

A Survey of Multimedia Big Data

Zaijian Wang^{1,*}, Shiwen Mao², Lingyun Yang¹, Pingping Tang¹

¹The College of Physics and Electronic Information, Anhui Normal University, Wuhu 241000 China.

²Department of Electrical and Computer Engineering, Auburn University, Auburn, AL 36849-5201 USA.

* The corresponding author, email: wangzaijian@ustc.edu

Abstract: Multimedia big data brings tremendous challenges as well as opportunities for multimedia applications/services. In this paper, we present a survey and tutorial for multimedia big data. After discussing the characteristics of multimedia big data such as human-centricity, multimodality, heterogeneity, unprecedented volume, and so on, this paper provides an overview of the state-of-the-art of multimedia big data, reviews the latest related technologies, and discusses the technical challenges. We conclude this paper with a discussion of open problems and future directions.

Keywords: Multimedia big data; human-centricity; heterogeneity; machine learning; multimodality

I. INTRODUCTION

Along with the growing prevalence of Internet-friendly, low-cost commodity devices with cameras (such as mobile phones, surveillance systems, etc.) and proliferation of social networks (such as Wechat, QQ, Tudou, Facebook, YouTube, Twitter, etc.), it is very easy for the ordinary user to generate and share all kinds of multimedia content. The user may have little knowledge of networking, but can handily post and upload user-generated content (including text, voice, photo, and video) to social networks through the Internet [1].

The enormous multimedia content offered on online social networks includes immense quantities of multimedia data, which come from various sources ranging from surveillance video and to social media. Since massive multimedia datasets are being generated, stored, transmitted, and processed, multimedia data has become a unique type of big data in the Big Data era [2]. Considering both the richness and scale of multimedia big data, it could provide tremendous opportunities to mine, model, learn and analyze for multimedia applications/services including multimedia retrieval, recommendations, searches, 3-D rendering, multimedia games, learning, advertisements, etc. Different from general, text-based big data, multimedia big data also brings about obvious challenges since it is characterized by both massive volume and partially unknown complex structures, which cannot be effectively handled by the traditional multimedia data processing technologies [3].

Multimedia big data can essentially be viewed as large-scale signals composed of partially unknown complex structures, which involve motion trajectory, time statistics, spatial statistics, human factors and inter-view correlations with structured singularities [4]. Multimedia big data is often unstructured, multimodal and heterogeneous, because of which it is difficult to represent and mod-

Received: Feb. 9, 2017

Revised: Jun. 4, 2017

Editor: Zan Li

This article contributes to the ongoing cross-disciplinary efforts in multimedia big data research by investigating the background and state-of-the-art of multimedia big data-relevant data analytic technologies, such as Knowledge Graph, Modeling, Divide-and-Conquer Anchoring, network traffic monitoring tools, representation, big data knowledge engineering framework, security, and so forth.

el. Since a computer cannot understand high-level semantics, there is a higher level of complexity involved in comprehending and cognition of multimedia big data. Multimedia big data has Quality of Experience (QoE) requirements, which often evolve with space and time. Multimedia big data often entails considerably more resources in terms of acquisition, storage, transmission, presentation and processing, including, for instance, the need for graphics processing unit (GPU) processing and parallel, distributed software [3].

The typical multimedia big data lifecycle is illustrated in figure 1, which comprises acquisition, compressing, storage, processing, understanding, assessment, computing, and security. In the acquisition stage, multimedia big data is often generated by heterogeneous sources, including ubiquitous portable mobile devices (such as smartphone, digital camera, individual digital device, etc.), video lectures, digital games, Internet, Internet of Thing (IoT), multimedia sensor, social media, and virtual words. After acquiring multimedia big data, one must compress it for efficient storage and communications, since its volume is beyond the ability of general software/networking tools. In the storage and processing stages, one faces the challenge of unprecedented data volumes. Furthermore, it is often infeasible to store all captured multimedia data in many scenarios. Effective multimedia big data analysis methods are needed to achieve fast storage and processing. Since there is the semantic gap between semantics and video visual appearance, automated understanding is required for multimedia big data when user receives multimedia big data. To improve us-

er's satisfactory level for multimedia big data services, we need to assess the quality of multimedia big data service. With proliferation of sensitive data, multimedia big data also needs to be managed under sound security during the entire multimedia big data lifecycle. In the typical multimedia big data lifecycle, computing has a profound effect, where almost each stage of the multimedia big data lifecycle demands real-time and effective computing, since multimedia big data applications have large volumes and needs to be processed in a tolerable period of time.

In the next sections, we provide detailed discussions of each of the multimedia big data lifecycle stages. In particular, we present the state-of-the-art, as well as future research directions in the field of multimedia big data with a review of the existing literature. We provide an overview of this emerging field, survey contemporary approximation techniques such as multi-level 3-D shape feature extraction [7], and a spatial multimedia big data framework [22]. We also discuss the open problems and possible future directions in this area. We hope this paper can serve as a useful references for readers who are interested in working in the exciting area of multimedia big data, which is still at its infancy stage with tremendous research opportunities.

The remainder of this paper is organized as follows. In Section II, we provide some preliminaries of big data. In Section III, we provide a definition of multimedia big data and review its key lifecycle components, as well review related work in each component area. We review several multimedia big data applications in Section IV, and discuss important open problems and future directions in Section V. Finally, we summarize this paper in Section VI.

II. BIG DATA PRELIMINARIES

Recently, enormous datasets are described as "Big Data," which can offer potentially huge values to researchers and businesses [40], [41]. Big data research and applications are

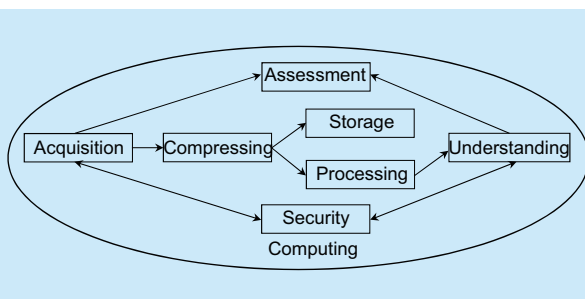


Fig. 1. The typical multimedia big data lifecycle.

accelerated by many government agencies and industry, while different definitions of big data are given. In big data, the data is often characterized by “4V” or “5V,” where “4V” indicates volume (quantity of data), velocity (fast generation of new data), veracity (quality of the data), and variety (data from different categories); and “5V” includes “4V” and value (huge value but very low density) [19]. In general, the datasets with growing data scale cannot be handled (i.e., perceived, captured, managed, and processed, etc.) by traditional database technologies within an acceptable scope for big data [40], [41].

In big data applications, it is the most critical problem to discover values from datasets with a rapid generation, enormous scale, and various types (including structured, semi-structured and unstructured data). Especially, people often cannot obtain the eventual size of the dataset since the data is often accumulated over time. Apart from masses of data, big data also has some other features as follows:

- the size of data is too big to be handled by the existing network tools;
- the data are streamed at rates faster than that can be handled by the existing network systems;
- the quality of data should be considered;
- the data has different types and modalities;
- the value of data should be considered, from heterogeneous, autonomous sources.

The readers can refer to [19], [40], [41] for more details.

III. MULTIMEDIA BIG DATA

3.1 Definition and features of multimedia big data

At present, although the importance of multimedia big data has been recognized, there is no clear definition of multimedia big data. Different from big data with 4V and 5V, multimedia big data is a type of datasets, in which the data is human-centric, heterogeneous, and has more media types and higher volume than the

typical big data. Multimedia big data analysis requires much more sophisticated algorithms and much more computing resources than the typical big data. For multimedia big data, an obvious feature is to effectively understand the media context and content, while another important feature is to design human-centric, high-level semantics. Some typical characteristics of multimedia big data are described as follows.

- Compared to non-multimedia big data, multimedia big data consists of more data types, which are more helpful for human understanding than for machines (for example, people often like video more than text, but it is more difficult for a machine to understand video than text). From media type to media content, multimedia big data focuses on humans, not machines.
- Datasets are mainly composed of different types of video data, such as camera video, interactive video, social video, immersive video (also called stereoscopic 3-D video), 3-D virtual worlds, and so forth. Therefore, multimedia big data has a higher level of complexity than typical big data, which is usually text-based.
- Multimedia big data can be acquired from heterogeneous sources, including ubiquitous portable mobile devices (such as the smartphone, digital camera, individual digital device, etc.), video lectures, digital games, the Internet, the Internet of Things (IoT), multimedia sensor, social media, and virtual worlds. Heterogeneous sources make multimedia big data to be unstructured, heterogeneous, and multimodal. Therefore, it is very challenging to represent and model multimedia big data since these data come from different sources or spaces (for example, social, cyber, and physical). Meanwhile, a sizable portion of multimedia big data is missing and multimedia big datasets are often incomplete due to possibly disparate origins.
- Since multimedia big data often evolve over time and space, multimedia big data analysis needs to effectively understand the

context and content in different scenarios.

- Multimedia big data have unprecedented data volumes, in which many applications tap into large, fast-moving, complex data streams and apply advanced analytical techniques. Furthermore, it is often infeasible to store all captured multimedia data in many cases. Different from traditional methods that can obtain only pieces of knowledge with uncertainty, incompleteness, and varying levels of quality from single sources [14], the large-scale multimedia big data inevitably contain communication errors, corrupted measurements, and even suffer from malicious attacks [28].
- Since multimedia data pass through a network at high velocity, multimedia big data requires to be processed rapidly and continuously, subject to storage volume and time constraints. For stored multimedia big data, real-time computing is often needed since the data is too big to move. Moreover, multimedia big data is typically dynamic, i.e., transient in nature, and also exhibits a wide variety since it often evolves with time and space.
- With the proliferation of sensitive video data, multimedia big data need to be managed under tight security. It is critical to consider traitor-tracing and secure multimedia sharing issues for multimedia big data.

From the above characteristics discussed, we can observe that multimedia big data involves working with sensitive data, analyzing complex and heterogeneous data, managing distributed data under security and performance constraints, and there are some specific requirements that emerge from multimedia big

data applications. There are obvious differences between typical big data and multimedia big data. Table 1 provides a comparison of traditional datasets, big data, and multimedia big data. We then discuss the components in the multimedia big data lifecycle (see figure 1) in the remainder of this section.

3.2 Acquisition

In the typical multimedia big data lifecycle, the first phase is to acquire multimedia big data. In this stage, multimedia big data can be acquired from heterogeneous sources, including ubiquitous portable mobile devices (such as a smartphone, digital camera, individual digital device, etc.), video lectures, digital games, the Internet, the Internet of Thing (IoT), multimedia sensor, social media, and virtual worlds. Heterogeneous sources generate multimedia big data that are unstructured, heterogeneous, and multimodal. Different multimedia types are generated, such as video (including camera video, interactive video, social video, and immersive video [17]), audio, speech, images, text, documents, HyperText Markup Language (HTML), geospatial data, tables, graphics, 2D graphs, 3-D virtual worlds, and other forms of multimedia information. Many video coding standards have been proposed. Therefore, multimedia big data have a much higher level of complexity than typical big data, which is usually text-based.

Multimedia big data tends to be unstructured, heterogeneous, and multimodal, since it comes from different sources or spaces (e.g., social, cyber, and physical spaces). Different media types with different formats have different distribution characteristics in volume,

Table I. Comparison of traditional datasets, big data, and multimedia big data.

	Data source	Data volume	Eventual data size	Interpreting video	Data structure	Interactive response	Real-time	Complexity	Human-centric	Response
Traditional dataset	Single, centralized	Small	Certain	No	Text-data	No	No	Low	No	No
Big data	Heterogeneous, distributed	Midium	Uncertain	No	Text-data	No	Yes	Medium	No	No
Multimedia big data	Heterogeneous, distributed	Big	Uncertain	Yes	Video-data	Yes	Yes	High	Yes	Yes

size, and so forth. Even if analyzed in the same space, they exhibit different characteristic patterns. Because users like unstructured data that is suitable for human understanding, unstructured data is rapidly growing in both quantity and quality. It is important to design new constructing and representing methods that can provide analytics of complex and heterogeneous data spaces.

Different from traditional methods that typically can obtain only pieces of knowledge with uncertainty, incompleteness, and at varying levels of quality from a single source [14], when designing new methods to efficiently acquire useful knowledge from heterogeneous multimedia information sources with complex and evolving relationships, one should keep in mind that large-scale multimedia big data inevitably contain communication errors, corrupted measurements, and even suffer from malicious attacks [28]. It is very difficult to represent and model multimedia big data. For example, unstructured multimedia big data is difficult to be understood by machines. Furthermore, most studies on traffic recognition focus on graph structure, rather than video structure. In particular, the hidden video content information is usually ignored in different scenarios.

Since the data collected from diverse domains are heterogeneous, the authors in [34] show that the data variables can be divided into different groups. For a particular problem, different forms can record the multiple views by considering each variable group as a particular view. The information acquired from each view may not be sufficient for effective learning, but the integration of multiple features from diverse information sources is both valuable and necessary. In [35], dynamic unstructured data is modeled as stream characterized by high volume, sequential access and dynamic evolution. To describe the properties and interests of a user, the paper utilizes a user profile in which users opinions about resources are expressed by descriptive labels (i.e., tags). Each descriptive label is regarded as a separate dimension in the vector space model.

To organize multimedia resources, a semantic link network model is proposed in [24], which generates the association relation among multimedia resources. For communication between geographically distributed cloud data centers, paper [52] focuses on multi-rate Bandwidth on-Demand (BoD) services which are offered by the Dense Wavelength Division Multiplexing (DWDM) layer. Paper [56] proposed a mobile content distribution scheme, which is based on roadside parking cloud (RPC) and mobile cloud (MC), where RPC is formed by the parked car on the roadside and MC is formed by moving cars on the road. This scheme is based infrastructure to relieve content distribution over VANETs with the high mobility of vehicles and intermittent connectivity.

The characteristics of different methods are summarized in Table 2. From the table, we find that different methods have different views. It is difficult to classify them into distinct categories. However, Table 2 shows that new methods need to consider in their design factors such as multimodal, unstructured, heterogeneous, geographically distributed data, dynamic evolution, interests of a user, temporal and spatial information, and semantics.

3.3 Compressing

After acquisition, multimedia big data must

Table II. Characteristics of the acquisition methods.

Method	Characteristics
BigKE [14]	Fragmented knowledge modeling from multiple autonomous information sources. Feature streams derived from temporal and spatial information
Statistical learning tools [28]	Consider communication errors, corrupted measurements, and even suffer from attacks
Multi-view Intact Space Learning algorithm [34]	For a particular problem, different forms can take in the multiple views by considering each variable group as a particular view
Tag assignments stream clustering [35]	Use locality-sensitive hashing to model dynamic unstructured data
Semantic link network model [24]	Organize multimedia data with their semantics
BoD broker [52]	Focuses on multi-rate Bandwidth on-Demand (BoD) services
Mobile content distribution scheme [56]	Based on infrastructure to relieve content distribution over VANETs

be compressed since it has an unprecedented volume, which brings great challenges to traditional data storage and processing systems. Under limited storage and computation resources, it is important to effectively compress the raw data with signal processing or transformation techniques.

There are many challenges to effectively compress multimedia big data, such as:

- multimedia big data is more complex than typical, text-based big data, in which structured, semi-structured, and unstructured data need to be handled;
- because multimedia big data has a large volume, effective multimedia big data compressing approaches are needed to achieve fast speed;
- Sometimes a sizable portion of multimedia big data is missing and multimedia big datasets are often incomplete due to possibly disparate origins.

The typical approaches include Compressive Sensing and Wavelet Transform for big data reduction [38]. The authors in [39] use a feature descriptor to achieve high compression ratio, which belongs to multimedia coding approaches. To design an effective high-level 3-D shape feature, the work [7] presents a multi-level 3-D shape feature extraction framework, which is a three-level feature extraction framework employing deep learning [8]-[13]. Firstly, the intrinsic properties are represented by low-level 3-D shape descriptors that are extracted. Then, the paper discovers middle-level patterns to explore geometric relationships among words by encoding low-level 3-D shape descriptors into

geometric bag-of-words. After that, a Deep Belief Network is applied to learn the structural high-level shape features. The authors show that the learned high-level features are highly discriminative [7].

To discover a latent representation of the data, paper [34] integrates the encoded complementary information in multiple views. The authors show theoretically that abundant information can be acquired by combing multiple views for latent intact space learning. To extract knowledge from Big Data, the authors in [62] propose an architecture to capture, process, analyze and visualize data coming from multiple streaming multimedia TV stations and radio stations. The paper emphasizes techniques such as diarization and the optimization of the number of Hadoop nodes, provisioned from Cloud infrastructures, to deliver enhanced performance.

The characteristics of different compressing methods are summarized in Table 3. From Table 3, it can be seen that the above mentioned methods are focused on high-level features and integrated the features in multiple views. It is vital to provide effective description for high-level features.

3.4 Storage

With massive multimedia big data being continuously generated, it needs to be effectively stored after compressing. Since multimedia big data has an unprecedented data volume and is in various media types, it is important to store the data effectively for various multimedia big data applications. With the growing quantity and quality of unstructured data, multimedia big data brings about big challenges to traditional data storage systems that are based on the NoSQL model. In many scenarios, it is often infeasible to store all observed multimedia data. For new data storage systems, practicability and cost are also important design factors.

In a recent work [36], the authors present a processing pipeline that is capable of seamlessly combining the NoSQL storage model and a Big Data processing platform,

Table III. Characteristics of the compressing methods.

Method	Characteristics
Typical approaches [39]	Compressive Sensing and Wavelet Transform for big data reduction
Feature descriptor [40]	Compact descriptors to achieve compression
Multi-level 3-D shape feature extraction framework [7]	Multi-level 3-D shape feature extraction
Multi-view Intact Space Learning algorithm [34]	Integrate the encoded complementary information in multiple views
Hadoop framework [62]	Extract knowledge from Big Data

e.g., MapReduce. To address data processing and network overload with rapid data volume [49], this paper proposed a hybrid-stream big data analytics model to perform big data video analysis, that contains data preprocessing, data classification, data recognition and data load reduction. Paper [57] introduces shallow-learning-based hashing and deep-learning-based hashing demonstrating state-of-the-art techniques for enabling efficient multimedia storage, indexing, and retrieval. Paper [58] focuses on content-centric networking (CCN) that is one of the key candidate Information-centric networking (ICN) architectures. In this paper, the effect of several caching strategies in the CCN domain is studied from the perspective of network and server overhead. The in-network caching performance of CCN under several popular cache replication methods are analyzed comprehensively.

In [60], a NoSQL approach is proposed to store big data effectively. NoSQL is designed for distributed data stores for very large scale data needs, can take advantage of scaling out - which means spreading the load over many commodity systems. Paper [61] presents a high-performance and scalable distributed storage and computing system for IMS services through different scenarios of service provisioning, storing and computing processes. Since a database containing a large number of 3-D models needs very high spatial cost of storing all the aggregated feature vectors, paper [73] presented a Sum of Sparse Binary codes (SSB) aggregation algorithm. The proposed algorithm encodes a local feature into a highly sparse binary code. By summing into a compact feature vector, a set of sparse binary codes are aggregated efficiently.

The characteristics of different methods are summarized in Table 4. From the table, it can be seen that new storage methods should have high-performance and scalable distributed storage. Meanwhile, the limitations in cloud computing and the IoT should be considered.

3.5 Processing

For multimedia big data, one of the most

Table IV. Characteristics of the storage methods.

Method	Characteristics
MARISSA [36]	Integrate NoSQL data stores with MapReduce under non-Java application scenarios
Hybrid-stream big data analytics model [49]	Key frame extraction
Multimedia Hashing [57]	Create effective compact hash codes
CCN [58]	Content-centric networking
NoSQL approach [60]	Distributed data stores for very large scale data
Distributed storage and computing system [61]	Combine the IP Multimedia Subsystem (IMS) platform and the Hadoop system for use in the distributed storage of IMS service resources
SSB [73]	Encodes a local feature into a highly sparse binary code

fundamental tasks is to explore the large volumes of data and extract useful information/knowledge for future actions. In this step, multimedia big data is often fused together to generate a coherent decision. Because multimedia big data applications/services are typically real-time, it is very important to process effectively large-scale high-dimensional data with a tolerable processing time. Several challenges posed by multimedia big data need be addressed, as follows.

- With continuously generated multimedia big data in real-time by many data sources, multimedia big data needs to be processed rapidly and continuously subject to storage and time constraints;
- multimedia big data is typically dynamic in which the data is transient in nature, and also have a wide variety since it often evolves with time and space;
- Automated and intelligent analytical techniques should be developed to extract knowledge from structured, semi-structured, and unstructured data by handling a very large number of multimedia data in real-time;
- multimedia big data should be processed by parallel/distributed, real-time online/streamed algorithms;
- since multimedia big data systems need large-scale computation, the allocation of storage, computation, and networking/communication resources should be optimized. To process large-scale high-dimensional

data, Divide-and-Conquer Anchoring is proposed to tackle nonnegative matrix factorization in high-dimensions in [33]. The authors in [33] show that Divide-and-Conquer Anchoring can locate the anchors precisely, even when the high dimensional features suffer from serious incompleteness. In this paper, the high-dimensional anchoring problem is divided into a few cheaper sub-problems to seek anchors of data projections in low-dimensional random spaces. Results indicate Divide-and-Conquer Anchoring performs well on various datasets.

To solve several challenges posed by crowdsourced multimedia data (such as social sensors, vehicle sensors, human sensors, physical sensors, etc.), the authors in [22] present a spatial multimedia big data framework after analyzing such multimodal and diversified crowdsourced multimedia data. According to [22], the big data framework can collect multimedia data from vehicles, social networks, and a very large crowd equipped with multi-sensory smartphones. Furthermore, paper [22] indicates that a very large number of multimedia spatio-temporal queries can be handled in real-time with the proposed framework. In [43], the authors design a novel community-centric framework for community activity prediction based on big data analysis. An effective approach is presented to extract community activity patterns by analyzing the big data collected from both the physical world and virtual social space. The proposed approach is evaluated with a real dataset collected over a 15-month period. In order to reduce the computational, storage, and communications

bottlenecks in big data, paper [29] proposes to exploit recent advances in convex optimization algorithms. Based on their investigations, the authors show how to design a convex optimization approach, suggest unconventional computational choices, and show how to make key structure-dependent algorithmic approximation tradeoffs.

In [66], an efficient object-aware anomaly detection scheme is proposed, that specifically focuses on certain object categories. In this paper, the normal events are represented by extracting histograms of block-motion trajectories and clustering them. To analyze a massive volume of big data and describe human behavior in the social area in real time address, a new concept of SmartBuddy is proposed in [75], which focuses on the analysis, the ecosystem provided by wearable devices, smart cities and big data to determine human behaviors as well as human dynamics. Since current various personalized recommendation systems ignore that the accelerated proliferation of social media data has led to the big data era, a cloud-assisted differentially private video recommendation system based on distributed online learning is proposed in [76]. The proposed framework modeled service vendors as distributed cooperative learners. The proposed model considers the sparsity and heterogeneity of big social media data.

The characteristics of different methods are listed in Table 5. The future methods should consider the sparsity and heterogeneity of multimedia big data, and handle spatio-temporal information in real-time. The computational, storage, and communications bottlenecks in big data should be reduced.

3.6 Understanding

With the rapid growth of video resources, there is a compelling need to resolve automated understanding of raw multimedia data based on visual appearances with a tolerable processing time. Due to the semantic gap between high-level semantics and low-level features, it is very difficult to understand multimedia big data with a machine. Especially,

Table V. Characteristics of the processing methods.

Method	Characteristics
Spatial multimedia big data framework [22]	Handle spatio-temporal queries in real-time
Divide-and-Conquer Anchoring [33]	Tackle nonnegative matrix factorization in high-dimensions
Object-aware anomaly detection scheme [66]	Confine the analysis to moving objects and avoid irrelevant background dynamics
SmartBuddy [75]	To analyze a massive volume of big data and describe human behavior in the social area
Cloud-assisted differentially private video recommendation system [76]	Consider the sparsity and heterogeneity of big social media data

some multimedia big data often evolve with time and space. It is thus important to develop cross-media and multi-modal analytical tools, and intelligent analytical techniques to understand the semantics of multimedia data, with the explosive growth of videos with no meta-data provided.

Knowledge Graph is proposed in [31], [32] to extract knowledge relations from structured, semi-structured, and unstructured text. A Knowledge Graph consists of a lot of concepts, entities, and relationships, in which various things in the real world are reflected by the concepts and entities. Knowledge Graph would help user cognize knowledge completely by creating a complete knowledge system. Before a Knowledge Graph is constructed, the entities and relationships can be extracted from open sources based on information extraction and pattern matching technologies. Paper [32] describes the following five steps to construct the Knowledge Graph:

- to crawl large-scale network corpus;
- to extract the entities and the relationships;
- to build the multidimensional graph model based on the corpus;
- to mine the multidimensional relationships based on the model; and
- quantify the uncertainty in the knowledge graph.

Knowledge Graph has been applied in many applications, including Semantic Search, information fusion of broad-spectrum knowledge, big data analytics, deep reading, among others. Among the current research on Knowledge Graph, automatic construction of Knowledge Graph is the main method for handling large-scale Knowledge Graph. However, more research is needed on Knowledge Graph for unstructured data, since many applications (such as FreeBased and Wikipedia) only focus on structured text and semi-structured text data.

For multimedia big data classification, a scalable classifiers ensemble framework in [48], that integrate the outputs from multiple classifier. In the proposed framework, features are extracted from the Region of Action

(ROA) in order to capture the action related information. The classification models chosen are Sparse Representation Classification (SRC) and Hamming Distance Classification (HDC). To mine rare class in multimedia big data which is imbalanced data caused by many real-world applications, paper [50] proposes a novel concept correlation analysis strategy framework through utilizing inter-concept correlations between the retrieval scores and labels. For summarizing User-generated Videos (UGVs) by considering both the representativeness and the quality of the selected segments from an original video, paper [55] considered and integrated semantics, emotions, and shooting quality since rich semantic and emotional contents are contained in UGVs.

In order to bridge the semantic gap between high-level concepts and low-level visual features, paper [59] proposes a novel system of discovering negative correlation for semantic concept mining and retrieval, that designed to adapt to Hadoop MapReduce framework further extended to utilize Spark. Since it is difficulty to directly tackle big data case due to their high computing time and memory usage, paper [63] proposes a comparative evaluation for the simplification of deep compositional features by exploring the existing vector quantization and binarization techniques. In this paper, a dedicated image searching framework is proposed to evaluate all the techniques in terms of computational cost, memory usage and discrimination preserving. To detect video content with various video transformations, paper [71] presents a video copy detection method based on sparse representation of MPEG-2 spatial and temporal features. In this paper, the global feature (HSVcolor histograms) and local feature (ORB features) are extracted from the key frames extracted based on visual saliency model.

The characteristics of different methods are listed in Table 6. The above mentioned methods should consider sparse representation of spatial and temporal features by extracting knowledge relations from structured, semi-structured and unstructured text.

Table VI. Characteristics of the understanding methods.

Method	Characteristics
Domain knowledge Graph Method [32]	Knowledge graph
Classifiers ensemble framework [48]	Features are extracted from the Region of Action (ROA) in order to capture the action related information
Novel concept correlation analysis strategy framework [50]	To mine rare class in multimedia big data which is imbalanced data
UGV summarization framework [55]	Consider and integrate semantics, emotions, and shooting quality
Novel ICF (Integrated Correlation Factor) algorithm [59]	Discovering negative correlation for semantic concept mining and retrieval
Dedicated image searching framework [63]	The simplification of deep compositional features by exploring the existing vector quantization and binarization techniques
Video copy detection method [71]	Sparse representation of MPEG-2 spatial and temporal features

3.7 Assessment

With the exponential growth of multimedia big data applications and services, it is very important for the service providers to cater for the Quality of Experience (QoE) of end users [72]. For video applications, it is crucial to evaluate video quality from the users' perspective. Because user experience is subjective, it is difficult to accurately evaluate the QoE of multimedia big data applications/services. By comparing applications user received with the corresponding original ones, evaluation scores can be obtained with subjective or objective tests. A subjective test is usually conducted in the laboratory environment, which often entails high cost in terms of time, money, and human power. By identifying the objective QoS parameters that contribute to user perceptual quality, an objective test can be developed based on the principle of the Human Visual System (HVS). Subjective test cannot be used for real-time evaluation in general. Its results are often used as ground truth to validate the performance of an objective test. Most of the objective test methods depend on subjective test results to train model parameters.

Challenges on assessment of multimedia big data include the following.

- It is hard to quantify and measure user experience levels.
- How to identify and weigh errors according to the temporal and spatial features of multimedia big data applications?
- How to monitor and control multimedia big data applications in real-time?
- How to develop QoE prediction models that do not depend on the original videos, network, and spatio-temporal features from distorted multimedia big data applications?
- How to define a model to evaluate and combine the contributions of each influence factor?
- Various factors should be considered in relevant scenarios.
- How to design a benchmark for customers?
- How to design an effective automated means to evaluate user experience?
- How to apply effective mechanism to correlate QoS/QoE metrics?
- How to assess the QoE quickly and accurately under various standards?

It is vital to monitor and analyze network traffic for improved user experience and optimizing network resource allocation [26]. For network traffic monitoring and analysis, existing methods usually rely on a high performance server with a large storage capacity, but this approach is not scalable for analyzing a large volume of traffic data in detail. In [27], the authors propose an open-source distributed computing platform based on Hadoop, which is an open-source distributed computing platform for big data processing [45]. Since Hadoop is a Java-based framework, users are required to write MapReduce code in Java, where MapReduce developed by Google is a major computing model for big data applications [46]. Since Hadoop has low cost and high efficiency, it is widely used to conduct large-scale data processing. As a suitable platform for large-scale network measurements, Hadoop has several valuable features, such as low-cost, scale-out capability, high fault tolerance, and efficient distributed parallel computing. However, Hadoop is originally developed for batch-oriented tasks (such as web page

indexing and text mining). It is a challenge to tailor Hadoop for network measurements [27].

The QoE problem has been studied in many recent works. For example, paper [77] discusses the relationship between user's QoE and the status of IPTV set-top box. Some statistical analyses are conducted. By utilizing a classical imbalanced dataset, this paper designs a cost coefficient for Adaboost model to achieve a higher accuracy. Paper [78] employs multiple machine learning (ML) algorithms (Decision Trees, Artificial Neural Networks, Gaussian Naive Bayes classifiers, and Support Vector Regression machines), and uses empirical measurements based on network metrics (e.g., delay, packet loss, and packet interarrival) and subjective opinion scores reported by actual users.

To assess the QoE of high-definition (HD) video streaming services, paper [79] proposed an objective QoE assessment and prediction method for HD video stream service that evaluated the user experience according to image damage accumulation (IDA). This paper established a mapping relationship between IDA and mean opinion scores. To provide good service and improve user's QoE, paper [80] studies the relationship between alarming data from IPTV set-top boxes and the user's QoE, uses decision tree video to model dataset, adopts information entropy minimization heuristic to discretize continuous-valued attributes. Paper [81] proposes a novel decision-theoretic approach called CaQoEM for QoE prediction. In CaQoEM, Bayesian networks and utility theory are used to predict users' QoE under uncertainty.

Based on random neural networks, paper [82] proposes a cross-layer prediction model for estimating the perceptual quality of mobile video in no reference mode. To improve the quality of IPTV service and the users satisfaction, paper [83] proposes a new attribute called user's viewing custom from the users point of view. For subjective video quality evaluation, paper [83] creates a mapping between viewing time ratio and users QoE. Paper [84] employs several machine learning (ML) algorithms for

Table VII. Characteristics of the assessment methods.

Method	Characteristics
Open-source distributed computing platform [28]	Hadoop is used to conduct large-scale data processing
Adaboost model [77]	Statistical analyses
Modular algorithm for user-centric QoE prediction [78]	Uses empirical measurements based on network metrics and subjective opinion scores reported by actual users
Objective QoE assessment and prediction method for HD video stream service [79]	Objective QoE assessment, establish a mapping relationship between image damage accumulation and mean opinion scores
Decision tree [80]	The relationship between alarming data from IPTV set-top boxes and the user's QoE
CaQoEM [81]	Decision-theoretic approach
Cross-layer prediction model [82]	Estimate the perceptual quality of mobile video in no reference mode
Improved CART algorithm [83]	Subjective video quality evaluation, creates a mapping between viewing time ratio and user's QoE
MLQoE [84]	For user-centric analysis and prediction of the QoE in video streaming

predicting the QoE of the video streaming service.

The characteristics of different methods are summarized in Table 7. From the table, we find that machine learning (ML) algorithms are often adopted and more human user-related metrics are used.

3.8 Computing

It is known that a medium-quality response obtained quickly can achieve a better QoE than a high-quality response obtained slowly [28]. Multimedia applications/services are typically real-time. While multimedia big data is continuously generated in real-time by many data sources, multimedia big data requires to be processed rapidly and continuously subject to storage and time constraints. Moreover, multimedia big data is typically highly dynamic, i.e., the data is transient in nature, and also has a wide variety since it often evolves with time and space. Therefore, it is necessary to design parallel/distributed, real-time online/streamed processing for analysis, mining, and learning algorithms [15]. Considering the large-scale computation involved, multimedia big data systems need to optimize storage,

computation, and networking/communication resources. In addition, GPU computing brings about further challenges since multimedia data pass through the network at high velocity.

Paper [16] presents a cloud-based system to harvest the information in the global camera networks in an efficient manner. In the proposed system, the multimedia big data from many cameras can be simultaneously analyzed by an application programming interface. In the global camera network (from instructional video in MOOC (massive open online course) to commercial movies, from surveillance videos to home videos taken at birthday parties), multimedia big data have wide variety and also pass through network at high velocity. Since storing multimedia data requires large capacity, the cloud-computing technology is applied for computing and storage [16].

To solve several challenges posed by crowdsourced multimedia data (such as social sensors, vehicle sensors, human sensors, physical sensors, etc.), paper [22] presents a spatial multimedia big data framework after analyzing such multimodal and diversified crowdsourced multimedia data. According to [22], the big data framework can collect multimedia data from vehicles, social networks, and a very large crowd equipped with multi-sensory smartphones. Furthermore, paper [22] shows that a very large number of multimedia spatio-temporal queries can be handled in real-time with the proposed framework [22].

In order to mitigate the computational, storage, and communications bottlenecks in big data, paper [29] exploits the recent advances in convex optimization algorithms. Based on the investigation, the authors show that the big data problem is to design a convex optimization algorithm and suggest several unconventional computational choices, and how to make key structure-dependent algorithmic approximation tradeoffs.

To synchronize multimedia data (such as audio and video with the 3-D depth skeletal data of a patient performing therapy at home), the authors in [53] propose a therapy recorder that can perform the two-tier synchronization

process and create the synchronized multimedia therapy session file, and a therapy player that can unpack the complex session file and separate the media files while keeping the synchronization among the media.

Paper [54] presents an open-source platform enabling developers to create and deploy Real-Time multimedia Communications (RTC) applications with advanced media processing capabilities, which introduces the concept of Media Pipeline and follows a Platform as a Service (PaaS) scheme. In paper [64], a big data analysis platform is presented for facilitating the public cultural services, which based on standard Cloud Computing platform and Hadoop ecosystem. To effectively and efficiently retrieve desired images from a large image database, paper [65] proposes three types of content-based image retrieval and a conceptual feature to improve the quality of content-based image retrieval. The content-based image retrieval determines the most relevant images according to the content comparisons between a query image and searched images.

Paper [67] presents a precision recommendation method based on big data technology for massive public digital cultural resources, which help people find the resources they are really interested in. In this paper, the public digital cultural resources are classified. To analyze big traffic data for intelligent transportation with different designs of vehicles in the same class and the similarity of shape and textures between different classes, paper [68] proposed a novel method using local and structural features for vehicle classification in real-time traffic system. This paper obtained local features by using scale invariant feature transform (SIFT), used an efficient L2-norm sparse coding technique to reduce computational cost, extracted vehicle building structures as structural features, and utilized linear support vector machine (SVM) as the classifier.

To process efficiently such a large volume of non-linear and non-stationary data, paper [70] presents a new parallel signal data anal-

ysis algorithm that utilizes General-Purpose Graphics Programming Unit (GPGPU) to improve data scalability and reduce computation time for large non-linear and non-stationary datasets. To fuse the big data technology into the process of aircraft health management, a design of a big data analysis and application platform is proposed in paper [74] for civil aircraft health management. Based on the data available to us now and foreseeable future, the proposed platform is designed in order to meet the current demands.

The characteristics of different methods are listed in Table 8. From Table 8, the future methods should fuse the big data technology, improve data scalability, reduce computation time and utilize Cloud Computing. Local and structural features should be applied into classifying traffics in order to process multimedia big data timely.

3.9 Security

Because most multimedia big data owners must rely on professionals to analyze such data, the potential safety risks arise since some datasets include private video or sensitive video, such as nude photo/video [42]. With the proliferation of sensitive data, multimedia big data should be managed under sound security. It is critical and urgent to consider traitor tracing and secure multimedia sharing issues for multimedia big data. To securely manage access to multimedia big data, we need to study the composition and enforcement of privacy policies. Furthermore, when we design distributed data management under security, we should consider normal users with very little or no domain knowledge or expertise, and design a new method to manage privileges of multimedia big data by a simple drag and drop mechanism [23].

At present, the existing works in the security domain of big data include digital watermarking, authorization, context-aware policies, authentication, and MPEG encryption schemes. Paper [23] includes some most relevant works in multimedia big data security. For sharing of multimedia big data in a secure

Table VIII. Characteristics of the computing methods.

Method	Characteristics
Cloud-based system [16]	Harvest the information in the global camera networks
Spatial multimedia big data framework [22]	Handle a very large number of multimedia spatio-temporal queries
Divide-and-Conquer Anchoring [33]	Tackle nonnegative matrix factorization in high-dimensions
Therapy recorder [53]	To synchronize multimedia data
PaaS [54]	An open-source platform
Big data analysis platform [64]	Based on standard Cloud Computing platform and Hadoop ecosystem
Content-based image retrieval [65]	Three types of content-based image retrieval and a conceptual feature
Precision recommendation method [67]	The public digital cultural resources are classified
A novel method using local and structural features [68]	Use local and structural features for vehicle classification in real-time traffic system
New parallel signal data analysis algorithm [70]	To utilize General-Purpose Graphics Programming Unit (GPGPU) to improve data scalability and reduce computation time for large non-linear and non-stationary datasets
Big data analysis and application platform [74]	To fuse the big data technology into the process of aircraft health management

and private manner, a framework is proposed in [23], where privacy requirements are captured in an access control system. The authors in [14] propose a big data knowledge engineering framework (BigKE), in which some functions are executed including fragmented knowledge modeling, online learning from multiple information sources, automated demand-driven knowledge navigation, and non-linear fusion on fragmented knowledge. The paper thinks that the fragmented knowledge is part of the migration puzzle, where each piece does not provide the whole picture, but some limited information. Through utilizing the observed phenomenon where user-shared images show users' similarities, the authors in [18] present a multimedia big data system, which can be as an alternative to user-generated tags and social graphs for gender identification and follower/followee recommendation.

The authors in [19] apply big data technologies in Virtual Physiological Human (VPH) to produce robust and effective medicine solutions. The paper thinks that big data technologies have great potential in the domain of

computational biomedicine. In [20], an integrated approach is proposed, in which a lot of technologies are combined including Social Network Analysis, Business Intelligence, Internet of Things, Big data, GeoSpatial information processing, structure and unstructured content analysis with Semantic techniques, and multimedia resources. By combining associative in-memory technologies, cloud computing, collaborative processes, and semantic technologies, this paper tries to merge and analyze information from different sources.

Considering that home based therapy monitoring has come to a reality, the authors in [21] present a gesture controlled e-therapy online framework, which can monitor occupational and physical therapy exercise using multimedia data acquired by different sensors (such as Leap, Myo, and Kinect2). The multimedia data is stored in a big data repository. For sharing of multimedia big data in a secure and

private manner, a framework is proposed in [23], where privacy requirements are captured in an access control system. In the proposed system, user can compose conflict free policies for their online multimedia data. For secure multimedia sharing and traitor tracing issues, the authors in [25] present a scheme, which implements the Tree-Structured Harr (TSH) transform in a homomorphic encrypted domain for fingerprinting using social network analysis in order to protect media distribution in social network.

The authors in [51] proposes an analysis model for multimedia data encryption optimization, introduces a general-purpose lightweight speed tunable video encryption scheme, and proposes a series of intelligent selective encryption control models. In [69], a new scheme based on the Z-curve is proposed for protecting the privacy and integrity of spatial data stored in the Cloud while being able to execute range queries efficiently. To support answering range query over encrypted data set, the proposed scheme suggests a new index structure. For answering range queries over spatial data stored on the Cloud, a distributed algorithm is described in this paper.

The characteristics of different methods are listed in Table 9. From the table, the new methods should consider multimedia big data is high-volume, real-time, dynamic and heterogeneous.

IV. MULTIMEDIA BIG DATA APPLICATIONS

In addition to the technical challenges, multimedia big data also enables many news applications. We briefly review some such applications reported in the literature in this section.

Through utilizing the observed phenomenon where user-shared images show users' similarities, paper [18] presents a multimedia big data system, which can serve as an alternative to user-generated tags and social graphs for gender identification and follower/followee recommendation. According to [18], it is the first attempt to prove such a phenomenon for

Table IX. Characteristics of the security methods.

Method	Characteristics
BigKE [14]	Fragmented knowledge modeling, online learning from multiple information sources, automated demand-driven knowledge navigation, and nonlinear fusion on fragmented knowledge
Multimedia big data system [18]	Bag-of-features tagging (BoFT) is used to analyze images which are annotated with labels
Big Data for Personalized Healthcare [19]	Apply big data technologies in Virtual Physiological Human (VPH)
CHIS-BI [20]	To focus on behavior analysis performed by monitoring, combining and comparing information from different types of populations and visits
Multimedia Big Data E-Therapy Framework [21]	To monitor occupational and physical therapy exercise using multimedia data acquired by different sensors
SUDRS [23]	Consider normal users with very little or no domain knowledge or expertise, and design a new method to manage privileges of multimedia big data by simple drag
A scheme which implements the Tree-Structured Harr (TSH) transform [25]	for fingerprinting using social network analysis in order to protect media distribution in social network Implement the Tree-Structured Harr (TSH) transform in a homomorphic encrypted domain
Multi-level Intelligent Selective Encryption Control Model [51]	Propose an analysis model for multimedia data encryption optimization
New scheme based on the Z-curve [69]	A new scheme based on the Z-curve

user-shared images along with more prediction methods. In this paper, bag-of-features tagging (BoFT) is used to analyze images that are annotated with labels.

Paper [19] applies big data technologies in Virtual Physiological Human (VPH) to produce robust and effective medicine solutions. This application incorporates some specific requirements as follows: sensitive data, complex and heterogeneous data spaces (including non-textual information), distributed data management under security and performance constraints, and so forth. In [19], a guiding example is provided to illustrate the big data VPH, while the big data problem is considered in the light of the “5V” definition. The paper shows that big data technologies have a great potential in the domain of computational biomedicine [42]. Meanwhile, five major problems are described in the paper, which have intense on-going research activities.

Paper [14] proposes a big data knowledge engineering framework (BigKE), in which some functions are executed including fragmented knowledge modeling, online learning from multiple information sources, automated demand-driven knowledge navigation, and nonlinear fusion on fragmented knowledge. The paper shows that the fragmented knowledge is part of the migration puzzle, where each piece does not provide the entire picture, but some limited information. The framework is composed of three tiers. The first tier is fragmented knowledge modeling and online learning from multiple information sources. The second tier is nonlinear fusion of fragmented knowledge, which fuses this fragmented knowledge obtained from multiple local sources to acquire a complete set of knowledge. Since this knowledge is likely to be biased and inaccurate, which is resulted by the limited scope of its analysis, the united representation and foundation for the construction of knowledge ontology are provided in BigKE’s nonlinear fusion process. The third tier is automated demand-driven knowledge navigation, which takes the unified representation and the fused knowledge ontology, and

carries out demand-driven knowledge navigation for personalized services.

In [20], an integrated approach is proposed, in which several technologies are combined including Social Network Analysis, Business Intelligence, IoT, Big data, GeoSpatial information processing, structure and unstructured content analysis with Semantic techniques, and multimedia resources. By combining associative in-memory technologies, cloud computing, collaborative processes, and semantic technologies, this paper aims to merge and analyze information from different sources. However, the proposed system is focused on behavior analysis performed by monitoring, combining, and comparing information from different types of populations and visits.

Considering that home based therapy monitoring has become a reality, paper [21] presents a gesture controlled e-therapy online framework, which can monitor occupational and physical therapy exercises using multimedia data acquired by different sensors (such as Leap, Myo, and Kinect2). The multimedia data is stored in a big data repository. From the platform some therapeutic information can be mined, such as the most appropriate therapy regime, ethnicity, geo-spatial location and disability level.

V. OPEN PROBLEMS AND OUTLOOK

At present, much literature in the field has discussed multimedia big data-related technologies and proposed some methods and techniques. However, multimedia big data-related research and applications are still in the infancy stage, and only limited work has been done in the field. Multimedia big data-related technologies are lacking on a systemic and further research. There are many obstacles in the development of multimedia big data applications. Some potential future research directions are provided as follows.

a) Nonlinear Knowledge Fusion: Generally, the knowledge of multimedia big data is fragmented since this data is often obtained from heterogeneous, autonomous sources,

which utilize different technologies to acquire, record, represent, and process data. To mine potential and useful information from fragmented data with lower accuracy, nonlinear knowledge fusion plays a vital role in the multimedia big data era. Research is needed so that a new knowledge fusion method can be designed. This method would solve the problems related to heterogeneous multimedia data types (such as camera video, interactive video, social video, immersive video, etc.), intrinsic semantic associations in video data, relationship networks among data, the relationship between video and the user's background, and so forth. The new method should provide useful knowledge by linking data through complex relationships.

b) Demand-driven Multimedia Big Data Mining: Because the knowledge from multimedia data often evolves with time, space, context, circumstance, the observed objects and the information extracted are often transient. In particular, a user's requirements often evolve with time context, circumstances, scenarios, content, and the individual's profile. Therefore, the new algorithm should be capable of predicting future trends and should analyze and model dynamic knowledge together with dynamic requirements simultaneously.

c) Dynamic Video Summarization Methods: Extracting information from dynamic videos containing rich information is confronted with a huge challenge. The new method, mentioned previously, should also deliver good performance. As part of the new method, a video should be effectively presented by utilizing a series of video frames. In addition, some new smoothing measures should be implemented to make the final summarization look natural. It is an especially yet important challenging problem to achieve automatic comprehension of raw multimedia solely based on their visual appearance. Intelligent methods to represent a vast amount of videos with no metadata should be developed for some typical applications (which include video surveillance, video resources browsing and indexing system, IoT, criminal investigation systems, intrusion de-

tection systems, sport events detection, and many others).

d) Real-time, Stream, Distribution, Parallel, Online Analytic Methods: With the explosion of multimedia applications (such as surveillance video, medical images, entertainment and social media, voice and video), their data volume grows to an extent that cannot be handled effectively by traditional multimedia processing and analysis systems [6]. The new method should effectively manage multimedia big data, which becomes hard to store and even harder for multiple scans. Many data or performance measures are non-smooth, non-linear, non-convex and non-decomposable over samples. In many multimedia big data applications, it is impossible to obtain the eventual data scale since the multimedia data is accumulated over time. Furthermore, multimedia big data usually exists in a distributed manner where different owners may hold different parts of the multimedia big data. In this case, different sources have different degrees of importance for different analytical goals. Data at different places may bring significant noise since they differ in quality.

Moreover, for the same multimedia data, different users might have different requirements. The new method needs to consider the computational and storage loads of multimedia big data. In an efficient data-processing platform, streamed, distributed, and parallel techniques should be adopted to handle multimedia big data. An effective and efficient multimedia data processing approach is critical for real-time multimedia big data applications where multimedia data is streamed in nature. The complexity and scalability problems encountered in the development of multimedia big data applications should be a focus of future research.

e) Human-centric Analysis Methods: Although multimedia analysis technologies develop rapidly, there is still a lack of understanding of user experience. User experience is an important component in multimedia big data research since users with different intentions usually have different criterion for the

achievement of satisfactory results. We need to design new methods to effectively express and represent images or video, and make it easier to understand the user experience. Meanwhile, the new methods should be capable of discovering users' latent intention from limited observed multimedia big data, since multimedia big datasets are often incomplete, due to possibly disparate sources. In new human-centric analysis methods, user profiles, behaviors, and social networks should be considered to model the user's intention [44].

f) Security: Since multimedia big data has an unprecedented volume and is usually stored in a distributional manner, security is actually a long-standing problem that still remains open. For multimedia big data, we should consider that different analyzers should have different access rights, which should be warranted by different data owners. Meanwhile, we should also leverage the sources without access to the entire data. The new method should avoid the violation of privacy concerns. Due to its large scale and high diversity, a multimedia big data safety mechanism should be designed, which includes effective cryptography approaches, access control, safety management, and safety communication. In addition, some characteristics should be considered, including attack characteristics, loophole characteristics, virus characteristics, and so forth.

g) Network Design: With more and more network equipment being deployed, it is necessary to make efforts to alleviate the peak pressure of network infrastructure and improve the utilization ratio in order to cope with the continuous growth of user demand for rich media experience. Moreover, existing infrastructures adopt the relatively static mechanism of system resource provisioning, which cannot react fast enough for various emerging multimedia applications. To design multimedia big data networks with sound performance, we need to consider application characteristics, users' demand for an enhanced QoE and rich service, a fundamental tradeoff between the system cost and user requirements, as well as utilizing various enabling technologies to low-

er the system maintenance cost. Furthermore, multimedia big data network design should meet heterogeneity, reliability, scalability, security, and usability needs, and, furthermore, be operated with acceptable costs. Re-programmable technologies may offer advanced solutions to reduce the cost of deploying and operating multimedia big data networks [17].

VI. CONCLUSIONS

The emergence of multimedia big data not only brings about challenges but also opens great opportunities for the development of multimedia big data applications. In the multimedia big data era, user requirements will drive the progress of future multimedia technologies. Multimedia big data will not only bring about impacts on human living and thinking, but also affect social and economic development.

As Alan Turing said, "we can only see a short distance ahead, but we can see plenty there that needs to be done [47]." This article contributes to the ongoing cross-disciplinary efforts in multimedia big data research by investigating the background and state-of-the-art of multimedia big data-relevant data analytic technologies, such as Knowledge Graph, Modeling, Divide-and-Conquer Anchoring, network traffic monitoring tools, representation, big data knowledge engineering framework, security, and so forth. It offers fundamental insights into the various open problems and latent directions for the development of multimedia big data technologies.

ACKNOWLEDGEMENTS

This work was supported in part by the National Natural Science Foundation of China (NO. 61401004, 61271233, 60972038), Plan of introduction and cultivation of university leading talents in Anhui (No.gxfxZD2016013), the Natural Science Foundation of the Higher Education Institutions of Anhui Province, China (No. KJ2010B357), Startup Project of Anhui Normal University Doctor Scientific

Research (No.2016XJJ129), the US National Science Foundation under grants CNS-1702957 and ACI-1642133, and the Wireless Engineering Research and Education Center at Auburn University.

References

- [1] P. Zhou, Y. Zhou, D. Wu, and H. Jin, "Differentially private online learning for cloud-based video recommendation with multimedia big data in social networks," *IEEE Transactions on Multimedia*, vol. 18, no. 6, pp. 1217-1229, Mar. 2016.
- [2] Y. Lu, X. Wang, W. Zhang, H. Chen, L. Peng, and W. Zhao, "Performance analysis of multimedia retrieval workloads running on multicores," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 11, pp. 3323-3337, Nov. 2017.
- [3] W. Zhu, P. Cui, Z. Wang, and G. Hua, "Multimedia big data computing," *IEEE MultiMedia*, vol. 22, no. 3, pp. 96-106, July/Sept. 2015.
- [4] H. Xiong, "Structure-based learning in sampling, representation and analysis for multimedia big data," in *Proc. 2015 IEEE International Conference on Multimedia Big Data (BigMM)*, Beijing, China, Apr. 2015, pp. 24-27.
- [5] M. Mardani, G. Mateos, and G.B. Giannakis, "Subspace learning and imputation for streaming big data matrices and tensors," *IEEE Transactions on Signal Processing*, vol. 63, no. 10, pp. 2663-2677, Mar. 2015.
- [6] S.-C. Chen, R. Jain, Y. Tian, and H. Wang, "Special issue on multimedia: the biggest big data," *IEEE Transactions on Multimedia*, vol. 17, no. 1, pp. 1401-1403, Sept. 2015.
- [7] S. Bu, Z. Liu, J. Han, J. Wu, and R. Ji, "Learning high-level feature by deep belief networks for 3-D model retrieval and recognition," *IEEE Transactions on Multimedia*, vol. 16, no. 8, pp. 2154-2167, Dec. 2014.
- [8] 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.
- [9] 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.
- [10] 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.
- [11] X. Wang, L. Gao, and S. Mao, "PhaseFi: Phase fingerprinting for indoor localization with a deep learning approach," in *Proc. IEEE GLOBECOM 2015*, San Diego, CA, Dec. 2015, pp. 1-6.
- [12] X. Wang, L. Gao, S. Mao, and S. Pandey, "DeepFi: Deep learning for indoor fingerprinting using channel state information," in *Proc. IEEE WCNC 2015*, New Orleans, LA, Mar. 2015, pp. 1666-1671.
- [13] X. Wang, C. Yang, and S. Mao, "ResBeat: Resilient breathing beats monitoring with online bimodal CSI data," in *Proc. IEEE GLOBECOM 2017*, Singapore, Dec. 2017, pp. 1-6.
- [14] X. Wu, H. Chen, G. Wu, J. Liu, Q. Zheng, X. He, A. Zhou, Z. Zhao, B. Wei, M. Gao, Y. Li, Q. Zhang, S. Zhang, R. Lu, and N. Zheng, "Knowledge engineering with big data," *IEEE Intelligent Systems*, vol. 30, no. 5, pp. 46-55, Sept./Oct. 2015.
- [15] Y. Xue, P. Zhou, T. Jiangs, S. Mao, and X. Huang, "Distributed learning for multi-channel selection in wireless network monitoring," in *Proc. IEEE SECON 2016*, London, UK, June 2016, pp. 1-9.
- [16] W.-T. Su, Y.-H. Lu, and A.S. Kaseb, "Harvest the information from multimedia big data in global camera networks," in *Proc. 2015 IEEE International Conference on Multimedia Big Data (BigMM)*, Beijing, China, Apr. 2015, pp. 184-191.
- [17] Y. Wen and W. Zhu, "Fuelling big data intelligence into future multimedia system: Reflection and outlook," in *Proc. 2015 IEEE International Conference on Multimedia Big Data (BigMM)*, Beijing, China, Apr. 2015, pp. 1-4.
- [18] M. Cheung, J. She, and Z. Jie, "Connection discovery using big data of user-shared images in social media," *IEEE Transactions on Multimedia*, vol. 17, no. 9, pp. 1417-1428, Sept. 2015.
- [19] M. Viceconti, P. Hunter, and R. Hose, "Big data, big knowledge: Big data for personalized healthcare," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 4, pp. 1209-1215, July 2015.
- [20] A. Chianese, P. Benedusi, F. Marulli, and F. Piccialli, "An associative engines based approach supporting collaborative analytics in the Internet of Cultural Things," in *Proc. 2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC)*, Krakow, Poland, Nov. 2015, pp. 533-538.
- [21] A.M. Qamar, S.O. Hussain, B. Sadiq, A.R. Khan, M.A. Rahman, and S. Basalamah, "A multimedia big data e-therapy framework," in *Proc. 2015 IEEE International Conference on Multimedia Big Data (BigMM)*, Beijing, China, Apr. 2015, pp. 288-289.
- [22] B. Sadiq, F. Ur Rehman, A. Ahmad, "A spatio-temporal multimedia big data framework for a large crowd," in *Proc. 2015 IEEE International Conference on Big Data*, Santa Clara, CA, Oct. 2015, pp. 2742-2751.
- [23] A. Samuel, M.I. Sarfraz, H. Haseeb, S. Basalamah, and A. Ghafoor, "A framework for composition and enforcement of privacy-aware and context-driven authorization mechanism for multimedia big data," *IEEE Transactions on Multimedia*, vol. 17, no. 9, pp. 1484-1494, Sept.

- 2015.
- [24] C. Hu, Z. Xu, Y. Liu, L. Mei, L. Chen, and X. Luo, "Semantic link network-based model for organizing multimedia big data," *IEEE Transactions on Emerging Topics in Computing*, vol. 2, no. 3, pp. 376-387, Sept. 2014.
- [25] C. Ye, Z. Xiong, Y. Ding, J. Li, G. Wang, X. Zhang, and K. Zhang, "Secure multimedia big data sharing in social networks using fingerprinting and encryption in the JPEG2000 compressed domain," in *Proc. 2014 IEEE 13th International Conference on Trust, Security and Privacy in Computing and Communications*, Beijing, China, Sept. 2014, pp. 616-621.
- [26] Z. Wang, Y. Dong, S. Mao, and X. Wang, "Internet multimedia traffic classification from QoS perspective using semi-supervised dictionary learning models," *IEEE/CIC China Communications*, to appear.
- [27] J. Liu, F. Liu, and N. Ansari, "Monitoring and analyzing big traffic data of a large-scale cellular network with Hadoop," *IEEE Network*, vol. 28, no. 4, pp. 32-39, July/Aug. 2014.
- [28] K. Slavakis, G.B. Giannakis, and G. Matesos, "Modeling and optimization for big data analytics: (Statistical) learning tools for our era of data deluge," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 18-31, Sept. 2014.
- [29] V. Cevher, S. Becker, and M. Schmidt, "Convex optimization for big data: Scalable, randomized, and parallel algorithms for big data analytics," *IEEE Signal Processing Magazine*, vol. 31, no. 5, pp. 32-43, Sept. 2014.
- [30] X. Wu, X. Zhu, G.-Q. Wu, and W. Ding, "Data mining with big data," *IEEE Trans. Knowledge and Data Eng.*, vol. 26, no. 1, pp. 97-107, Jan. 2014.
- [31] M. Schuhmacher and S.P. Ponzetto, "Knowledge-based graph document modeling," in *Proc. 7th ACM Int'l Conf. Web Search and Data Mining (WSDM'14)*, New York, NY, Feb. 2014, pp. 543-552.
- [32] Q. Lv, L. Xu, J. Yu, etc. "Research on domain knowledge graph based on the large scale online knowledge fragment," in *Proc. 2014 IEEE Workshop on Advanced Research and Technology in Industry Applications (WARTIA'14)*, Ottawa, ON, Canada, Sept. 2014, pp. 312-315.
- [33] T. Zhou, W. Bian, and D. Tao, "Divide-and-conquer anchoring for near-separable nonnegative matrix factorization and completion in high dimensions," in *Proc. 13th IEEE Int'l Conf. Data Mining (ICDM'13)*, Dallas, TX, Dec. 2013, pp. 917-926.
- [34] C. Xu, D. Tao, and C. Xu, "Multi-view intact space learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 12, pp. 2531-2544, Dec. 2015.
- [35] Z. Wu and M. Zou, "An incremental community detection method for social tagging systems using locality-sensitive hashing," *Neural Networks*, vol. 58, no. 1, pp. 12-28, Oct. 2014.
- [36] E. Dede, B. Sendir, P. Kuzlu, J. Weachock, M. Govindaraju, and L. Ramakrishnan, "Processing cassandra datasets with Hadoop-Streaming based approaches," *IEEE Transactions on Services Computing*, vol. 9, no. 1, pp. 46-58, June 2016.
- [37] Y. Wen, X. Zhu, J.J.P.C. Rodrigues, and C.-W. Chen, "Cloud mobile media: Reflections and outlook," *IEEE Transactions on Multimedia*, vol. 16, no. 4, pp. 885-902, June 2014.
- [38] M.K. Jeong, J.-C. Lu, X. Huo, B. Vidakovic, and D. Chen, "Wavelet-based data reduction techniques for process fault detection," *Technometrics*, vol. 48, no. 1, pp. 26-40, Feb. 2006.
- [39] L.-Y. Duan, J. Lin, J. Chen, T. Huang, and W. Gao, "Compact descriptors for visual search," *IEEE Multimedia*, vol. 21, no. 3, pp. 30-40, Jan. 2014.
- [40] M. Chen, S. Mao, and Y. Liu, "Big data: A survey," *Springer Mobile Networks and Applications Journal (MONET)*, vol. 19, no. 2, pp. 171-209, Apr. 2014.
- [41] M. Chen, S. Mao, Y. Zhang, and V.C.M. Leung, *Big Data: Related Technologies, Challenges and Future Prospects*, New York, NY: Springer, 2014.
- [42] M. Chen, Y. Qian, J. Chen, K. Hwang, S. Mao, and L. Hu, "Privacy protection and intrusion avoidance for cloudlet-based healthcare big data sharing," *IEEE Transactions on Cloud Computing*, vol. PP, no. 99, pp. 1-1, Oct. 2016. DOI: 10.1109/TCC.2016.2617382.
- [43] Y. Zhang, M. Chen, S. Mao, L. Hu, and V.C.M. Leung, "CAP: Crowd activity prediction based on big data analysis," *IEEE Network*, Special Issue on Networking for Big Data, vol. 28, no. 4, pp. 52-57, July/Aug. 2014.
- [44] M. Chen, Y. Hao, S. Mao, D. Wu, and Y. Zhou, "User intent-oriented video QoE with emotion detection networking," in *Proc. IEEE GLOBECOM 2016*, Washington DC, Dec. 2016, pp. 1-6.
- [45] Apache Hadoop, "Welcome to Apache Hadoop!" [online] Available: \url{http://hadoop.apache.org}. Retrieved 2016-08-25.
- [46] J. Dean and S. Ghemawat, "MapReduce: Simplified data processing on large clusters," in *Proc. Sixth USENIX Symposium on Operating System Design and Implementation (OSDI'14)*, San Francisco, CA, Dec. 2004, pp. 1-10.
- [47] A.M. Turing, "Computing machinery and intelligence," *Mind - A Quarterly Review of Psychology and Philosophy*, vol. 59, no. 236, pp. 433-460, Oct. 1950.
- [48] Y. Yan, Q. Zhu, M.-L. Shyu, and S.-C. Chen, "A classifier ensemble framework for multimedia big data classification," *Proc. 2016 IEEE 17th International Conference on Information Reuse and Integration (IRI)*, Pittsburgh, PA, July 2016, pp. 615-622.
- [49] K. Wang, J. Mi, C. Xu, L. Shu, and D.-J. Deng,

- "Real-time big data analytics for multimedia transmission and storage," *Proc. 2016 IEEE/CIC International Conference on Communications in China (ICCC)*, Chengdu, China, July 2016, pp. 1-6.
- [50] Y. Yan and M.-L. Shyu, "Enhancing rare class mining in multimedia big data by concept correlation," *Proc. 2016 IEEE International Symposium on Multimedia (ISM)*, San Jose, CA, Dec. 2016, pp. 281-286.
- [51] C. Xiao, L. Wang, Z. Jie, and T. Chen, "A multi-level intelligent selective encryption control model for multimedia big data security in sensing system with resource constraints," *Proc. 2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing (CSCloud)*, Beijing, China, June 2016, pp. 148-153.
- [52] A. Yassine, A.A.N. Shirehjini, and S. Shirmohammadi, "Bandwidth on-demand for multimedia big data transfer across geo-distributed cloud data centers," *IEEE Transactions on Cloud Computing*, vol. PP, no. 99, pp. 1-1, 2016.
- [53] M.A. Rahman and A. Alelaiwi, "A synchronized multimedia in-home therapy framework in big data environment," *Proc. 2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, Seattle, WA, July 2016, pp. 1-6.
- [54] B. Garcia, M. Gallego, L. Lopez, G.A. Carella, and A. Cheambe, "NUBOMEDIA: An elastic PaaS enabling the convergence of real-time and big data multimedia," *Proc. 2016 IEEE International Conference on Smart Cloud (SmartCloud)*, New York, NY, Nov. 2016, pp. 45-56.
- [55] B. Xu, X. Wang, and Y.-G. Jiang, "Fast summarization of user-generated videos: Exploiting semantic, emotional, and quality clues," *IEEE MultiMedia*, vol. 23, no. 3, pp. 23-33, July/Sept. 2016.
- [56] H. Gong, L. Yu, N. Liu, and X. Zhang, "Mobile content distribution with vehicular cloud in urban VANETs," *IEEE/CIC China Communications*, vol. 13, no. 8, pp. 84-96, Sept. 2016.
- [57] W. Liu and T. Zhang, "Multimedia hashing and networking," *IEEE MultiMedia*, vol. 23, no. 3, pp. 75-79, July/Sept. 2016.
- [58] D. Kim, Y.-B. Ko, and S.-H. Lim, "Comprehensive analysis of caching performance under probabilistic traffic patterns for content centric networking," *IEEE/CIC China Communications*, vol. 13, no. 3, pp. 127-136, Mar. 2016.
- [59] Y. Yan, M.-L. Shyu, and Q. Zhu, "Negative correlation discovery for big multimedia data semantic concept mining and retrieval," *2016 IEEE Tenth International Conference on Semantic Computing (ICSC)*, Laguna Hills, CA, Feb. 2016, pp. 55-62.
- [60] A. Gueidi, H. Gharsellaoui, and S. Ben Ahmed, "A NoSQL-based approach for real-time managing of embedded data bases," *Proc. 2016 World Symposium on Computer Applications & Research (WSCAR)*, Masaken Al Mohandesin, Egyptpp, Mar. 2016, pp. 110-115.
- [61] Y. Seraoui, M. Bellafkih, and B. Raouyane, "A high-performance and scalable distributed storage and computing system for IMS services," *Proc. 2016 2nd International Conference on Cloud Computing Technologies and Applications (CloudTech)*, Marrakesh, Morocco, May 2016, pp. 335-342.
- [62] J. Herrera and G. Molto, "Detecting events in streaming multimedia with big data techniques," *Proc. 2016 24th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP)*, Heraklion Crete, Greece, Feb. 2016, pp. 345-349.
- [63] S. Qiu, S. Wei, and Y. Zhao, "A comparative evaluation: Different methods for simplifying the deep compositional features," *Proc. 2016 IEEE International Conference on Multimedia and Expo (ICME)*, Seattle, WA, July 2016, pp. 1-6.
- [64] G. Zhang, Y. Yang, X. Zhai, W. Huang, and J. Wang, "Public cultural big data analysis platform," *Proc. 2016 IEEE Second International Conference on Multimedia Big Data (BigMM)*, Taipei, Taiwan, Apr. 2016, pp. 398-403.
- [65] J.-H. Su, T.-P. Hong, Y.-T. Chang, and H.-Y. Tung, "Un-supervised, semi-supervised and supervised image retrieval based on conceptual features," *Proc. 2016 IEEE Second International Conference on Multimedia Big Data (BigMM)*, Taipei, Taiwan, Apr. 2016, pp. 129-132.
- [66] X. Zang, G. Li, Z. Li, N. Li, W. Wang, "An object-aware anomaly detection and localization in surveillance videos," *Proc. 2016 IEEE Second International Conference on Multimedia Big Data (BigMM)*, Taipei, Taiwan, Apr. 2016, pp. 113-116.
- [67] X. Zhai, F. Jin, J. Wang, Y. Yang, Q. Yao, Q. Qiu, and G. Zhang, "A kind of precision recommendation method for massive public digital cultural resources: A preliminary report," *Proc. 2016 IEEE Second International Conference on Multimedia Big Data (BigMM)*, Taipei, Taiwan, Apr. 2016, pp. 56-59.
- [68] Z.Q. Xiang, X.L. Huang, and Y.X. Zou, "An effective and robust multi-view vehicle classification method based on local and structural features," *2016 IEEE Second International Conference on Multimedia Big Data (BigMM)*, Taipei, Taiwan, Apr. 2016, pp. 68-73.
- [69] M.N. Ba-Hutair and I. Kamel, "A new scheme for protecting the privacy and integrity of spatial data on the cloud," *2016 IEEE Second International Conference on Multimedia Big Data (BigMM)*, Taipei, Taiwan, Apr. 2016, pp. 394-397.
- [70] C.-C. Chen, C.-Y. Shen, M.-S. Chen, "Massive parallelism for non-linear and non-stationary data analysis with GPGPU," *Proc. 2016 IEEE International Conference on Big Data (Big Data)*, Washington DC, Dec. 2016, pp. 329-334.

- [71] D. Ren, L. Zhuo, H. Long, P. Qu, and J. Zhang, "MPEG-2 video copy detection method based on sparse representation of spatial and temporal features," *Proc. 2016 IEEE Second International Conference on Multimedia Big Data (BigMM)*, Taipei, Taiwan, Apr. 2016, pp. 233-236.
- [72] Z. He, S. Mao, and T. Jiang, "A survey of QoE driven video streaming over cognitive radio networks," *IEEE Network*, vol. 29, no. 6, pp. 20-25, Nov./Dec. 2015.
- [73] T. Furuya and R. Ohbuchi, "Aggregating sparse binarized local features by summing for efficient 3D model retrieval," *Proc. 2016 IEEE Second International Conference on Multimedia Big Data (BigMM)*, Taipei, Taiwan, Apr. 2016, pp. 314-321.
- [74] J. Chen, Z. Lyu, Y. Liu, J. Huang, G. Zhang, J. Wang, and X. Chen, "A big data analysis and application platform for civil aircraft health management," *Proc. 2016 IEEE Second International Conference on Multimedia Big Data (BigMM)*, Taipei, Taiwan, Apr. 2016, pp. 404-409.
- [75] A. Paul, A. Ahmad, M.M. Rathore, and S. Jabbar, "Smartbuddy: Defining human behaviors using big data analytics in social internet of things," *IEEE Wireless Communications*, vol. 23, no. 5, pp. 68-74, Oct. 2016.
- [76] P. Zhou, Y. Zhou, D. Wu, and H. Jin, "Differentially private online learning for cloud-based video recommendation With multimedia big data in social networks," *IEEE Transactions on Multimedia*, vol. 18, no. 6, pp. 1217-1229, June 2016.
- [77] R. Huang, X. Wei, C. Lv, X. Li, and S. Zhang, "Prediction model for user's QoE in imbalanced dataset," *Proc. 2015 First International Conference on Computational Intelligence Theory, Systems and Applications (CCITSA)*, Yilan, Taiwan, Dec. 2015, pp. 41-45.
- [78] P. Charonyktakis, M. Plakia, I. Tsamardinos, and M. Papadopouli, "On user-centric modular QoE prediction for VoIP based on machine-learning algorithms," *IEEE Transactions on Mobile Computing*, vol. 15, no. 6, pp. 1443-1456, June 2016.
- [79] Y. Geng, L. Meng, Y. Wang, Y. Yang, and Z. Qu, "QoE assessment and prediction method for high-definition video stream using image damage accumulation," *IEEE/CIC China Communications*, vol. 13, no. 7, pp. 48-59, Sept. 2016.
- [80] S. Sun, X. Wei, L. Wang, and J. Chen, "Association analysis and prediction for IPTV service data and user's QoE," *Proc. 2015 International Conference on Wireless Communications & Signal Processing (WCSP)*, Nanjing, China, Oct. 2015, pp. 1-5.
- [81] K. Mitra, A. Zaslavsky, and C. Ahlund, "Context-aware QoE modelling, measurement, and prediction in mobile computing systems," *IEEE Transactions on Mobile Computing*, vol. 14, no. 5, pp. 920-936, Dec. 2015.
- [82] E. Danish, M. Alreshoodi, A. Fernando, B. Alzahrani, and S. Alharthi, "Cross-layer QoE prediction for mobile video based on random neural networks," *Proc. 2016 IEEE International Conference on Consumer Electronics (ICCE)*, Antalya, Turkey, Nov. 2016, pp. 227-228.
- [83] H. Meng, R. Huang, X. Wei, Y. Qian, and Q. Liu, "QoE prediction model for IPTV based on machine learning," *Proc. 2016 8th International Conference on Wireless Communications & Signal Processing (WCSP)*, Yangzhou, China, Oct. 2016, pp. 1-5.
- [84] M. Plakia, M. Katsarakis, P. Charonyktakis, M. Papadopouli, and I. Markopoulos, "On user-centric analysis and prediction of QoE for video streaming using empirical measurements," *Proc. 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX)*, Lisbon, Portugal, June 2016, pp. 1-6.

Biographies



Zaijian Wang, received his BE degree (2002) from Anhui Polytechnic University, MSc degree (2005) from University of Science and Technology of China, and PhD degree (2015) from Nanjing University of Posts and Telecommunications.

His current research interests focus on multimedia big data, end-to-end QoS provisioning and wired/wireless multimedia streaming. He is currently an associate professor at College of Physics and Electronic Information, An'hui Normal University, Wuhu, China.



Shiwen Mao, (S'99-M'04-SM'09) received his Ph.D. in electrical and computer engineering from Polytechnic University, Brooklyn, NY (now Tandon School of Engineering of New York University, New York, NY) in 2004. Currently,

he is the Samuel Ginn Distinguished Professor and Director of Wireless Engineering Research and Education Center at Auburn University, Auburn, AL, USA. His research interests include wireless networks, multimedia communications, and smart grid. He is a Distinguished Lecturer of IEEE Vehicular Technology Society. He is on the Steering Committee of IEEE Transactions on Multimedia and IEEE Transactions on Network Science and Engineering. He is on the Editorial Board of IEEE Transactions on Multimedia, IEEE Internet of Things Journal, IEEE Multimedia, IEEE/CIC China Communications, among others. He serves as TPC Chair of IEEE INFOCOM 2018, Area TPC Chair of IEEE INFOCOM 2017 and 2016, Technical Program Vice Chair for Information Systems (EDAS) of IEEE INFOCOM 2015, symposium co-chairs for many conferences, including IEEE ICC, IEEE GLOBECOM, ICCCN,

among others, Steering Committee Voting Member for IEEE ICME and AdhocNets, and in various roles in the organizing committees of many conferences. He is the Chair of IEEE ComSoc Multimedia Communications Technical Committee. He received the 2015 IEEE ComSoc TC-CSR Distinguished Service Award, the 2013 IEEE ComSoc MMTC Outstanding Leadership Award, and the NSF CAREER Award in 2010. He is a co-recipient of the Best Demo Award from IEEE SECON 2017, the Best Paper Awards from IEEE GLOBECOM 2016, IEEE GLOBECOM 2015, IEEE WCNC 2015, and IEEE ICC 2013, and the 2004 IEEE Communications Society Leonard G. Abraham Prize in the Field of Communications Systems.



Lingyun Yang, received her B.S. degree from HeNan Normal University, Xinxiang, China, in 2004 and M.S. degree in communication and information system from Nanjing University of Science and Technology, Nanjing, China in 2006. She is currently working toward the Ph.D. degree at Nanjing University of Posts and Telecommunications, Nanjing, China. She is also currently an teacher with the College of Physics and Electronic Information, Anhui Normal University, Wuhu, China. Her current research interests focus on end-to-end QoE provisioning and wired/wireless multimedia streaming



Pingping Tang, was born in Anhui, China in 1981. She received her M.S. degree in computer science from Southeast University in 2006 and is currently studying for her Ph.D. degree in communications engineering from Nanjing University of Posts and Communications. From 2002 to 2006, she was a webpage designer working in Anhui Normal University and was primarily responsible for creating and maintaining official websites of the university, with the tools of Dreamweave, Photoshop, Firework, JSP, ASP, Java builder, etc. Since 2007, she has been a university lecturer at Anhui Normal University, teaching several courses in the Electronics and Communications Engineering Department, including C language, Single Chip Computer, Computer Technology, Principle of Microcomputer, Embedded System, etc. Her previous research was mainly concerned with computer science, while later her research interest became communication engineering, and now she is a Ph.D. candidate at Nanjing University of Posts and Communications, engaging in the business of communications.