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

*MMTC Communications – Review*

Vol. 7, No. 3, June 2016

IEEE COMMUNICATIONS SOCIETY

## TABLE OF CONTENTS

# Message from the Review Board Directors

Welcome to the June 2016 issue of the IEEE ComSoc MMTC **Communications – Review**. As of this issue we will adopt this new name instead of the previously used R-Letter

This issue comprises **four reviews** in the area of *mulsemedia streaming*, *panoramic projections for 360° video*, *uncompressed video streaming*, and *image clustering*.

We hope that this issue **stimulates your research in the area of multimedia communication**.

An overview of all reviews is provided in the following:

The **first paper**, published in *IEEE Transactions on Multimedia* and *edited by Frank Hartung*, describes means for QoE-optimized streaming of multi-sensorial media addressing all human senses.

The **second paper**, published in the Proceedings of the *IEEE International Symposium on Mixed and Augmented Reality (ISMAR) 2015* and *edited by Alisa Devlic, Gwendal Simon, and Xavier Corbillon*, provides a framework to evaluate omnidirectional video coding schemes.

The **third paper** is *edited by Carl James Debono* and has been published within the *IEEE Transactions on Multimedia*. It addresses the streaming of uncompressed video over 60GHz wireless networks.

Finally, the **forth paper**, published in the Proceedings of the *International Conference on Computer* and *edited by Jun Zhou*, describes an approach how to cluster 100 million images on a single machine in less than one hour.

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

Finally, we would like to highlight upcoming conferences in 2016 which are related to MMTC:

- **IEEE ICME**, July 11-15, Seattle, USA: http://www.icme2016.org/
- **ACM Multimedia**, October 15-19, Amsterdam, The Netherlands: http://www.acmmm.org/2016/
- **IEEE GLOBECOM**, December 4-8, Washington, DC, USA: http://globecom2016.ieee-globecom.org/

**IEEE ComSoc MMTC Communications – Review**

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

**Co-Director**: Yan Zhang
Simula Research Laboratory, Norway
Email: yanzhang@simula.no

# Addressing All Senses
# QoE Optimized Multi-Sensorial Media Streaming

*A short review for "Beyond Multimedia Adaptation: Quality of Experience-Aware Multi-Sensorial Media Delivery*
*(Edited by **Frank Hartung**)*

*Zhenhui Yuan, Gheorghita Ghinea, and Gabriel-Miro Muntean, "Beyond Multimedia Adaptation: Quality of Experience-Aware Multi-Sensorial Media Delivery", IEEE Transactions on Multimedia 17(1), 2015, pp. 104-117*

For many years, the term multimedia was mainly understood as the combination of video (image sequences), and sound. Only very exotic and rare applications used other additional media, appealing to other human senses than the eyes and ears. Ideas for such systems existed, but were hardly investigated in a systematic scientific context.

Meanwhile however, supported by the development of processor capabilities and communication bit rates, but also of sensors and actuators, concepts for multi-sensorial media systems are getting into focus again. Media that come into play besides audio and video (including all their extensions like 3D, 360 degree video, VR, spatial sound) are for example haptics [1], odors or scents [2], temperature, humidity, and air motion. Among them, haptics is certainly the most popular; even more so, if vibration (as used in smartphones) is counted into that area. The term mulsemedia (multiple sensorial media) has been coined to describe such systems that address art least three human senses [3]. Possible applications are immersive entertainment, or immersive communications [4].

The authors describe a system for optimized delivery of mulsemedia. This is in the spirit of joint source-channel coding, where source rate and transmission rate (by coding) are jointly optimized to deliver the optimal end-to-end perceived quality. Here, the rate allocation between different media is added as a third dimension, or optimization axis. To achieve this, it is necessary to develop a strategy for rate allocation among mulsemedia, that means an understanding which senses should be supplied with stimuli when the rate is constrained, and not all mulsemedia can be transmitted.

To get this basic understanding of the importance of different senses, the authors have performed subjective tests regarding sensorial effects. Short sequences from commercial movies were encoded at different video qualities, and in addition augmented with mulsemedia effects, namely with olfaction, haptics (vibration), and air vent with a fan. Different scenes from the movies were augmented with different combinations of these additional effects. Test persons were then exposed to the mulsemedia-augmented scenes. Afterwards, the test persons evaluated their experience with a qualitative questionnaire. From the results, the authors draw a few qualitative conclusions:

- The majority of users (around 85%) appreciate and enjoy the use of mulsemedia effects.
- If enhanced with mulsemedia, the differences between high and average video quality are less perceived. Mulsemedia can thus "cover" a decrease in video quality. This is in fact an important new qualitative insight.
- Users prefer the use of haptics over air effects, and air effects over olfactory effects
- Synchronization between sensorial effects and multimedia content needs to be precise, especially when olfaction is included.
- Unpleasant smells such as methane and rubbish annoy the users and result in reduced user enjoyment levels.

Based on these results, the authors propose an optimized mulsemedia system, called "ADAptive MulSemedia delivery solution" (ADAMS), that adapts to varying transmission conditions. Earlier multimedia transmission system would possibly have dropped "additional" mulsemedia first, before video quality is reduced. Based on the insights from the subjective tests, ADAMS prioritizes mulsemedia higher. In other words, video quality is decreased first, then sensorial content is dropped in inverse order of user preference (i.e., olfaction, air, and haptic), before the video is eventually dropped totally and audio only is delivered.

An ADAMS server first describes to the ADAMS client which mulsemedia components are available in a presentation. This meta description is based on MPEG-7. Based on user preferences signaled back

from the client to the server, and on available bandwidth, the server selects the mulsemedia components to be transmitted for the presentation. All involved media are then transmitted as data packets (for the mulsemedia data, a special packet format based on MPEG-7 XML descriptions is used). A packet scheduler decides which packets to transmit, based on short-term network condition feedback. The priorities of the different mulsemedia, as determined in the subjective tests and also as possibly signaled by the client, are taken into consideration. The result is that video quality is sacrificed first in the presence of network bottlenecks, and mulsemedia transmission later. Synchronization of different mulsemedia is also considered by adding appropriate time delays. However, propagation times of some of the media (e.g. olfactory smells) are sometimes hard to control, and are also difficult to standardize.

The authors have tested their system with a limited number of 16 test persons. The ns2 network simulator was used to create a controlled environment where congestion occurs, resulting in the necessity of dropping packets from the mulsemedia presentations. The qualitative result was that the test persons evaluated the proposed ADAMS system better than previous multimedia delivery systems not optimized for mulsemedia, like the Quality-Oriented Adaptation Scheme (QOAS) system. While the pure audio/video quality was similar, the overall "user enjoyment experience" for the mulsemedia presentation was improved. As the authors phrase it, "ADAMS outperforms QOAS, as the percentage of users enjoying their experience in the "Agree" and "Strongly Agree" categories has increased by 10.7% and 7.5%, respectively.

Future research will have to extend the quantitative analysis and evaluation of mulsemedia systems, which provide additional media appealing to other senses than the eyes and ears. However, the current research, as described in the paper, is a promising early step into the direction of mulsemedia systems.

**References:**


[1] Eckehard Steinbach et al., "Haptic data compression and communication", *IEEE Signal Processing Magazine*, 28(1), 2011, pp. 87-96

[2] Gheorghita Ghinea and A. Oluwakemi, "Perceived synchronization of olfactory multimedia", *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 40(4), 2010, pp. 657-663

[3] Rajkumar Kannan, S.R. Balasundaram, and Frederic Andres, "The Role of Mulsemedia in Digital Content Ecosystem Design", *Proceedings of the ACM International Conference on Management of Emergent Digital Ecosystems*, 2010, pp. 264-266

[4] J. G. Apostolopoulos, P. A. Chou, B. Culbertson, T. Kalker, M. D.Trott, and S. Wee, "The road to immersive communication," Proc. IEEE, vol. 100, no. 4, Apr. 2012, pp. 974–990


**Frank Hartung** is a full professor of multimedia technology at FH Aachen University of Applied Sciences, Aachen, Germany. He received a MSc in electrical engineering from RWTH Aachen University, Germany, and a PhD in Telecommunications from University of Erlangen, Germany. He has been working with Ericsson Research, as a research team leader in Multimedia Technologies, from 1999 to 2011. In 2008, he was a visiting researcher at Stanford University, Palo Alto, USA, and in 2016, he was a visiting researcher at Eurecom, Sophia-Antipolis, France. His research interests include media security and forensics, networked multimedia, immersive multimedia communication, streaming, and mobile video. He has authored or co-authored more than 50 publications in this domain, and is the co-inventor of 22 granted patents. Dr. Hartung is a member of IEEE and VDE.

# Evaluating Impact of Panoramic Projections of 360-degree Videos on Coding Efficiency

*A short review for "A Framework to Evaluate Omnidirectional Video Coding Schemes"*
*(Edited by **Alisa Devlic**, **Gwendal Simon**, and **Xavier Corbillon**)*

*M. Yu, H. Lakshman and B. Girod, "A Framework to Evaluate Omnidirectional Video Coding Schemes," Mixed and Augmented Reality (ISMAR), 2015 IEEE International Symposium on, Fukuoka, 2015, pp. 31-36.*

360-degree videos (or omnidirectional videos) are gaining popularity with the arrival of virtual reality head-mounted displays (HMDs), offering users an immersive experience in which their head movements are tracked in three dimensions to control their point of view. Omnidirectional videos are large in size due to high resolution and frame rate, while only a portion of the images is displayed to a user on an HMD (referred to as a viewport). Since a high bitrate is required to represent such content, new techniques have to be developed for *preparing* 360-degree videos with respect to their characteristics.

360-degree videos are spherical in nature, while the existing video encoders work with regular rectangular-shaped videos. Therefore, they need to be mapped to a rectangular plane before the encoding takes place. This mapping of a spherical video to different panoramic projections might have large impact on video encoding, resulting in different video qualities and bitrates. This paper studies this problem.

The authors propose a unified framework for evaluating coding efficiency of various projections of omnidirectional videos: equirectangular, equal-area, dyadic and cubic, each of which has a different sampling pattern. Their goal is to compare the quality of different mappings to the original spherical video quality, in order to evaluate the impact of various panoramic projections on the coding efficiency of the video encoder. The main challenges that are investigated in this paper are: *i*) how to compare the quality of high resolution reference video with the same low resolution video, when both videos are encoded in different panoramic projections, and *ii*) how to evaluate the quality of spherical video when all points on the sphere do not have the same viewing probability.

The first challenge of comparing the quality of videos encoded in different projections and resolutions is to deal with a potentially unfair bias towards the projection and resolution in which the reference video is stored. The authors tackled this problem by encoding the reference video in higher resolution and remapping it to different panoramic projections at the lower resolution. The second challenge is related to the fact that points in the spherical video are not equally likely viewable. Authors address it by deriving the relative frequencies of different points from the head motion dataset and computing the average viewport quality.

The coding efficiency of a panoramic projection is evaluated in terms of bitrate savings relative to the equirectangular projection. For the video quality evaluation, the authors generated a viewport corresponding to head movement data, using which they computed the peak signal-to-noise ratio (PSNR). The obtained value represents the estimated quality of the view of the video presented to the user. Next, from all possible viewports the authors computed the weighted video quality. By averaging the quality of these viewports they obtained the so-called spherical PSNR (S-PSNR). An interesting result showed in the paper is that this metrics can estimate the average viewport quality, without requiring the head motion data in advance.

With this approach the authors compared different sphere-to-plane mappings. The framework uses the user-specific head mounted trajectories when they are available; otherwise it relies on general head motion statistics. The results show that equi-area projection has the largest bitrate savings when compared to equirectangular projection and the highest viewport quality of all projections. These results can be beneficial to content providers and multimedia research community to understand the potentials and weaknesses of different 360-degree video transformations.

In summary, the paper addresses an interesting and timely relevant topic. The proposed framework is the first step towards improving the coding efficiency and quality of user experience, as we expect more and more omnidirectional videos with increasing resolutions to be available on the Internet.

Unfortunately, the paper does not investigate the impact of the variability of parameters of different projections on coding efficiency, such as resolution, and image quality. Additionally, other panoramic projections need to be evaluated in the same manner. The results of these investigations might sparkle an interest into development of future 360-degree spherical video encoders.

**Alisa Devlic** is a Postdoctoral researcher at Telecom Bretagne. She received her Master degree in 2005 from University of Zagreb, Croatia, and PhD degree in 2015 from Royal Institute of Technology in Stockholm, Sweden. She worked as a researcher at Wireless@KTH, Ericsson Research, and Appear Networks. Her research interests include video delivery systems, service-oriented network architectures, energy profiles, and context-aware computing.

**Gwendal Simon** is Associate Professor at Telecom Bretagne. He received his Master Degree in Computer Science in 2000 and his PhD degree in Computer Science in December 2004 from University of Rennes 1 (France). From 2001 to 2006 he was a researcher at Orange Labs, where he worked on peer-to-peer networks and social media innovations. Since 2006, he has been Associate Professor at Telecom Bretagne, a graduate engineering school within the Institut Mines-Telecom. He has been a visiting researcher at University of Waterloo from September 2011 to September 2012. His research interests include large scale networks, distributed systems, optimization problems and video delivery systems.

**Xavier Corbillon** is a PhD student at Telecom Bretagne. He received his Master Degree in Telecomunication Engineering in 2014 from Telecom Bregagne. From 2014 to 2015 he was a research engineer at ABB research center in Baden (Switzerland), where he worked on a real-time embedded network simulator. In 2015 he worked on the Eropean project RITE for the Institut Mine-Telecom. His research interest include streaming of interactive video and optimization problems.

# Streaming Uncompressed Video over 60GHz Wireless Networks

*A short review for "Link Adaptation for High-Quality Uncompressed Video Streaming in 60-GHz Wireless Networks"*

*(Edited by **Carl James Debono**)*

*M. Choi, G. Lee, S. Jin, J. Koo, B. Kim, and S. Choi, "Link Adaptation for High-Quality Uncompressed Video Streaming in 60-GHz Wireless Network," IEEE Transactions on Multimedia, vol. 18, no. 4, pp. 627-642, April 2016.*

The millimeter-wave band is an attractive band for broadband communications. Recent advances in wireless technologies have made this band available where wide bandwidth is available for short range communications. The limited range allows for frequent spatial reuse of the frequencies available allowing for wide use in different applications. Transmission of video content in bandwidth-limited channels demands complex compression algorithms that increase latency and to a certain extent, depending on the lossy compression and its amount, distortions in the video content. Having a huge bandwidth available can provide the means to transmit uncompressed video content, even for high resolutions, leading to lower latency and improved quality.

Standardization groups have been working to develop 60GHz wireless standards in the last few years. IEEE 802.15.3c was the first IEEE standard in this band to offer data rates that exceed 1Gb/s [1]. Another standard is ECMA-387 [2] which employs ECMA-368 [3] medium access control (MAC) and the 60GHz physical layer. This standard also defines wireless high-definition multimedia interface (HDMI) [4] phase alternating line (PAL) interface that supports the streaming of uncompressed video. Other efforts include the IEEE 802.11ad [5] that is intended to extend the coverage of the wireless local area networks (WLAN) to the 60GHz band.

The MAC protocols of these standards allocate resources in advance for high-quality video streaming. This guarantees that there is no contention for the reserved blocks and large amounts of data can be streamed. The authors of the original paper adopt the ECMA-387 standard for their study. This choice was influenced by the definition of HDMI PAL which supports uncompressed video streaming.

Most wireless channels have limited bandwidth available and cannot sustain uncompressed high definition video streams at gigabits per second (Gbps) rates. Thus, an attractive characteristic of the millimeter band is that it allows multi-Gbps links. This allows streaming of uncompressed content removing the overhead incurred during encoding and decoding of the video. In doing so, the quality of the video is also improved as distortions are removed and error propagation due to video encoding dependencies cannot occur.

To ensure reliable transmission in the lossy wireless environment, the system employs a number of modulation and coding schemes (MCS). The use of a more robust MCS provides more protection of the content resulting in a better quality of service. This however comes at the expense of more bandwidth requirements. This might be an issue as the band is unlicensed and can be shared with other devices. Having a higher MCS allows better use of the resources but the video quality might degrade as the probability of transmission errors increases. This means that a link adaptation algorithm is necessary to select the appropriate MCS to trade video quality and efficient use of resources.

The authors of the original paper propose two MAC layer link adaptation policies. These are: (1) maximize the video quality for the given wireless resources, and (2) minimize the resources needed to sustain a particular video quality. They also propose a new metric called "expected peak signal-to-noise ratio (ePSNR)" to determine the MCS value. Moreover, unequal error protection (UEP) [6, 7] is employed to protect more important information.

The policy to maximize video quality means that the transmitter has to find a configuration that provides maximum ePSNR for the allocated resources. This outputs the best MCS that satisfies this constraint. The other policy tries to minimize the resources needed by each device. This is especially important when the capacity must be shared with other users of the network. The more resources saved, the more users the system can accommodate. Another constraint imposed by the shared medium is that the algorithm must be distributed such that multiple streaming services can be satisfied for multiple users. This is satisfied by the distributed resource

management scheme developed by the original authors. The algorithm strives to maximize PSNR of ongoing streams if resources are available. Failing this, it administers new streams by guaranteeing at least a fair quality between users. This means that the impact on ongoing streams is minimized.

The results in the original paper show that the latency of the video is significantly reduced with figures reporting values lower than 50ms. Furthermore, the reported results show that high quality video can be achieved with efficient resource allocation.

Millimeter wave communication is a very promising solution for broadband video transmission. However, the size of the data and the errors on the channel still limit the quality of experience of the users. Some better means of error protection and scalable solutions that can free resources when needed to allow more users need to be devised. Other resource allocation schemes that guarantee fairness amongst users accessing services of such networks while being energy efficient are still an open area of research.

**References:**


[1] T. Baykas, C.-S. Sum, Z. Lan, J. Wang, M.A. Rahman, H. Harada, S. Kato, "IEEE 802.15.3c: The first IEEE wireless standard for data rates over 1 Gb/S," *IEEE Communications Magazine*, pp. 114–121, July 2011.

[2] *ECMA International 387, High rate 60GHz PHY, MAC and HDMI PAL*, standard ECMA-387, 2nd ed., December 2010.

[3] *ECMA International 368, High rate ultra wideband PHY and MAC*, standard ECMA-368, 3rd ed., December 2008.

[4] *High-definition multimedia interface (HDMI) specification*, version 1.4, May 2009.

[5] *Wireless LAN medium access control (MAC) and Physical layer (PHY) Specifications: Enhancements for very high throughput in the 60GHz band part 11*, IEEE P802.11ad, December 2012.

[6] H. Singh, J. Oh, C. Kweon, X. Qin, H.-R. Shao, and C. Ngo, "A 60GHz wireless network for enabling uncompressed video communications," *IEEE Communications Magazine*, vol. 46, no. 12, pp. 71–78, December 2008.

[7] Z. Lan, J. Wang, C.-S. Sum, T. Baykas, C. Pyo, F. Kojima, H. Harada, and S. Kato, "Unequal error protection for compressed video streaming on 60GHz WPAN sytsem," in *Proc. of the International Wireless Communications and Mobile Computing Conference*, pp. 689–693, August 2008.


**Carl James Debono** (S'97, M'01, SM'07) received his B.Eng. (Hons.) degree in Electrical Engineering from the University of Malta, Malta, in 1997 and the Ph.D. degree in Electronics and Computer Engineering from the University of Pavia, Italy, in 2000.

Between 1997 and 2001 he was employed as a Research Engineer in the area of Integrated Circuit Design with the Department of Microelectronics at the University of Malta. In 2000 he was also engaged as a Research Associate with Texas A&M University, Texas, USA. In 2001 he was appointed Lecturer with the Department of Communications and Computer Engineering at the University of Malta and is now an Associate Professor. He is currently the Head of the Department of Communications and Computer Engineering at the University of Malta.

Prof. Debono is a senior member of the IEEE and served as chair of the IEEE Malta Section between 2007 and 2010. He was the IEEE Region 8 Vice-Chair of Technical Activities between 2013 and 2014. He has served on various technical program committees of international conferences and as a reviewer in journals and conferences. His research interests are in wireless systems design and applications, multi-view video coding, resilient multimedia transmission, and modeling of communication systems.

# Clustering 100 Million Images on A Single Machine in Less Than One Hour

*A short review for "Web-Scale Image Clustering Revisited"*
*(Edited by **Jun Zhou**)*

*Yannis Avrithis, Yannis Kalantidis, Evangelos Anagnostopoulos, Ioannis Z. Emiris. "Web-Scale Image Clustering Revisited", International Conference on Computer, pages 1502-1510, 2016.*

The big data era has created great opportunities and challenges to image retrieval and mining. In practice, images are represented as vectorized descriptors. Matching or retrieving millions of them relies on approximate nearest neighbor (ANN) techniques which aim to to find an approximate matching rather than the exact one. Many methods have been proposed for this purpose, for example, locality-sensitive hashing [1] and kd-tree based search [2].

An important step to boost the performance of ANN search is data clustering, which groups images based on the similarities of their descriptors. Then the matching can first be made between a query and the cluster centers and then expands to the data point within the nearest clusters. However, when the dimension of the image descriptors is large, significant computation and storage costs may incur, making the scaling of clustering to web-scale data difficult.

In this paper, an inverted-quantized k-means method is introduced to tackle this scalability problem. Extended from inverted approximate k-means approach that searches for the nearest data points using clustering centers as queries, this method seeks a combination of two-dimensional grid data representation for subspace quantization and a probabilistic model for overlap cluster estimation and automatic cluster merging. Encouraging performance of this method has been achieved, with 100 million images being clustered on a single machine in less than one hour.

This method starts from the data representation step that is based on multi-indexing [3]. The general idea is to decompose a data vector into two sub-vectors in two orthogonal subspaces. Vectors in these two subspaces can be used to produce two codebooks whose codewords are assigned with indices. This allows a grid be generated so that each data point can be assigned to the nearest codewords and take their indices as the representation. In this way, each data is associated with a cell in the grid and data distribution over the grid can be obtained. An advantage of this step is that the descriptors of the original data are no longer needed as the new representation is simply based on cell indices, the probability distribution that data fall into each cell, and the mean of all data in each cell.

With the data quantization done, a cluster center can be calculated as the weighted average of all samples in the cluster. To this stage, an inverted fast search strategy is adopted to avoid traditional time consuming data assignment process during the clustering iterations. It takes advantage of the fixed grid representation, and seeks cells in the neighborhood of each cluster center for inverse assignment. This step is similar to ranked retrieval [4].

The introduced method also allows dynamic cluster merging by adopting an expanding Gaussian mixtures approach [5]. This is based on center to center searching, so that neighboring clusters can be found and their overlapping estimated. If too much overlapping between two clusters is detected, they are merged.

Authors presented the performance of their method on three datasets that contain from 500K to 100 million images respectively. On these datasets, image features are either extracted using convolutional neural networks or SIFT approach for codebook generation. When evaluated by efficiency, this method is about 70 times faster than a distributed implementation of k-means clustering using 300 machines on the 100 million dataset, though with sacrificed precision.

In summary, this paper has most required components for developing a fast clustering method on huge image datasets. These include subspace based grid data quantization so as to reduce dimensionality of image descriptors, inverted cluster neighborhood search to avoid time consuming data to cluster assignment process, and dynamically reduction of the number of clusters for improved efficiency in each clustering step. The whole process forms a useful foundation for other large scale data matching, retrieval, and mining tasks, in which clustering huge

amount of data is probably one of the most time consuming tasks.

**References:**


[1] P. Indyk and R. Motwani, "Approximate nearest neighbors: towards removing the curse of dimensionality," in Proceedings of the Annual ACM symposium on Theory of computing, 1998, pp. 604–613.

[2] J. Beis and D. Lowe, "Shape indexing using approximate nearest-neighbour search in high-dimensional spaces," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1000–1006, 1997.

[3] A. Babenko and V. Lempitsky, "The inverted multi-index," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3069-3076, 2012.

[4] A. Broder, L. Garcia-Pueyo, V. Josifovski, S. Vassilvitskii, and S. Venkatesan. "Scalable k-means by ranked retrieval," in Proceedings of the 7th ACM international conference on Web search and data mining, pp. 233-242, 2014.

[5] Y. Avrithis and Y. Kalantidis. "Approximate Gaussian mixtures for large scale vocabularies," In Proceedings of the 12th European conference on Computer Vision, Part III, pp 15-28, 2012.


**Jun Zhou** received the B.S. degree in computer science and the B.E. degree in international business from Nanjing University of Science and Technology, China, in 1996 and 1998, respectively. He received the M.S. degree in computer science from Concordia University, Canada, in 2002, and the Ph.D. degree in computing science from University of Alberta, Canada, in 2006.

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

# 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:

Christian Timmerer (christian.timmerer@aau.at) and Yan Zhang (yanzhang@simula.no).

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

**Chair**: Yonggang Wen, Singapore
**Steering Committee Chair**: Luigi Atzori, Italy
**Vice Chair – North America**: Khaled El-Maleh, USA
**Vice Chair – Asia**: Liang Zhou, China
**Vice Chair – Europe**: Maria G. Martini, UK
**Vice Chair – Letters**: Shiwen Mao, USA
**Secretary**: Fen Hou, China
**Standard Liaison**: Zhu Li, USA

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.