Zhang, and W. Gao, "Blind quality assessment of tone-mapped images via analysis of information, naturalness and structure, " IEEE T-MM, Mar. 2016.
[6] A. K. Moorthy and A. C. Bovik, "Blind image quality assessment: From scene statistics to perceptual quality," IEEE T-IP, Dec. 2011.
[7] M. A. Saad, A. C. Bovik, and C. Charrier, "Blind image quality assessment: A natural scene statistics approach in the DCT domain," IEEE T-IP, Aug. 2012.

[8] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," IEEE T-IP, Dec. 2012.
[9] G. Zhai, X. Wu, X. Yang, W. Lin, and W. Zhang, "A psychovisual quality metric in free-energy principle," IEEE TIP, Jan. 2012.
[10] K. Gu, G. Zhai, X. Yang, and W. Zhang, "A new reduced-reference image quality assessment using structural degradation model," IEEE ISCAS, May 2013.
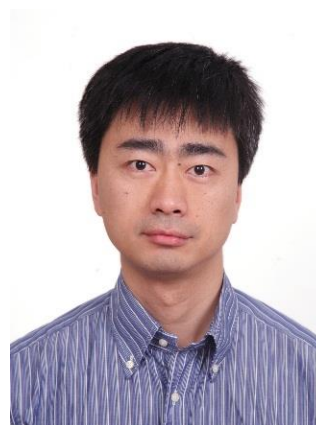
Lifeng Sun received the B.S. and Ph.D. from National University of Defense Technology China, in 1995 and 2000, respectively. He joined the Department of Computer Science and Technology (CST), Tsinghua University (THU), Beijing, China, in 2001. Currently, he is a full Professor in CST of THU. His research interests include networked multimedia, video streaming, 3D/multi-view video coding, multimedia content analysis, multimedia cloud computing and social media. He has authored or coauthored about 100 high quality papers, one chapter of a book. He got the Annual Best Paper Award of IEEE TCSVT (2010), Best Paper Award of ACM Multimedia (2012) and Best Student Paper Award of Multimedia Modeling (2015).

QoE-Centric Multimedia Delivery Collaboration Between Content And Network Providers

A short review for "QoE-aware Service Delivery: A Joint-Venture Approach for Content and Network Providers"

Edited by Wei Wang

The total data traffic over internet is changing from traditional web page traffic to multimedia traffic, due to the popularity of mobile smartphone content generators and various multimedia services across the internet [1]. In fact, the Quality of Experience (QoE) becomes an essential issue in wireless multimedia edge services in the cloud and the fogs [2], due to the popularity of user-generated multimedia contents, aka. self-medias.

The authors in this paper proposed a new approach for collaborative QoE management between Over The Top (OTT) content providers and Internet Service Providers (ISPs). One major contribution of this article is the QoE-centric OTT-ISP management, which is fundamentally different from traditional Quality of Service (QoS)-centric policies.

The authors proposed the QoE management model considering several important factors, such as the user churn, pricing and marketing. The authors also modeled the revenue, and determined whose maximization drove the collaboration between the OTT and ISP.

Usually ISPs are not in the loop of revenue generation between the content provide OTTs and the end users consuming the multimedia contents. However, the ISPs face the user churn risks when the service quality does not meet the users' expectation levels. In fact, both OTTs and ISPs are financially affected by the user churn reactions.

Recent research works have identified that QoE modeling and evaluation is complex and hard, which not only include network service and packet delivery, but also multimedia application configuration settings and users' subjective aspects [3]. Both the applications and the providers are key players involved in QoE delivery.

Additionally, users' subjective aspects should also be included in this consideration. Quality and pricing are considered as major factors causing user churn [4]. For example, if user A pays a higher price to secure a channel for high quality movie, and user B pays a lower prices for a best-effort service. It is very possible that user A gives a higher chance of churn reaction when the network throughput service is not stable. Therefore, QoE and pricing are not isolated factors when considering OTT and ISP providing media service to end user consumers [5] [6] [7].

The authors referred to Paris Metro pricing model [8] to start their revenue modeling, with higher pricing indicating higher expected QoE. The revenue was modeled as the total number of users multiplies by the unit price of a specific service class that user chose. The authors also includes the user churn dynamics in the modeling, letting the number of users in the current iteration multiply by a user churn function, and finally adding any new users who recently join this service class.

The user churn function was modeled as a sigmoid function with a special factor denoting "half of paid users leave this service level". A factor of sensitivity was also introduced with regards to the paid price, representing the fact that users who paid more typically expect a higher service quality. The target revenue maximization was performed during a reference period of time. The Yamagishi video model [9] was used in their QoE modeling.

The authors also acknowledged an important factor of Network Neutrality (NN) when they presented their models. In order to provide the

openness in the internet, the users should have equal access to the contents, and the ISP should avoid discriminating the users from any application providers or high profile payers. Therefore, the NN is an important factor to keep in mind when designing the QoE management policies.

References:

[1] "Cisco visual networking index: Global mobile data traffic forecast update, 2015-2020," white paper, 2016.

[2] W. Wang, Q. Wang, K. Sohraby, ``Multimedia Sensing as a Service (MSaaS): Exploring Resource Saving Potentials of at Cloud-Edge IoTs and Fogs,'' IEEE Internet of Things Journal, in press.

[3] "Qualinet white paper on definitions of quality of experience," 2013.

[4] B.-H. Chu, M.-S. Tsai, and C.-S. Ho, "Toward a hybrid data mining model for customer retention," Knowledge-Based Systems, vol. 20, no. 8, pp. 703–718, 2007.

[5] I. Abdeljaouad and A. Karmouch, "Monitoring iptv quality of experience in overlay networks using utility functions," Journal of Network and Computer Applications, vol. 54, pp. 1–10, 2015.

[6] W. Dai, J. W. Baek, and S. Jordan, "Modeling the impact of qos pricing on isp integrated services and ott services," in Network and Service Management (CNSM), 11th Int. Conf. on. IEEE, 2015, pp. 85–91.

[7] S. Wahlmueller, P. Zwickl, and P. Reichl, "Pricing and regulating quality of experience," in Next Generation Internet (NGI), 2012 8th EURO-NGI Conference on. IEEE, 2012, pp. 57–64.

[8] A. Odlyzko, "Paris metro pricing for the internet," in Proceedings of the 1st ACM conf. on Electronic commerce. ACM, 1999, pp. 140–147.

[9] K. Yamagishi and T. Hayashi, "Parametric packet-layer model for monitoring video quality of iptv services," in Communications, 2008. ICC'08. IEEE International Conference on. IEEE, 2008, pp. 110–114.

Wei Wang is an Assistant Professor of Computer Science at San Diego State University. He received his Ph.D. degree in Computer Engineering with Significant Computer Science Emphasis from University of Nebraska - Lincoln, NE, USA. He worked as a Software Engineer with Nokia Siemens Networks Ltd (former Siemens Communications Ltd) in Beijing, China, 2005, and as a Co-op with National Broadcasting Research Center in conjunction with Tongshi Data Communications Co., Xian, China, 2001-2005. His research interests include wireless multimedia communications, QoE-QoS issues in network economics, cyber physical system healthcare, and breast cancer imaging. He serves as an Associate Editor of Wiley Security and Communication Networks Journal (SCIE-Index), the Guest Editor of several journals and special issues such as IEEE ACCESS, Springer MONET, IJDSN, etc, the web cochair of IEEE INFOCOM 2016-17, the program chair of ACM RACS 2014-16, the track-chair of ACM SAC 2014-16, workshop co-chair of ICST BodyNets 2013, the chair of IEEE CIT-MMC track 2012, the vicechair of IEEE ICCT-NGN track 2011, and the program chair of the ICST IWMMN 2010, a Technical Program Committee (TPC) member for many international conferences such as IEEE INFOCOM, GLOBECOM and ICC. He also serves as the Co-Director of IEEE MMTC Review Board 2017-2018, and the Co-Director of the IEEE MMTC Publicity Board 2014-2016. He is a Senior Member of IEEE.

Towards Efficient Resource Allocation in Media Cloud for Mobile Social Users

A short review for "Game Theoretic Resource Allocation in Media Cloud with Mobile Social Users"

Edited by Qing Yang

Due to the increasing amount of mobile traffic and the population of mobile users in recent years, it becomes more important to provide mobile social users with efficient multimedia services than before [1]. Media cloud has been advocated to provide resources such as capacity, bandwidth, buffer, etc. to improve the quality of experience (QoE) [2] for mobile social users, instead of only depending on the local mobile devices. In the media cloud, the broker can act as a proxy to be placed close to mobile social users. With the broker, media cloud can help users to save their resource with the result that users can obtain multimedia services faster than contacting the remote multimedia content servers.

However, due to the limited resource to allocate among media cloud, brokers and mobile social users, there is a challenge efficiently allocate resource among these three parties. Conventional resource allocation schemes cannot be directly used because social features in media cloud with mobile social users, the affection and competition among different parties have not been considered enough. Some related works studied the resource allocation for cloud computing and mobile networks [3] [4]. But, few of them focus on the resource allocation problem by considering social features, while most of them mainly discuss the behaviors of servers rather than media cloud. Besides, the efficient employment of cloud brokers to allocate resource is not mentioned either. Thus, the design of resource allocation scheme is still an open issue.

In this paper, considering the competitions among three parties for cloud resource and the social features of mobile social users, the authors present a novel resource allocation scheme to maximize the utilities of three parties. The media cloud can determine the price of its cloud resource to obtain revenue. The brokers then buy resource to satisfy the demand of social users. Based on the information of others' strategies, the mobile social user can select the optimal broker and send the media tasks to the broker. According to the interactions among three parties on cloud resource, the problem of resource allocation is formulated as a four-stage Stackelberg game. To implement the proposal, an iteration algorithm is proposed to obtain the Stackelberg equilibrium.

Thus, the authors' primary contribution is to propose a framework of resource allocation among media cloud, brokers, and mobile social users. With brokers as proxies, the media cloud can allocate virtual resource to mobile social users through brokers. And mobile social users in the community with the similar interest can optimally select a broker to maximize their revenue based on the social features and others' selections on brokers.

To capture the interactions among media cloud, brokers, and mobile social users, the problem of resource allocation is formulated as a four-stage Stackelberg game. Firstly, the media cloud determines the price of its cloud resource to obtain the maximum profit. Secondly, based on the price announced by the media cloud, each broker can buy a certain size of cloud resource to process the media tasks from mobile social users. Third, each broker determines the price to optimize its revenue. A non-cooperative game is employed to study brokers' interaction. Finally, each mobile social user selects an optimal broker to obtain the satisfied QoE and sends media tasks to the selected broker. Mobile social users in the same community can obtain the information of others, where an evolution game is used to study the selections of mobile social users.

Furthermore, based on the analysis of four-stage Stackelberg game, an iteration algorithm is presented to obtain the Stackelberg equilibrium through the backward induction method to implement the proposed scheme. The media cloud announces the price to the brokers. The media cloud can set the time period to update

both the amount and the price of cloud resource to obtain the maximum utility.

Extensive simulation experiments are carried out to demonstrate the performance of the proposal. Simulation results show mobile social users can obtain the required resource according to their demands with the proposed game-based cloud resource allocation. The price can be determined reasonably and all parties can obtain maximum utilities. Compared with other schemes, the proposal can achieve the highest Media Response Ratio (MRR). It can be known that all mobile social users can select the best strategies to obtain the optimal utility. Each broker can decide the optimal price and size of cloud resource to maximize its revenue. To obtain the maximum profit, the media cloud can determine the optimal price of cloud resource.

Media cloud is a promising solution to provide mobile social users with satisfied multimedia services, facing with the explosive growth in both the volume of multimedia and the demand of QoE. Media cloud can allocate cloud resource through brokers to reduce the required resource for mobile social users and provide users with better services than the remote multimedia content servers. In summary, the proposed game-based cloud resource allocation scheme is demonstrated to well allocate the resource and maximize all parties' utilities.

Acknowledgement:

The editor thanks Dr. Zhou Su for providing critical information about the reviewed paper.

References:

[1] "Cisco Syst., Inc., San Jose, CA, USA, "Cisco Visual Networking Index: Global mobiledata traffic forecast update, 2014–2019."

[2] Y. Wu, C. Wu, B. Li, and L. Zhang, "Scaling social media applications into geo-distributed clouds," *IEEE/ACM Trans. Netw*., vol. 23, no. 3, pp. 689–702, Jun. 2015.

[3] S. Zhan and S. Chang, "Design of truthful double auction for dynamic spectrum sharing," in *Proc. IEEE Int. Symp. Dyn. Spectrum Access Netw*., Apr. 2014, pp. 439–448.

[4] I. Stanojev and A. Yener, "Relay selection for flexible multihop communication via competitive spectrum leasing," in *Proc. IEEE Int. Conf. Commun*., Jun. 2013, pp. 5495–5499.
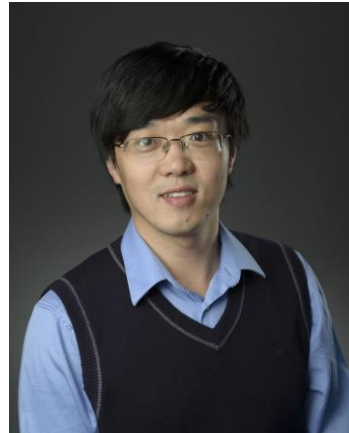
Qing Yang, Ph.D, is a RightNow Technologies Assistant Professor in the Gianforte School of Computing at Montana State University. He received B.S. and M.S. degrees in Computer Science from Nankai University and Harbin Institute of Technology, China, in 2003 and 2005, respectively. He received his Ph.D degree in Computer Science from Auburn University in 2011. His research interests include Internet of Things, online social network, trust management, and vehicular networks. He has published papers in prestigious journals such as IEEE Transactions on Dependable and Secure Computing, IEEE Network Magazine and IEEE Communications Magazine, in prestigious conferences such as IEEE INFOCOM, IEEE CNS and IEEE SECON.

A New Framework for Image Dataset Construction with Web Images

A short review for "Exploiting Web Images for Dataset Construction: A Domain Robust Approach"
Edited by Dongdong Zhang

Labelled image datasets have played a critical role in high-level image understanding. In the early years, manual annotation was the most important way to construct image datasets. (e.g., STL-10 [1], CIFAR-10 [2], PASCAL VOC 2007 [3], ImageNet [4] and Caltech-101 [5]). However, the process of manual labelling is both time-consuming and labor intensive. With the development of the Internet, methods of exploiting web images for automatic image dataset construction have recently become a hot topic in the field of multimedia processing.

Schroff et al. [6] adopted text information to rank images retrieved from a web search and used these top-ranked images to learn visual models to re-rank images once again. Li et al. [7] leveraged the first few images returned from an image search engine to train the image classifier, which uses incremental learning to refine its model. With the increase in the number of positive images accepted by the classifier, the trained classifier will reach a robust level for this query. Hua et al. [8] proposed to use clustering based method to filter "group" noisy images and propagation based method to filter individual noisy images.

The advantage of these methods [6]-[8] is that the need for manual intervention is eliminated. However, the domain adaptation ability is limited by the initial candidate images and the iterative mechanism in the process of image selection. In order to obtain a variety of candidate images, Yao et al. [9] proposed the use of multiple query expansions instead of a single query in the process of initial candidate images collection, then using an iterative mechanism to filter noisy images. The automatic works discussed here mainly focus on accuracy and scale in the process of image dataset construction, which often results in a poor performance on domain adaptation. To address this issue, this article proposed a novel image dataset construction framework that can be generalized well to unseen target domains.

The main contributions of this article include:

1. Based on multiple query expansions and multi-instance learning, the authors proposed the first method for automatic domain-robust image dataset construction, which considers the source of candidate images and retains images from different distributions. So the dataset constructed by their approach efficiently alleviates the dataset bias problem.

2. To suppress the search error and noisy query expansions induced noisy images, the authors formulate image selection as a multi-instance learning problem and propose to solve the associated optimization problems by the cutting-plane and concave-convex procedure (CCCP) algorithm, respectively.

3. The authors have released their image dataset DRID-20 on Google Drive. The diversity of DRID-20 will offer unparalleled opportunities to researchers in the multi-instance learning, transfer learning, image dataset construction and other related fields.

To demonstrate the effectiveness of the proposed approach, the authors choose 20 categories in PASCAL VOC 2007 as the target categories for the construction of DRID-20. They compare the image classification ability, cross-dataset generalization ability and diversity of their dataset with three manually labelled and three automated datasets.

For image classification ability, DRID-20 outperforms the automated datasets in terms of average accuracy in 20 categories. For cross-dataset generalization, DRID-20 outperforms CIFAR-10, STL-10, ImageNet, Optimol, Harvesting and AutoSet in terms of average cross-dataset performance. That is because DRID-20 constructed by multiple query expansions and MIL selection mechanisms has much more effective visual patterns than other datasets given the same number of training samples.

For dataset diversity, the lossless JPG file size of the average image for each category reflects the amount of information in an image and a diverse image dataset will result in a blurrier average image, the extreme being a gray image. Meanwhile, an image dataset with limited diversity will result in a more structured, sharper average image. DRID-20 has a slightly smaller JPG file size than ImageNet and STL-10 which indicates the diversity of the dataset. It can be seen that the average image of DRID-20 is blurred and it is difficult to recognize the object, while the average image of ImageNet and STL-10 is relatively more structured and sharper.

The authors also compare object detection ability of the collected data with other baseline methods. They selected PASCAL VOC 2007 as the test data since it is recent state-of-the-art weakly supervised and web-supervised methods have been evaluated on this dataset. In most cases, the proposed method surpasses the results obtained from WSVCL, IDC-MTM, which also uses web supervision and multiple query expansions for candidate images collection. The explanation for this is that the authors use different mechanisms for the removal of noisy images. Compared to WSVCL, IDCMTM which uses iterative mechanisms in the process of noisy images filtering, their approach applies an MIL method for removing noisy images. This maximizes the ability to retain images from different data distributions while filtering out the noisy images.

References

[1] A. Coates, A. Ng, H. Lee, "An analysis of single-layer networks in unsupervised feature learning," International Conference on Artificial Intelligence and Statistics, 215–223, 2011.

[2] A. Krizhevsky and G. Hinton, "Learning multiple layers of features from tiny images," Citeseer, 2009.

[3] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," International Journal of Computer Vision, 88(2): 303–338, 2010.

[4] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," IEEE International Conference on Computer Vision and Pattern Recognition, 248–255, 2009.

[5] G. Griffin, A. Holub, P. Perona, "Caltech-256 object category dataset."

[6] F. Schroff, A. Criminisi, and A. Zisserman. "Harvesting image databases from the web," IEEE Transactions On Pattern Analysis and Machine Intelligence, 33(4): 754–766, 2011.

[7] L. Li and L. Fei-Fei, "Optimol: automatic online picture collection via incremental model learning," International Journal of Computer Vision, 88(2): 147–168, 2010.

[8] X. Hua and J. Li. "Prajna: Towards recognizing whatever you want from images without image labeling," AAAI International Conference on Artificial Intelligence, 137–144, 2015.

[9] F. Shen, X. Zhou, Y. Yang, J. Song, H. Shen, and D. Tao, "A fast optimization method for general binary code learning," IEEE Transactions on Image Processing, 25(12): 5610–5621, 2016.

Dongdong Zhang is an associate professor of Computer Science department at Tongji University, P. R. China. She received her Ph.D. in Electronic Engineering from Shanghai Jiao Tong University in 2007. She worked as a postdoc research fellow with Chinese University of Hong Kong, 2007-2008. Her research interest includes image communication and processing, perceptual video coding, 3D video quality evaluation, 3D video coding and rendering, HEVC, scalable video coding, video processing and analysis etc. She is a recipient of Shanghai Outstanding Doctoral Dissertation and Excellent young talents of Tongji University.

Approaching New Limits of Synchrony with Multi-sensorial Media

A short review for "Digital Lollipop: Studying Electrical Stimulation on the Human Tongue to Simulate Taste Sensation"

Edited by Carsten Griwodz

The field of multimedia research is carrying with it a name that comes with an implicit claim that our research is concerned with several forms of media. Ordinarily, we understand here only auditory and visual media, and claim that text, images, videos, songs and speech are so severely different types of content that we should consider them different media and consequently, combination of these multimedia. However, these content types are all perceived with only our eyes and ears (with the addition of fingers in case of the few people who are capable of reading Braille). The human senses extend beyond this: we know about vision, hearing, touch, smell and taste.

The authors of "Digital Lollipop: Studying Electrical Stimulation on the Human Tongue to Simulate Taste Sensations" are continuing their research into the human sense that has attracted the least of all senses. Whereas vision and hearing are stimulated by multimedia devices in everyday use, haptics research that focuses on touch has produced results ranging from force-feedback devices [1] to vibrotactile vests [2], and smell has received a moderate amount of attention [3][4], the generation of stimuli that allow us to emulate taste has been absent, apart from the authors' earlier work [5][6][7]. In the current paper, the authors present the "Digital Lollipop", which is a device meant to explore the human ability to recognize direct electrical stimuli on the tongue as taste. The name of the device is meant to imply that it should eventually resemble a classical sweet lollipop, although it's current instance is still rather far from this vision. The contraption consists of two differently-sized electrodes that can stimulate the top and bottom of the tongue.

The authors have made extensive literature study on the existing work on electrical stimulation of the human tongue, and designed the device appropriately. They are conducted a set of user studies to determine how humans experience tiny electrical stimuli to the tongue. These user studies are prime examples of well-designed user studies and careful interpretation of the results, where the authors make all efforts to remove a bias from their interpretation of their observations. Even if a reader might find the results of the authors' experiments not terribly promising for an all-electronic solution to the taste problem, it is this assessment procedure that makes the paper highly worth reading for new students who are faced with the task to design a user study experiment outside the ordinary tracks of our field. Ranasinghe and Do find at the end of their first user study, that only a small range of electrical stimuli is recognized as valid taste signals. A majority of the testers (90%) recognized the taste as "sour", but 70% tasted saltiness, 50% bitterness and 5% sweetness. Consequently, it is hard to claim that the electrical signals are interpreted in a consistent manner. Continuing with the majority group only, the authors designed and conducted a test that would estimate the intensity of the sourness by comparing it with the taste of lime juice at a measurable pH value. While the result of dual stimuli (the most robust of all subjective tests) was very stable, indicating that the relative intensity of the taste experience could be confirmed, they observed also that the absolute current required for a comparably sour taste was individual. Furthermore, some of their testers reported a numbness of their tongue after the test, and it was reported by several that the lime juice tasted more natural; although all efforts had been made to isolate the effect of smell from the comparison.

We can probably conclude that the direct stimulation of the tongue through electrodes is not the final word for the delivery of the new multimedia experience of taste to end users. Still, the authors have demonstrated that not attempts of recreating a taste experience is not hopeless even when we attempt to deliver it without a recreation of taste by delivering a chemical

substance. However, for future experiments, the authors promote a new approach that relies on a less dogmatic approach free of chemicals, and propose a future experimental setup including a chemical delivery mechanism for additional smell.

The authors may be going too far in their desire to create a purely electrical interface for delivering taste. The inconsistency of results, the strongly personalized interpretation of the signals, and the limitation to the sour taste, all seem to indicate that this particular approach to conveying taste over temporal and spatial distances may not be dealt with by purely electronic means. In designing future experiments, the authors are therefore also proposing a system that mixes chemical, smell-inducing components with the electrical stimulation devices. But should we really expect that the goal of taste transfer can be solved by direct stimulation of the tongue? Loud-speakers are transforming electronically communicated data into sound waves, video are transformed into colored lights, models may be printed by 3D printers, text and images may be printed out. Perhaps a transfer into another medium, even if it requires containers of consumable resources like the ink of a printer, will turn out to be the key to success also in taste transfer.

But what the authors have most certainly done is to remind us that this sense exists and that simulating it remotely should be a desirable prospect for multimedia researchers. They have demonstrated that it is possible to start small with a limited set of tools, and most of all, they demonstrate with their paper an excellent methodology for thorough test design, a thoroughly conducted investigation, and self-critical evaluation of the results. Even if a researcher is not willing to immediately adopt taste as a new research medium, it would still be most highly recommended to learn from the authors' approach to conduct and evaluate an investigation into any form of new media presentation and consumption in our world of multiple media.

References:

[1] L. Buoguila, M. Ishii, and M. Sato, "Multi-modal haptic device for large-scale virtual environments," in Proceedings of the eighth ACM international conference on Multimedia, 2000, pp. 277–283.

[2] A. Morrison, C. Manresa-Yee, and H. Knoche, "Vibrotactile Vest and The Humming Wall," in Proceedings of the XVI International Conference on Human Computer Interaction, 2015, pp. 1–8.

[3] N. Murray, Y. Qiao, B. Lee, A. K. Karunakar, and G.-M. Muntean, "Subjective evaluation of olfactory and visual media synchronization," in Proceedings of the 4th ACM Multimedia Systems Conference, 2013, pp. 162–171.

[4] Zhenhui Yuan, G. Ghinea, and G.-M. Muntean, "Beyond Multimedia Adaptation: Quality of Experience-Aware Multi-Sensorial Media Delivery," IEEE Trans. Multimed., vol. 17, no. 1, pp. 104–117, Jan. 2015.

[5] N. Ranasinghe, A. D. Cheok, O. N. N. Fernando, H. Nii, and P. Gopalakrishnakone, "Digital taste: Electronic stimulation of taste sensations," in Lecture Notes in Computer Science, 2011, vol. 7040 LNCS, pp. 345–349.

[6] N. Ranasinghe, R. Nakatsu, H. Nii, and P. Gopalakrishnakone, "Tongue mounted interface for digitally actuating the sense of taste," in Proceedings of International Symposium on Wearable Computers, 2012, pp. 80–87.

[7] N. Ranasinghe, A. D. Cheok, O. N. N. Fernando, H. Nii, and G. Ponnampalam, "Electronic taste stimulation," Proc. 13th Int. Conf. Ubiquitous Comput. , pp. 561–562, 2011.

Carsten Griwodz is chief research scientist in the Media Department at the Norwegian research company Simula Research Laboratory AS, Norway, and professor at the University of Oslo. His research interest is the performance of multimedia systems. He is concerned with streaming media, which includes all kinds of media that are transported over the Internet with a temporal demands, including stored and live video as well as games and immersive systems. To achieve this, he wants to advance operating system and protocol support, parallel processing and the understanding of the human experience. He was area chair and demo chair of ACM MM 2014, and general chair of ACM MMSys and NOSSDAV (2013), co-chair of ACM/IEEE NetGames (2011), NOSSDAV (2008), SPIE/ACM MMCN (2007) and SPIE MMCN (2006), TPC chair ACM MMSys (2012), and systems track chair ACM MM (2008). More information can be found at http://mpg.ndlab.net.

Paper Nomination Policy

Following the direction of MMTC, the Communications – Review platform aims at providing research exchange, which includes examining systems, applications, services and techniques where multiple media are used to deliver results. Multimedia includes, but is not restricted to, voice, video, image, music, data and executable code. The scope covers not only the underlying networking systems, but also visual, gesture, signal and other aspects of communication. Any HIGH QUALITY paper published in Communications Society journals/magazine, MMTC sponsored conferences, IEEE proceedings, or other distinguished journals/conferences within the last two years is eligible for nomination.

Nomination Procedure

Paper nominations have to be emailed to Review Board Directors: Pradeep K. Atrey (patrey@ albany.edu), Qing Yang (qing.yang@montana. edu), Wei Wang (wwang@mail.sdsu.edu), and Jun Wu (wujun@tongji.edu.cn). The nomination should include the complete reference of the paper, author information, a brief supporting statement (maximum one page) highlighting the contribution, the nominator information, and an electronic copy of the paper, when possible.

Review Process

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

Best Paper Award

Accepted papers in the Communications – Review are eligible for the Best Paper Award competition if they meet the election criteria (set by the MMTC Award Board). For more details, please refer to http://mmc.committees.comsoc. org/.

MMTC Communications – Review Editorial Board

Multimedia Communications Technical Committee Officers

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