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

MMTC Communications – Review

Vol. 9, No. 6, December 2018

IEEE COMMUNICATIONS SOCIETY

TABLE OF CONTENTS

Message from the Review Board Directors

Welcome to the December 2018 issue of the IEEE ComSoc MMTC Communications – Review.

This issue comprises three reviews that cover multiple facets of multimedia communication research including indoor location determination via deep learning, incentive-based D2D video distribution, and 2D to 3D video conversion. These reviews are briefly introduced below.

The first paper, published in IEEE Transactions on Vehicular Technology and edited by Dr. Xiaohu Ge, designed an algorithm to leverage channel state information and deep learning for indoor localization.

The second paper is published in IEEE International Conference on Image Processing and edited by Dr. Gwendal Simon. It proposes a novel method for Plenoptic point cloud compression, leveraging different colors depending on the position of the observer.

The third paper, published in IEEE Transactions on Multimedia and edited by Dr. Carsten Griwodz, investigates the use of rendered 3D scenes to emulate depth on a large scale and use this knowledge for estimating depth within the specific context of the input video.

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

IEEE ComSoc MMTC Communications – Review Directors

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

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

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

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

A Deep Learning based Approach for Indoor Localization

A short review for "CSI-based fingerprinting for indoor localization: A deep learning approach"

Edited by Dr. Xiaohu Ge

X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," IEEE Transactions on Vehicular Technology, vol.66, no.1, pp.763—776, Jan. 2017. DOI: 10.1109/TVT.2016.2545523.

Location is important information for many mobile applications, for example, navigation and tracking. Since most mobile users are indoors, indoor localization is of great interest. Accurate indoor localization can enable traditional and new applications such as navigation in a stadium or exhibition hall, location-based advertisement, access control to wireless networks or information based on location, and even faster beamforming/beam tracking in 5G mmWave networks. Although having been studied for decades, there is still a great need for accurate and robust solutions for complex indoor environments.

Among various indoor localization techniques, WiFi based fingerprinting has many advantages. For example, it does not require access to GPS or cellular base stations, and WiFi access is ubiquitous in many indoor environments. There is also no need for accurate propagation models as in ranging based schemes. In such schemes, survey data is collected for chosen positions in the off-line phase. During the on-line phase, a mobile device records its realtime WiFi signal and compares it with stored survey data, to find the best match and determine its location [1]. Many existing schemes are based on received signal strength (RSS), which only represent coarse channel information [2]. The new trend is to move from RSS to channel state information (CSI), which represents fine-grained channel information and is now available for several WiFi cards [3]. The challenge is how to effectively process the much larger CSI data for accurate indoor localization in realtime.

In this paper, the authors propose a novel deep-learning-based fingerprinting scheme, termed DeepFi, to address the challenge [4]. The deep-learning-based scheme can fully exploit the rich features of WiFi CSI data and obtain the optimal weights as fingerprints [5]. The authors also incorporate a greedy learning algorithm to reduce computational complexity. In particular, the authors first present three hypotheses on CSI, which justify the feasibility of using CSI for more accurate indoor location. The authors then present the DeepFi system design, which includes an offline training phase and an online localization phase. In the training phase, CSI information for all the subcarriers from the three antennas if the WiFi card is collected from accessing the device driver and is analyzed with an autoencoder with four hidden layers. The proposed greedy learning algorithm uses a stack of restricted Bolzmann machines (RBMs) to train the deep network in a layer-by-layer manner to reduce complexity. Moreover, for each layer of the RBM model, the authors adopt the contrastive divergence with one-step iteration (CD-1) method to update weights, which has lower time complexity than other schemes, such as Markov chain Monte Carlo. In the online phase, a probabilistic fusion method based on radial basis function (RBF) is developed for location estimation. To reduce the computational complexity, packets are divided into several batches of equal size. Because packets are processed in parallel in batches, the processing time can be significantly shortened when dealing with a large amount of packets.

The proposed DeepFi scheme is implemented with a laptop and wireless router with low-cost Intel 5300 WiFi cards. The authors conducted extensive experiments in two representative indoor environments, i.e., a living room and a computer laboratory to validate its performance. DeepFi is shown to outperform several existing RSS and CSI-based schemes in both experiments. The effects of different DeepFi system parameters and different propagation environments are also evaluated in the experiments. The experimental results clearly confirm that DeepFi can perform well in these scenarios.

The major contribution of this paper is to propose the first deep learning based design for indoor WiFi fingerprinting. It is worth noting that this work was conducted before the release of AlphaGo in Oct. 2015 and the release of TensorFlow in Nov. 2015, which triggered the great interest on applying deep learning/machine learning to networking problems. This work is the first to introduce deep learning to solving indoor localization problems and clearly demonstrates the feasibility and the high potential of the deep learning based approach.

Following this work, the authors have published a body of work on applying various deep learning algorithms to indoor fingerprinting [6,7,8], which also led to several US provisional patents. This work also triggered considerable interest in the community. Since its publication in Jan. 2017, this paper has been the Top 1 most downloaded in most of the months (except for two months, when it was the Top 2 most downloaded) among all papers published in *IEEE Transactions on Vehicular Technology*. In a short period of less than two years, this paper has received 149 citations, while its conference preliminary version received 75 citations, according to Google Scholar (as of Dec. 1, 2018).

In summary, this paper made a great contribution in presenting the first deep learning based WiFi indoor fingerprinting solution. The proposed approach is implemented with commodity WiFi and demonstrated to achieve an accurate and robust localization performance. It is also well received in the community, as indicated by the high download and citation numbers in less than two years.

References:

[1] A. Khalajmehrabadi, N, Gatsis, and D. Akopian, "Modern WLAN fingerprinting indoor positioning methods and deployment challenges," *IEEE Communications Surveys & Tutorials*, vol.19, no.3, pp.1974-2002, Third Quarter 2017.

[2] X. Ge, B. Yang, J. Ye, G. Mao, C.-X. Wang and T. Han, "Spatial spectrum and energy efficiency of random cellular networks," *IEEE Transactions on Communications*, vol. 63, no. 3, pp. 1019 - 1030, March 2015.

[3] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Computing Surveys*, vol.46, no.2, pp.25:1–25:32, Nov. 2013.

[4] X. Wang, L. Gao, S. Mao, and S. Pandey, "CSI-based fingerprinting for indoor localization: A deep learning approach," *IEEE Transactions on Vehicular Technology*, vol.66, no.1, pp.763-776, Jan. 2017.

[5] X. Wang, X. Wang, and S. Mao, "RF sensing for Internet of Things: A general deep learning framework," *IEEE Communications*, vol.56, no.9, pp.62-67, Sept. 2018.

[6] X. Wang, L. Gao, and S. Mao, "CSI phase fingerprinting for indoor localization with a deep learning approach," *IEEE Internet of Things Journal*, vol.3, no.6, pp.1113-1123, Dec. 2016.

[7] X. Wang, L. Gao, and S. Mao, "BiLoc: Bi-modality deep learning for indoor localization with 5GHz commodity Wi-Fi," *IEEE Access Journal*, vol.5, no.1, pp.4209-4220, Mar. 2017.

[8] X. Wang, X. Wang, and S. Mao, "Deep convolutional neural networks for indoor localization with CSI images," *IEEE Transactions on Network Science and Engineering*, in press. DOI: 10.1109/TNSE.2018.2871165.

Xiaohu Ge (M'09–SM'11) received the Ph.D. degree in communication and information engineering from the Huazhong University of Science and Technology (HUST), China, in 2003. He was a Researcher with Ajou University, South Korea, and Politecnico di Torino, Italy, from 2004 to 2005. He has been with HUST since 2005. He is currently a Full Professor with the School of Electronic Information and Communications, HUST. He is also an Adjunct Professor with the Faculty of Engineering and Information Technology, University of Technology Sydney, Australia. He has published over 200 papers in refereed journals and conference proceedings, and has been granted about 15 patents in China. His research interests include mobile communications, traffic modeling in wireless networks, green communications, and interference modeling in wireless communications. He received the Best Paper Awards from the IEEE GLOBECOM 2010. He served as a General Chair for the 2015 IEEE International Conference on Green Computing and Communications in 2015. He serves as an Associate Editor for the IEEE WIRELESS COMMUNICATIONS, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY.

Plenoptic Point Cloud Compression

A short review for "Compression of Plenoptic Point Cloud Using the Region-Adaptive Hierarchical Transform"

Edited by Gwendal Simon

G. Sandri, R. de Queiroz, and P. A. Chou, "Compression of Plenoptic Point Cloud Using the Region-Adaptive Hierarchical Transform," in IEEE International Conference on Image Processing, Athens, Greece, 2018.

The world of motion picture is at the dawn of a major disruption. The set of technologies that humans have developed for capturing a sequence of images from real scenes, compressing this flow, delivering it, and displaying it is culminating today when a standard phone allows anybody to generate a video that can be watched by potentially millions of people all over the world. This amazing set of technologies is however showing its limits when one would like to also get a sense of the volume in the captured scene. The fundamentals of volumetric videos are essentially different from what has been studied and implemented in the past half-century. Researchers have to reinvent the full chain of acquisition, delivering, and rendering, with regards to the lessons learnt from the remarkable success of video delivery, but without fear of challenging some of the cornerstones that enabled our current multimedia era. This huge challenge is what makes the community of multimedia systems pretty excited right now.

This paper, which was presented at this year's edition of the IEEE ICIP conference, is about the compression of point cloud. The research community has been active in the past five years in the field of compressing the set of points that can be captured by the cameras that feature depth acquisition. A large part of the excitement comes from the availability of technologies that enable relatively cheap hardware to acquire in real time any surrounding (real world scene around) at centimeter and even millimeter scales. The huge set of captured "points" (often referred to as *voxels* for volume pixels) is expected to become the input data, which will then be manipulated by some software to offer volumetric experience, regardless of the actual applications that are here envisioned. From this perspective, the compression of the said "point cloud" is a key element of the future chain of volumetric multimedia entertainment.

As far as I can observe it, the vast majority of the research in the area of point cloud compression has focused on reducing the voxel structure representation load, while enabling various features such that offering levels of details locally. In short, most of the papers I read state that the information that is actually stored at the level of a voxel is a vague color and material data. This paper is one of the first work I have paid attention to that deals with the compression of this peculiar color information. Although I have only a loose familiarity with point cloud compression and light field area, I was able to appreciate that this compression problem is not trivial. Yet another exciting problem in the unpaved road toward volumetric media entertainment.

The main idea behind plenoptic information is that one given point in the space can have a different color depending on the position of the observer. This color depends on many parameters, including light and materials. The fact is that, if one would like to generate a new view from anywhere in the space, the more we know about the voxel color, the better. This paper thus aims at compressing the so-called plenoptic information of every voxel. The contribution complements the other studies on geometrical point cloud compression and presents a high interest for the 3D renderer in charge of ingesting the color data to typically reconstruct the scene or render it from new viewpoint.

The contribution in itself is rather simple: three compression methods have been designed by the authors. The first one projects the captured colors in a cylinder around the voxel, the two others leverage the intersection of rays and the faces around the voxel center (cube or sphere). The proposals are evaluated across a set of representative datasets and the two latter appear to perform better. The authors also show that the compression is much higher than just storing the color from all cameras without exploiting the underlying geometry.

The paper reads well, and the fundamental idea behind plenoptic compression is the kind of fascinating problem that is thought-provoking. How to exploit the wide set of information that one may acquire from a large set of depth-enabling cameras is just the beginning of the story that should eventually lead the humans to the implementation of a totally revamped multimedia entertainment chain. This paper is a contribution to this field, with regards to multi-view object acquisition.

Gwendal Simon is a Full Professor at IMT Atlantique, an elite technological university in France. His research interests include multimedia delivery systems (video and gaming) and network management. He has written more than 80 scientific papers, 4 of them having been awarded as best papers in prestigious conferences. His impactful research has also resulted in six patents, several software and datasets (including the awarded Solipsis p2p virtual world), and contributions to multiple innovative collaborative projects (including two awarded projects). He has advised 13 PhD students and he has directed a research lab (including 6 post-docs and research engineers). He graduated from University Rennes 1 (France). He obtained a PhD in Computer Science in 2004 and a Habilitation in 2015. From 2001 to 2006 he was a researcher at the research center of Orange (then France Telecom). Since 2006, he has been Associate Professor, and then a Full Professor at IMT Atlantique. He was a visiting researcher at University of Waterloo in 2011/2012 and he held a position of senior scientist at Adobe in 2018/2019.

Application-specific 2D-to-3D Video Conversion

A short review for "Data Driven 2-D-to-3-D Video Conversion for Soccer"

Edited by Dr. Carsten Griwodz

K. Calagari, M. Elgharib, P. Didyk, A. Kaspar, W. Matusik and M. Hefeeda, "Data Driven 2-D-to-3-D Video Conversion for Soccer," IEEE Transactions on Multimedia, vol.20, no.3, pp.605—619, Mar. 2018. DOI: 10.1109/TMM.2017.2748458.

The transport and presentation of stereoscopic content is today well-standardized through the MPEG family of codecs, and a large share of recent television hardware can display this kind of content, either through auto-stereoscopic displays or by means of polarized glasses. Similarly, this kind of content is available to computer users, who may also use headsets of various cost and complexity. The adoption of stereoscopic viewing habits by end-users, however, depends on the availability of compelling content. Since the generation of this kind of content requires planning and the appropriate hardware during the recording of live action film, and depth information is hardly available for content pre-dating the technology, there have been many attempts to understand depth in 2D content and use this understanding to generated depth information. This kind of 3D content generation has been raised to popularity especially with the proliferation of devices.

Many of these approaches rely on motion in the recording itself without [1-7] or with user input [8-9], while the most aggressive ones attempt generic object understanding and add depth information directly to frames [10-16]. The method proposed in this journal paper is restricted to a single context, namely soccer, and takes a unique approach for the estimation of depth. The exciting new contribution of "*Data Driven 2-D-to-3-D Video Conversion for Soccer*" is the use of rendered 3D scenes to emulate depth on a large scale and use this knowledge for estimating depth within the specific context of the input video.

The authors give a good insight into existing work, but present a very unexpected and innovative approach to solving the creation of high-quality 3D information: they use one of today's excellent football games to create a large database of realistic player poses. Obviously, the 3D information from arbitrary camera angles is known in this database, and where researchers in other works struggled to build databases from real-world ground truth, the authors generate a database of arbitrary size. They study the appropriate database size and find an applicable limit. This approach by itself is already a great contribution because it exploits the opportunities of the football scenario and is highly innovative.

An essential decision for the accuracy of the depth map creation is the understanding that the scene shown in every single frame can be reconstructed from the database. The method starts with the selection of a set of similar frames that are selected using classical similarity for structure (GIST) and color. In principle, the football scenario would allow for a transformation of the selected frame into the exact spatial football field position of the query image, but without a pre-existing 3D structure, it would break the coherence of the most salient elements, the players, due to parallax effects. The authors choose to forego this older idea, and work with the assumption that the pictures are locally consistent. Instead, they cut the images in tiny blocks (9x9 in the paper), and match these blocks using SIFT features in the block and a surrounding region. A frame that is reconstructed from these matched blocks does already provide an image that is highly similar to the input image, but carries depth information from the database. The patchwork of blocks, however, does not satisfy the quality of depth estimation yet, and more importantly, cannot guarantee coherence across frames.

So, the authors add a step that remedies the temporal coherence. Transferring the depth map from the database to the query image, they use Poisson reconstruction to smooth and adapt locally the depth of objects in non-flat regions. This reconstruction process is extended to the time domain to ensure consistent depth maps across several frames. Within each frame, object masks are used to enforce obviously correct discontinuities, in particular at the boundaries separating players from the field. They are then dealing separately with those regions where domain-specific knowledge reveals that they are flat, and chances for mismatches in SIFT feature matching are high.

The authors do not only demonstrate that this approach provides detailed depth maps for players on the football field, they have also designed the process to run efficiently in parallel. All relevant features for the database images are precomputed, while the Poisson reconstruction process is split into temporal windows for a higher degree of parallelism, with a final averaging steps between the windows' depths.

Very interesting aspects of the approach taken by the authors are in the place where attractive domain knowledge is actually ignored. Looking at a frame in its entirety, it would not be trivial to estimate individual players' depth in isolation from other, because many game situations such as tackles, corners and penalties lead to a real-world situation where players are physically connected. Instead, the authors look at the in terms of small, square pixel blocks that loose semantic meaning for a human observer, and transfer depth information mechanically from similar blocks in other images. Depth discontinuities may then occur where players occlude each other, but because of the authors' enhancement of the Poisson reconstruction to the time dimension, correct discontinuities at the occlusion line between players will fade unless they can be separated by object masks.

This explains why the authors went to considerable lengths to create good object masks. Firstly, they used domain knowledge to distinguish between close-ups of players and long shots. While the amount of grass in an image is an obvious give-away, close-ups don't necessarily have a lot of it, and player size in combination with audiences in the background are used to detect them. Shots of one type or another were separated using image similarity, a straightforward method that works in the domain because shots don't transition between totals and close-ups in football, and lighting doesn't change rapidly. After this coarse classification of shots, the two types are treated separately, and the distinct methods rely on considerable domain knowledge. In the long shots, field detection and warping is used to compensate for camera movement, and optical flow is combined with classical background detection to assign pixels as belonging to players or background. The authors demonstrate that both contributions are needed for the successful masking of all players. The close-up shots are treated in a very different manner, because advertising and audiences prevent any color-based method. Instead, and color-based optical flow is used to track motion in combination with feature-based motion. Combined with camera motion trajectories, this provides dense pixel trajectories, which allow the creation of quite accuracy player masks.

The paper rounds up with a very detailed study into the quality-of-experience that is provided by the proposed system. The competition in this case is not any of the existing generic methods for deriving depth information from 2D video. Instead, existing stereoscopic ground truth information is used, either from stereoscopic recordings or from virtual sequences, and compared with the depth recovery approach proposed by the authors. The results show that the method provides quite similar visual quality to original sequences, and improves over simplified methods.

The paper impresses with the number of steps that are combined to make maximum use of the domain knowledge, as well as the breadth of methods that have been put to use. While it is uncertain how general the proposed method for adding depth information to a football scenario can be applied in other contexts, it does demonstrate in an excellent way how important domain knowledge can be to derive depth information from 2D video. The highly innovative idea of extracting information from entirely virtual content, taken from computer games, adds a new dimension to the domain-specific content generation. It will be highly interesting to see whether a similarly inspired approach can be repeated in future research.

References:

[1] S. Patil and P. Charles, "Review on 2D-to-3D Image and Video Conversion Methods," *2015 International Conference on Computing Communication Control and Automation*, Pune, 2015, pp. 728-732.

[2] E. Rotem, K. Wolowelsky, D. Pelz, "Automatic video to stereoscopic video conversion", Proc.

SPIE 5664, Stereoscopic Displays and Virtual Reality Systems XII, 2005
[3] C. Varekamp, B. Barenbrug, 'Improved depth propagation for 2D to 3D video conversion using key-frames', IET Conference on Visual Media Production (CVMP), p. 29-29, 2007
[4] S. Knorr and T. Sikora, "An Image-Based Rendering (IBR) Approach for Realistic Stereo View Synthesis of TV Broadcast Based on Structure from Motion," *2007 IEEE International Conference on Image Processing*, San Antonio, TX, 2007, pp. VI - 572-VI - 575.
[5] K. Moustakas, D. Tzovaras and M. G. Strintzis, "Stereoscopic video generation based on efficient layered structure and motion estimation from a monoscopic image sequence," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 8, pp. 1065-1073, Aug. 2005.
[6] Y. Yan, F. Xu, Q. Dai and X. Liu, "A novel method for automatic 2D-to-3D video conversion," *2010 3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video*, Tampere, 2010, pp. 1-4.
[7] J. L. Herrera, J. Konrad, C. R. del-Bianco and N. García, "Learning-based depth estimation from 2D images using GIST and saliency," *2015 IEEE International Conference on Image Processing (ICIP)*, Quebec City, QC, 2015, pp. 4753-4757.
[8] M. Guttmann, L. Wolf and D. Cohen-Or, "Semi-automatic stereo extraction from video footage," *2009 IEEE 12th International Conference on Computer Vision*, Kyoto, 2009, pp. 136-142.
[9] Xi Yan, You Yang, Guihua Er and Qionghai Dai, "Depth map generation for 2D-to-3D conversion by limited user inputs and depth propagation," *2011 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, Antalya, 2011, pp. 1-4.
[10] J. L. Herrera, C. R. del-Blanco and N. García, "Automatic Depth Extraction from 2D Images Using a Cluster-Based Learning Framework," in *IEEE Transactions on Image Processing*, vol. 27, no. 7, pp. 3288-3299, July 2018.
[11] T. Leimkühler, P. Kellnhofer, T. Ritschel, K. Myszkowski and H. Seidel, "Perceptual Real-Time 2D-to-3D Conversion Using Cue Fusion," in *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 6, pp. 2037-2050, 1 June 2018.
[12] K. Karsch, C. Liu and S. B. Kang, "Depth Transfer: Depth Extraction from Video Using Non-Parametric Sampling," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 11, pp. 2144-2158, 1 Nov. 2014.
[14] J. Konrad, M. Wang and P. Ishwar, "2D-to-3D image conversion by learning depth from examples," *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, Providence, RI, 2012, pp. 16-22.
[15] Y. Kuznietsov, J. Stückler and B. Leibe, "Semi-Supervised Deep Learning for Monocular Depth Map Prediction," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017, pp. 2215-2223.
[16] P. R. Kapse and S. S. Hippargi, "Automatic 2D-to-3D image conversion using KNN search," *2016 IEEE International Conference on Advances in Electronics, Communication and Computer Technology (ICAECCT)*, Pune, 2016, pp. 210-215.

Carsten Griwodz is professor at the University of Oslo and adjunct research scientist at the Norwegian research company Simula Research Laboratory AS, Norway.
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: Qing Yang (qing.yang@unt.edu), Roger Zimmermann (rogerz@comp.nus.edu.sg), Wei Wang (wwang@mail.sdsu.edu), and Zhou Su (zhousu@ieee.org). The nomination should include the complete reference of the paper, author information, a brief supporting statement (maximum one page) highlighting the contribution, the nominator information, and an electronic copy of the paper, when possible.

Review Process

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

Best Paper Award

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

MMTC 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.