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

MMTC Communications – Review

IEEE COMMUNICATIONS SOCIETY

Vol. 12, No. 5, October 2021

TABLE OF CONTENTS

Message from the Review Board Directors

Welcome to the February 2021 issue of the IEEE ComSoc MMTC Communications – Review.

This issue comprises five reviews that cover multiple facets of multimedia communication research including data classification, human pose estimation, node classification, and multi-view video compression. These reviews are briefly introduced below.

The first paper, published in ACM The 26th Annual International Conference on Mobile Computing and Networking (MobiCom '20), Sep. 2020, and edited by Dr. Takuya Fujihashi, study focuses on the point cloud (PtCl) representation for volumetric video streaming.

The second paper is published in IEEE Wireless Communications and edited by Dr. Shengjie Xu. It studied the worker selection issues to ensure reliable federated learning in mobile networks.

The third paper, published in IEEE Transactions on Multimedia and edited by Dr. Jinbo Xiong, focus on a fundamental problem for video service enhancement value controlling whether communication should continue for another step.

The fourth paper, published in IEEE Transactions on Image Processing and edited by Prof. Cao Guitao. This paper proposes a novel study scenario of zero-shot learning, in which it is utilized to index images for scalable retrieval.

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

Zhisheng Yan
George Mason University, USA
Email: zyan4@gmu.edu

Yao Liu
Binghamton University, USA
Email: yaoliu@binghamton.edu

Wenming Cao
Shenzhen University, China
Email: wmcao@szu.edu.cn

Phoenix Fang
California Polytechnic State University, USA
Email: dofang@calpoly.edu/span>

A Visibility-Aware Mobile Volumetric Video Streaming

A short review for "ViVo: Visibility-Aware Mobile Volumetric Video Streaming"

Edited by Takuya Fujihashi

B. Han. Y. Liu, F. Qian, "ViVo: Visibility-Aware Mobile Volumetric Video Streaming," ACM The 26th Annual International Conference on Mobile Computing and Networking (MobiCom'20), Sep. 2020.

Immersive video streaming plays an important role in the next-generation mobile networks, i.e., 5G and 6G. Volumetric video streaming is one of the enabling technologies for MR. Although there are various volumetric data formats, this study focuses on the point cloud (PtCl) representation for volumetric video streaming.

Streaming PtCl videos is resource-demanding. Uncompressed PtCl data is often prohibitively large for streaming over wireless networks. Each PtCl frame consists of a set of unsorted and non-uniformly distributed 3D points with attributes of 3D coordinates and color components. The representation of each 3D point typically takes 15 bytes; 4 bytes for each attribute of the 3D coordinates, i.e., (X, Y, Z) and 1 byte for each color component. For example, a sender streams the PtCl video with 200K points per frame at 30 fps requires 15×200K×30×8=720Mbps of bandwidth. To efficiently compress such numerous and irregular structures of 3D points while keeping original 3D scenes/objects, compression/decompression solutions have been proposed [1-3] in recent years. However, the recent compression of the PtCl still cannot provide a better QoE for commodity mobile devices because of either high computational overhead and poor network conditions.

In this paper, the authors address the problems of PtCl streaming over wireless networks. The contributions of this paper are three-fold. The first contribution is the detailed investigation of PtCl encoding, decoding, segmentation, and viewport movement patterns in mobile devices. To discuss the state-of-the-art performance of the PtCl compression and segmentation, the authors carry out experimental evaluations using PtCl video datasets obtained from multiple Kinects. Each PtCl video is encoded and decoded using typical open-source PtCl compression solutions to compare the PtCl compression/decompression performance: Draco (k-d tree-based compression [4]), Point Cloud Library (octree-based compression [5]), and Limited Error Point Cloud Compression (extended version of limited error raster compression [6]). The authors compare the compression and decompression performance with/without segmentation (i.e., cell division) using the PtCl video datasets across the solutions. From the evaluation results, the required bitrate of PtCl video is up to 180Mbps even after the compression. Since the available datarate for wireless networks is still limited, such a high bitrate serves as a key motivation of visibility-aware optimizations to significantly reduce bandwidth for the PtCl video streaming.

In addition, they studied viewport movement and prediction for volumetric videos. The authors collect viewport trajectory of smartphone and MR headset, i.e., Magic Leap One, users during the PtCl video playback. They found three observations. The first one is the users seldom move vertically. The second one is the translational movement between the smartphone and headset users is different. Specifically, the movement appears straight in the smartphone users while natural body movement in the headset users. The third one is that the number of objects in the PtCl video affects movement.

In terms of the viewport prediction, they use two lightweight machine learning models. The authors use linear regression (LR) and multilayer perceptron (MLP) to predict each dimension of X, Y, Z, yaw, and pitch from the past translational movement. As a result, the authors found there is no qualitative difference between the accuracy of LR- and MLP-based viewport prediction and decided to use LR for viewport prediction.

The second contribution is to propose a volumetric streaming system called ViVo (Visibility aware Volumetric video streaming). The proposed ViVo can deliver high-quality volumetric content to commodity mobile devices. To further reduce bandwidth consumption for streaming PtCl videos, ViVo integrates three key optimizations: Viewport Visibility (VV), Occlusion Visibility (OV), and Distance Visibility (DV). The basic concepts of VV, OV, and DV are to fetch volumetric content that overlaps with the predicted viewport, to reduce the point density of each cell based on the occlusion level of the cell, and to adjust the point density of the cell based on the viewpoint-to-cell distance, respectively. The concept of VV is inspired by viewport adaptive 360-degree video streaming [7]. On the other hand, OV and DV are unique optimizations to the PtCl videos which consider the depth information of the PtCl data.

The third contribution is to implement the proposed ViVo for commodity devices. The authors implement the ViVo PtCl video player on Android devices and its video server on Linux. Specifically, they use off-the-shelf Android smartphones of SGS8 and SGS10 for the video players and server with Intel Xeon CPU E5-2680 v4@2.40GHz for the video server. We thoroughly evaluate the implemented ViVo over diverse wireless/mobile networks (including WiFi, LTE, and commercial mmWave 5G), headset/smartphone users using both objective metrics of structural similarity index (SSIM) and subjective metric of mean opinion score (MOS).

From evaluation results, the authors highlight the evaluation results as follows:

- When the bandwidth for the PtCl streaming is sufficiently high, each optimization of VV, OV, and DV can reduce the required bit-rate for the PtCl streaming without perceived quality loss (SSIM >0.99).
- The proposed ViVo can reduce the required bandwidth for the commercial 5G mmWave networks while maintaining good visual quality and short stall time.
- Even when the network bandwidth is constrained/fluctuated, the proposed ViVo can realize better subjective performance under the same required bit-rate.

References:

[1] Draco 3D Data Compression. https://google.github.io/draco/.

[2] Point Cloud Library (PCL). http://pointclouds.org/.

[3] Limited Error Point Cloud Compression. https://github.com/Esri/lepcc.

[4] D. Chen, Y.-J. Chiang, and N. Memon. Lossless Compression of Point-Based 3D Models. In Proceedings of the 13th Pacific Conference on Computer Graphics and Applications, 2005.

[5] T. Golla and R. Klein. Real-time Point Cloud Compression. In Proceedings of International Conference on Intelligent Robots and Systems, 2015.

[6] Limited Error Raster Compression. https://github.com/Esri/lerc.

[7] J. He, M. A. Qureshi, L. Qiu, J. Li, F. Li, and L. Han. Rubiks: Practical 360-Degree Streaming for Smartphones. In Proceedings of ACM MobiSys, 2018.

Takuya Fujihashi received the B.E. degree in 2012 and the M.S. degree in 2013 from Shizuoka University, Japan. In 2016, he received Ph.D. degree from the Graduate School of Information Science and Technology, Osaka University, Japan. He is currently an assistant professor at the Graduate School of Information Science and Technology, Osaka University since April, 2019. He was an assistant professor at the Graduate School of Science and Engineering, Ehime University, Japan from Jan. 2017 to Mar. 2019. He was research fellow (PD) of Japan Society for the Promotion of Science in 2016. From 2014 to 2016, he was research fellow (DC1) of Japan Society for the Promotion of Science. From 2014 to 2015, he was an intern at Mitsubishi Electric Research Labs. (MERL) working with the Electronics and Communications group. He selected one of the Best Paper candidates in IEEE ICME (International Conference on Multimedia and Expo) 2012. His research interests are in the area of video compression and communications, with a focus on multi-view video coding and streaming.

A Worker Selection Scheme for Reliable Federated Learning in Mobile Networks

A short review for "Reliable Federated Learning for Mobile Networks"

Edited by Shengjie Xu

J. Kang, Z. Xiong, D. Niyato, Y. Zou, Y. Zhang and M. Guizani, "Reliable Federated Learning for Mobile Networks," IEEE Wireless Communications, vol. 27, no. 2, pp. 72-80, April 2020.

This paper studied the worker selection issues to ensure reliable federated learning in mobile networks. A reputation-based scheme was designed to select reliable and trusted workers. In order to achieve efficient and secure reputation management, we calculated workers' reputation by using a multi-weight subjective logic model, and employed consortium blockchain to manage the reputation with tamper resistance and non-repudiation in a decentralized manner. The authors conducted a study and numerical results showed that our schemes can bring reliable federated learning to mobile networks.

The paper first introduced the background of emerging wireless networks. Mobile devices, such as smart phones or vehicles, equipped with a variety of sensors, generate a huge amount and diverse types of user data. Recently, for greatly improving mobile services and enabling smarter mobile applications, it is increasingly popular to utilize machine learning technologies to train models on such user data.

The paper then presented the background of federated learning. To address the privacy challenges, a decentralized machine learning paradigm namely, federated learning, has been proposed to enable mobile devices to collaboratively train a global model required by a central aggregator in a decentralized manner, without the need of centrally storing raw training data. In federated learning, mobile devices download a global model from the central aggregator in each iteration, and then train and improve the current global model by using their local raw data. The mobile devices send the local model updates to the central aggregator. By aggregating these local model updates, the central aggregator generates a new global model for the next iteration. Both the mobile devices and the central aggregator repeat the above process until the global model achieves a certain accuracy. This paradigm significantly reduces risks of privacy leakage by decoupling of model training from the need for accessing raw training data [1].

Although federated learning brings great benefits for mobile networks, it is still susceptible to various adversarial attacks in its primary stage. That is, during a federated learning process, data owners may mislead a global model by intentional or unintentional behaviors [2]. For intentional behaviors, an attacker can send malicious updates, that is, the poisoning attack, to affect the global model parameters resulting in the failure of current collaborative learning. The authors in [3] demonstrated the vulnerability of federated learning to sybil-based poisoning through experiments, and showed that existing defenses to such attacks are ineffective.

Motivated by the descriptions, the authors proposed that reputation can be used to provide solutions to select reliable and trusted workers for the federated learning tasks. Specifically, the authors presented reputation as a reliable metric to select trusted workers for reliable federated learning, for defending against unreliable model updates. In addition to that, authors also applied a multi-weight subjective logic model to design an efficient reputation calculation scheme according to both task publishers' interaction histories and recommended reputation opinions. Finally in order to achieve secure reputation management, the authors managed the reputation in a decentralized manner by employing the consortium blockchain deployed at edge nodes.

The authors reviewed the concept of federated learning and its four popular applications. Authors then reviewed the security challenges of federated learning, by highlighting three key challenges for worker selection: no reliable and fair metrics to evaluate workers, no efficient and universal worker selection schemes, and no timely monitoring methods for workers.

The authors then investigated the reputation management for reliable federated learning. In particular, the authors offered an overview of reputation management in crowdsensing. The mobile devices collect local sensing data and generate various user data from mobile applications. Mobile applications with federated learning perform model training by using these data without the need of data aggregation for privacy preservation [4]. A reputation-based worker selection scheme with consortium blockchain is presented with five steps.

To assess the trustworthiness of a worker candidate, reputation opinions from task publishers are considered collected and integrated into a composite reputation value of the worker candidate for secure worker selection. Based on this observation, the authors utilize the subjective logic model to calculate composite reputation values of worker candidates for efficiency.

Simulation was performed by using the MNIST dataset to evaluate the performance of the proposed schemes. Accuracy to the federated learning was shown with respect to different poisoning attack strengths and Earth Mover's Distance (EMD). There are three factors that affect the learning accuracy: EMD, attacker number, and attack strength. The proposed scheme can achieve a more accurate and fair reputation calculation, thereby leading to a more reliable worker selection in federated learning.

Finally, authors indicated several possible directions. More accurate and efficient validation schemes for non-IID datasets should be designed to improve the detection performance of poisoning attacks in the proposed worker selection schemes. In addition, efficient schemes for optimizing the number of workers are worth investigation in order to balance learning performance and resource cost.

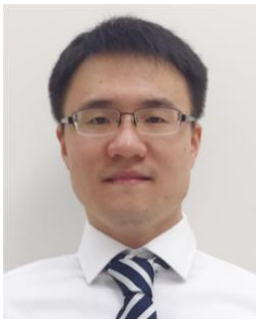
References:

[1] X. Zhu et al., "Blockchain-Based Privacy Preserving Deep Learning," Proc. Int'l. Conf. Information Security and Cryptology, Springer, Cham, 2018, pp. 370–83.

[2] M. Shayan et al., "Biscotti: A Ledger for Private and Secure Peer-to-Peer Machine Learning," 2018; available: https:// arxiv.org/abs/1811. 09904.

[3] C. Fung et al., "Mitigating Sybils in Federated Learning Poisoning," 2018; available: https://arxiv.org/abs/1808.04866.

[4] J. Kang et al., "Incentive Mechanism for Reliable Federated Learning: A Joint Optimization Approach to Combining Reputation and Contract Theory," IEEE Internet of Things J., vol. 6, no. 6, Dec. 2019, pp. 10700–714.



Shengjie Xu [SM'14-M'19] received a Ph.D. degree in Computer Engineering from University of Nebraska-Lincoln, and an M.S. degree in Telecommunications from University of Pittsburgh. Before that, he held a B.E. degree in Computer Science and Information Security. Presently, he is an Assistant Professor of Computer and Cyber Sciences in the Beacom College of Computer and Cyber Sciences at Dakota State University. His research interests include AI-driven cybersecurity, secure edge computing, and critical infrastructure protection. He serves as a Technical Editor for IEEE Wireless Communications and an Editor for International Journal of Sensor Networks. He was awarded the 2020 IET Journals Premium Award for Best Paper. He is a member of IEEE, ACM, and AAAI. He is also involved in IEEE Technical Committees, including Communications and Information Security (CISTC), Green Communications and Computing (TCGCC), and Big Data (TCBD). He holds professional certifications in cyber security and computer networking.

Hybrid Human-Artificial Intelligence-Based Video Service Enhancement

A short review for "Edge-Cloud Collaboration Enabled Video Service Enhancement: A Hybrid Human-Artificial Intelligence Scheme"

Edited by Jinbo Xiong

D. Wu, R. Bao, Z. Li, H. Wang, H. Zhang and R. Wang, "Edge-Cloud Collaboration Enabled Video Service Enhancement: A Hybrid Human-Artificial Intelligence Scheme," in IEEE Transactions on Multimedia, doi: 10.1109/TMM.2021.3066050.

Recently, there are more and more types of video services emerging, such as live steaming, short videos, and video on demand and etc., which leads to a more complex service provision for the service operators [1]. In order to guarantee diverse service requirements for different users, it is necessary to study how to utilize the limited communication, computing and cache (3C) resources of a considered network to serve as many people as possible [2-3]. Usually, researchers try to improve the quality of service (QoS) and quality of experience (QoE) of video services for users through video cache and video delivery [4-5]. However, most of works just study video cache and video delivery separately, which results in sub-optimization of network performance. Recently, some efforts have been devoted to design integration of video cache and delivery schemes. Nevertheless, the state-of-the-art schemes are still with high computation complexity, which may not stratify the requirement of delay-sensitive video services.

In this paper, the authors focus on a fundamental problem for video service enhancement, i.e., how to send videos to the users as soon as possible. An edge-cloud collaboration framework is proposed to enable jointly video cache and delivery optimization. Under the proposed framework, video cache and delivery processes can be decoupled into two optimization problem. Regarding video cache, the authors try to let as many users as possible be served at the network edge and as many cached videos as possible are hit by users at the same time. For video delivery, the objective is to improve the video coding rate for each user under fairness constraints. In particular, the authors analyze the shortcakes of artificial intelligence tool [6] and traditional optimization tool in network optimization, and then propose a novel hybrid human-artificial intelligence idea to deal with the video service enchantment problem. Specific contributions devoted by the authors can be summarized into the following three aspects.

Firstly, in order to improve the cache hit rate of edge caching, the authors propose to select candidate videos based on user interest. An intelligent factorization machine and multi-layer perceptron merging scheme is designed to represent both low-order and high-order features simultaneously. The proposed scheme is able to guarantee high prediction accuracy on user interest of different videos.

Secondly, in order to hit the video requests from as many users as possible, the authors further propose to cache content based on group interest. A social aware similarity model is formulated to characterize the similarity between individual user and the group. In addition, a group interest model is constructed with consideration of the impacts of user similarity, positive emotion and negative emotion. The proposed group interest model is able to guarantee high user hit rate and high content hit rate at the same time.

Thirdly, in order to improve the QoE of each user and guarantee user fairness at the same time. The authors propose to apply network calculus theory to derive a statistical delay guarantee model of video services. Then a double bisection exploration scheme is further proposed to guide video delivery. The computation complexity of the proposed scheme is proved to be logarithmic level.

Thereafter, the advantage of the proposed video cache and delivery scheme is validated based on a

real-world dataset. In detailed, the performance of video cache scheme is evaluated with the baseline schemes such as popularity-based caching scheme, individual interest-based scheme. In addition, the proposed video delivery scheme is compared with the ones while ideal edge caching, only cloud caching are assumed respectively. The authors carry out extensive experiments to verify that the proposed scheme performs better in terms of content hit rate, user hit rate, and video coding rate.

In summary, the proposed edge-cloud collaboration-based video service framework as well as the hybrid human-artificial intelligence schemes therein is proved to be applicable to improve QoS and QoE for users with diverse video service requirements. The authors make a breakthrough in the approach of video service enhancement, which provides valuable insights into delay-sensitive service optimization.

References:

[1] Y. Sani, A. Mauthe and C. Edwards, "Adaptive Bitrate Selection: A Survey," in IEEE Communications Surveys & Tutorials, vol. 19, no. 4, pp. 2985-3014, Fourthquarter 2017, DOI: 10.1109/COMST.2017.2725241.

[2] M. Sheng, C. Xu, J. Liu, J. Song, X. Ma and J. Li, "Enhancement for content delivery with proximity communications in caching enabled wireless networks: architecture and challenges," in IEEE Communications Magazine, vol. 54, no. 8, pp. 70-76, Aug. 2016, DOI: 10.1109/MCOM.2016.7537179.

[3] N. Li, Y. Hu, Y. Chen and B. Zeng, "Lyapunov Optimized Resource Management for Multiuser Mobile Video Streaming," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 6, pp. 1795-1805, Jun. 2019, DOI: 10.1109/TCSVT.2018. 2850445.

[4] S. Mehrizi, S. Chatterjee, S. Chatzinotas and B. Ottersten, "Online Spatiotemporal Popularity Learning via Variational Bayes for Cooperative Caching," in IEEE Transactions on Communications, vol. 68, no. 11, pp. 7068-7082, Nov. 2020, DOI: 10.1109/TCOMM. 2020. 3015478.

[5] M. Choi, A. No, M. Ji and J. Kim, "Markov Decision Policies for Dynamic Video Delivery in Wireless Caching Networks," in IEEE Transactions on Wireless Communications, vol. 18, no. 12, pp. 5705-5718, Dec. 2019, DOI: 10.1109/TWC.2019.2938755.

[6] J. Xiong, M. Zhao, M. Bhuiyan, L. Chen, Y. Tian, "An AI-enabled three-party game framework for guaranteed data privacy in mobile edge crowdsensing of IoT," in IEEE Transactions on Industrial Informatics, vol. 17, no. 2, pp. 922-933, Feb. 2021, DOI:10.1109/ TII.2019.2957130.

Jinbo Xiong, is a Professor and Ph.D. supervisor with the Fujian Provincial Key Laboratory of Network Security and Cryptology and the College of Computer and Cyber Security at Fujian Normal University. He received the M.S. Degree in Communication and Information Systems from Chongqing University of Posts and Telecommunications, Chongqing, China, in 2006, and the Ph.D. degree in Computer System Architecture from Xidian University, China, in 2013. He was a visiting scholar with the Department of Computer Science and Engineering at the University of North Texas, Denton, USA, from 2019-2020. His research interests include connected and autonomous vehicle, secure deep learning, cloud data security, mobile media management, privacy protection, and Internet of Things. He has published more than 100 papers in prestigious journals such as IEEE WCM, IEEE TII, IEEE TCC, IEEE IoT J, IEEE TNSE, FGCS and in major International conferences such as IEEE ICCCN, IEEE TrustCom, IEEE HPCC and IEEE ICPADS. He has applied 15 patents and two monographs in these fields. He is a member of IEEE and senior member of CCF.

A Novel Approach for Image Indexing in Zero-Shot Way

A short review for "Zero-Shot Learning to Index on Semantic Trees for Scalable Image Retrieval"

Edited by Guitao Cao

S. Kan, Y. Cen, Y. Cen, M. Vladimir, Y. Li and Z. He, "Zero-Shot Learning to Index on Semantic Trees for Scalable Image Retrieval," IEEE Transactions on Image Processing, vol. 30, pp. 501-516, 2021

Zero-shot learning is a powerful and promising learning paradigm, in which the classes covered by training examples are disjoint with the sample of target classification [1]. Zero-shot learning has been studied widely in image classification [2], stance detection [3], video classification [4], etc. In the existing research, zero-shot learning is generally exploited to supervised classification with auxiliary information. This paper proposes a novel study scenario of zero-shot learning, in which it is utilized to index images for scalable retrieval.

Due to the explosive growth of images nowadays, highly efficient image indexing and scalable retrieval methods have become necessary. There are many image indexing schemas proposed to solve the problem. Codebook-based indexing method [5] produces a set of representative centroid codewords in the high-dimensional feature space, with each feature to be indexed and quantized to the nearest codeword. This method requires a trade-off between efficiency and performance. Tree-based indexing is widely used for scalable and fast deployment at large scales [6][7]. During the study of this method, the time-constrained approximation method [8] has been found to produce better performance than the error-constrained approximate NN search [9]. Graph-based indexing schemes [10] aim to build greedy rooting navigation on a group of datasets to realize fast NN (Nearest Neighbor) retrieval. For a given adjacent graph, the search starts at a point and iteratively traverses the graph. At each step of the traversal, the algorithm checks the distance to the neighbor of the current base node, and then selects one of the neighbors as the next base node with the smallest distance. There are still distance calculations need to be reduced in this method.

The above methods are distance-based, thus they suffer from information loss during matching or quantization. The learning-based indexing methods like Prob-RAW [11] aim to learn the neighborhood relationships embedded in the index space and nearest neighbor probabilities based on the query feedback. However, this method requires the queries in the database to obtain the NN probabilistic nonlinear mapping.

The authors propose a zero-shot learning-based method, using LTI-ST network to obtain a hierarchical semantic tree structure for image indexing and retrieval, which represents the inherent correlation. In this zero-shot way, the index of test images can be generated automatically without any specific analysis.

Major contribution of this paper is to propose a zero-shot learning-based image index scheme, which do not require the analysis of test images. This scheme learns the feature embedding and indexing model in an end-to-end manner, avoids the spending of distance calculation with scalable implementation.

To obtain LTI-ST networks, the schema carries sematic tree encoding, generating the schema tree labels for each class of training images. The authors use deep neural network model with fine-tuning to encode each image into a feature vector. Then the vector is embedded to obtain higher category discrimination. Based on the semantic feature, the authors perform hierarchical clustering to construct the binary codeword tree, with the top layers showing significant difference between images and bottom layers showing subtle difference.

The LTI-ST networks include feature embedding network, labels prediction networks and index prediction network, this model can be used to directly predict the index of the input image. The labels prediction networks consist of hard prediction network, soft prediction network and ranked prediction network. They receive feature embedding and each output a feature vector. The

index prediction network uses the concatenation of the above four outputs as an input, and then output the image index.

The extensive experimental results indicate that the LTI-ST method outperforms the existing state-of-the-art indexing methods. Besides, the ablation studies demonstrate the effectiveness of individual components.

In summary, the proposed LTI-ST schema provides a novel method for image indexing and retrieval. With reducing the complexity of the method, it applies a direct way to predict the image's index. It also extends the study of zero-shot learning to image indexing, which avoids requirement of test images, making it scalable.

References:

[1] Wang, Wei , et al. "A Survey of Zero-Shot Learning: Settings, Methods, and Applications." ACM Transactions on Intelligent Systems and Technology, vol. 10, no. 2, pp. 1-37, 2019.

[2] Liu, Z. , et al. "Convolutional prototype learning for zero-shot recognition." Image and Vision Computing, vol. 98, no.3, 2020.

[3] Emily Allaway, Malavika Srikanth and Kathleen McKeown. "Adversarial Learning for Zero-Shot Stance Detection on Social Media." Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 2021.

[4] Brattoli, J. Tighe, F. Zhdanov, P. Perona, and K. Chalupka, "Rethinking zero-shot video classification: End-to-end training for realistic applications," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.(CVPR), Jun. 2020, pp. 4613–4623

[5] R. Liu, S. Wei, Y. Zhao, and Y. Yang, "Indexing of the CNN features for the large scale image search," Multimedia Tools Appl., vol. 77, no. 24, pp. 32107–32131, Dec. 2018.

[6] A. Beygelzimer, S. Kakade, and J. Langford, "Cover trees for nearest neighbor," in Proc. 23rd Int. Conf. Mach. Learn. (ICML), Pittsburgh, PA, USA, Jun. 2006, pp. 97–104

[7] C. Silpa-Anan and R. Hartley, "Optimised KD-trees for fast image descriptor matching," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Anchorage, AK, USA, Jun. 2008, pp. 24–26.

[8] J. S. Beis and D. G. Lowe, "Shape indexing using approximate nearestneighbour search in high-dimensional spaces," in Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., San Juan, Puerto Rico, Jun. 1997, pp. 1000–1006.

[9] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching fixed dimensions," J. ACM, vol. 45, no. 6, pp. 891–923, Nov. 1998.

[10] C. Fu, C. Xiang, C. Wang, and D. Cai, "Fast approximate nearest neighbor search with the navigating spreading-out graph," Proc. VLDB Endowment, vol. 12, no. 5, pp. 461–474, Jan. 2019

[11] C.-Y. Chiu, A. Prayoonwong, and Y.-C. Liao, "Learning to index for nearest neighbor search," IEEE Trans. Pattern Anal. Mach. Intell., vol. 42, no. 8, pp. 1942–1956, Aug. 2020.

Guitao Cao obtained her Ph.D. in 2006 from Shanghai Jiao Tong University with a focus on pattern recognition. She is currently a professor of Software Engineering Institute. East China Normal University (ECNU), Shanghai. ECNU is the top tier university in China with a high rank (Level A) in Software Engineering in China. She was also a visiting researcher with University of Missouri at Columbia. She has published decades of peer reviewed papers in top venues including IEEE Transactions on Cybernetics, IEEE Transactions on Multimedia, and IEEE Transactions on Biomedical Engineering. Prof. Cao is also the Principal Investigator for many research funding with major sponsors including the National Science Foundation of China, Ministry of Industry and Information Technology of the People's Republic of China, and Science Foundation of Shanghai. Her research interests include pattern recognition, image processing and machine learning.

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: Zhisheng Yan (zyan@gsu.edu), Yao Liu (yaoliu@binghamton.edu), Wenming Cao (wmcao@szu.edu.cn), and Phoenix Fang (dofang@calpoly.edu). 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: Jun Wu, Fudan University, China
Steering Committee Chair: Joel J. P. C. Rodrigues, Federal University of Piauí (UFPI), Brazil
Vice Chair – America: Shaoen Wu, Illinois State University, USA
Vice Chair – Asia: Liang Zhou, Nanjing University of Post and Telecommunications, China
Vice Chair – Europe: Abderrahim Benslimane, University of Avignon, France
Letters & Member Communications: Qing Yang, University of North Texas, USA
Secretary: Han Hu, Beijing Institute of Technology, China
Standard Liaison: Guosen Yue, Huawei, 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.