3DQoE-Oriented and Energy-Efficient 2D plus Depth based 3D Video Streaming over Centrally Controlled Networks

Yanwei Liu, *Member, IEEE*, Jinxia Liu, Antonios Argyriou, *Senior Member, IEEE*, Song Ci, *Senior Member, IEEE*

***Abstract*—IP networks have become the dominant platform for video delivery. However, bandwidth-hungry video is pushing networks to their limits: costs are rising for the operators and the viewing experience is not always satisfactory for the users. When considering 3D video delivery, the previous problems are exacerbated because of the higher volume of data that must be communicated, and the difficulty in characterizing the viewing experience of the end user. Consequently, network operators may be reluctant to deliver 3D video due to costs and unclear quality improvements to their users. In this setting the true immersive experience of 3D video remains elusive.**

In this work we focus on the efficient delivery of 3D video in terms of quality and energy cost over centrally controlled networks. As a representative example of a centrally controlled network, a software defined network (SDN) is assumed. Our approach is based on a comprehensive network-dependent 3D quality of experience (3DQoE) model and an energy cost model for 3D video streaming. By using the developed models, we formulate the problem of energy-efficient and 3DQoE-optimized 3D video flow path routing. The particular characteristic of video/depth rate allocation presented in 3D video is embedded seamlessly into the selection of the optimal routing paths for multiple 3D video streams. The formulated problem is NP-hard and is solved with a heuristic algorithm based on the branch and bound method after significant reduction of the solution search space. Extensive 3D video streaming experiments are conducted over an OpenFlow-based SDN with subjective and objective evaluations and they highlight the significant benefits of the proposed approach.

***Index Terms*—3D video streaming, centrally controlled networks, SDN, OpenFlow, 3D quality of experience, energy efficiency, rate allocation**

I. Introduction

THE advances in stereoscopic video capture, compression, and rendering have already resulted in the commercialization of 3D video in 3D theaters. More recently, 3D video has started entering the living room by enabling 3DTV. 3DTV is a natural extension to the 2DTV since it can provide a depth-enhanced visual experience similar to the one that humans perceive the real world. With the rapid advances in auto-stereoscopic displays (glass-free 3D), 3D is expected to be one of the most popular options for video viewing in the future [1].

Delivering 3D video to the home can be accomplished in several ways that include broadcasting, Internet, Cable TV, and Blue-ray disks. Among these techniques, Internet-based 3D video on demand (VoD) is a choice that allows maximum flexibility both for the users and the content owners. It is also well known that Internet has currently become the most popular and successful mechanism for delivering 2D video. However, the path towards Internet-based 3D VoD services involves several challenges. Improvements in communication networks have led to capacity increases, and video compression technologies have led to the need for lower bandwidth [2]. However, huge costs are involved for the network operators and content providers that must deliver these bandwidth-hungry video streams through their infrastructure. Even 2D video is putting a significant strain on networks today [3]. Consequently, 3D video delivery over IP networks [1][4] is expected to further exacerbate the problems in the video distribution. As a result it is critical for the operators to optimize both the quality of the delivered 3D video and also their costs.

To optimize the delivery of 3D video, one must know precisely the control knobs in the overall system architecture. By delving into the details of 3D video we note that unlike 2D video it involves at least two views. This means that the volume of the transmitted video data is typically higher. Even for 3D video that is represented as 2D video plus depth (2D plus depth), the additional depth map data [2] increase significantly the total volume of the 3D video data. Thus, it is critical to allocate judiciously the available bandwidth across different 3D video flows. A second observation is related to the time-dependent fluctuations of the available bandwidth. Similar to 2D video streaming, it is more efficient to deploy scalable video coding (SVC) to encode 2D plus depth based 3D video so as to generate multiple scalable streams [5][6] that can match the available bandwidth of the communication channel. A third observation regarding the delivery of 2D plus depth based 3D video is that the rate allocation between video and depth [7] is important for the final 3D visual experience. Even more importantly for scalable encoded video the rate allocation among different layers of video and depth will result in different 3D visual experiences.

Our previous discussion was related to the transmitted 3D video stream, i.e., the content. The network that will

This work was supported in part by NSFC under Grants 61771469 and 61472388, Zhejiang Provincial Natural Science Foundation of China under Grant LY17F010001. (*Corresponding author: Jinxia Liu.*)

Y. Liu is with State Key Lab of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China, and also with the School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail:liuyanwei@iie.ac.cn).

J. Liu is with the Zhejiang Wanli University, Ningbo 315100, China (e-mail: liujinxia1969@126.com).

A. Argyriou is with the Department of Electrical and Computer Engineering, University of Thessaly, Volos 38221, Greece (e-mail: anargyr@ieee.org).

S. Ci is with University of Nebraska-Lincoln, Omaha 68046, USA (email: sci@engr.unl.edu).

deliver these videos has also to be carefully considered. Today the operators of content delivery networks (CDN) typically transmit the data over different routing paths [8] in their network to maximize different efficiency metrics. This can be further optimized by a centrally controlled network [9] which enables the traffic flows to match the network resources via a centralized monitoring mechanism. One of the typical centralized network architectures is the software defined network (SDN) paradigm that allows the control of network flow routing through a centralized mechanism. OpenFlow [10] is a protocol for software defined networks (SDNs) that allows this management/control of multiple flows to maximize the network efficiency. For the particular type of traffic that we consider in this paper (scalable 3D video), the multiple scalable streams can be transmitted over different routing paths to ensure the optimization of a desired metric. Since both the quality of each network path and the importance of the several source streams are different, optimizing routing decisions in a SDN for 3D video streams can have a significant impact on the 3D visual experience and operator cost.

In this paper our goal is to optimize the 3D quality of experience (3DQoE) for the users and the cost of the operator that are naturally conflicting objectives. As we discussed in the previous paragraphs, the 3D video streaming system can be controlled by tuning rate allocation at the source and changing path routing inside the network. Before we are able to optimize the system, there is a need to model the impact of rate allocation and path routing on the 3DQoE and operator cost. Consequently, we need a 3DQoE model suitable for network resource allocation. Besides 3DQoE, the operator cost [11] is also another element that we consider. The operator cost for a deployed network is typically synonymous to energy consumption [11]. The reason is that the video flows are communicated with network devices, such as the routers and gateways that typically consume significant energy. Hence, minimizing the total energy consumption is also critical for the operator. In summary, we note that several degrees of freedom exist both at the disposal of the content creator/owner and that of the operator. To achieve the optimization goals it is critical to characterize their impact on the user perceived 3DQoE as well as the energy cost of the operator.

In this paper, we propose a 3DQoE-oriented and energy-efficient scalable 3D video streaming framework over centrally controlled networks that can provide fine-grained mechanisms for monitoring and controlling individual network flows. Specifically, we utilize SDN as an instance of a centrally controlled network to design the optimization framework for 3D video streaming. We emphasize that while the techniques outlined in this work are applicable to any centrally controlled network architecture, the proposed optimization framework fits particularly well into SDNs. The specific contributions of this paper are as follows.

1) A network-dependent 3DQoE model based on Choquet integral is proposed to predict the 3DQoE of different 3D streaming scenarios. After mimicking the non-additive fusion process of the 3D perception of the cyclopean view, depth map and content characteristics in the human brain, the proposed 3DQoE model is transformed into an additive Choquet integral form. By utilizing fuzzy measurements, the 3DQoE can be predicted in terms of the left view video quality, depth quality and the 3D content type. Our model connects the different qualities of left view video and depth to the communication network.
2) We propose a 3DQoE-oriented and energy-efficient optimization framework for scalable 3D video streaming over an OpenFlow-based SDN. First, we formulate the problem of joint 3D video rate allocation and path selection. Then we show that the problem is NP-hard and we present a greedy branch and bound algorithm for solving it. With the proposed algorithm, the multiple streams of 3D video are dynamically distributed over different routing paths to avoid network congestion. The proposed framework and the associated solution algorithm take into account: first the optimal routing for multiple 3D video streams in the available routing paths, and second the application layer rate-distortion information (video/depth rate allocation) of the 3D video.
3) Our optimization framework is enhanced by embedding the energy consumption cost during the delivery (routing) of the 3D video flow. By modeling the energy cost of each routing path, the optimal trade-off between 3DQoE and energy cost can be identified. For each routing path, we establish an energy cost model that characterizes the piecewise linear relationship between the energy costs of network devices and the volume of video traffic.

To the best of our knowledge, our work is the first that jointly optimizes 3DQoE and energy consumption cost for 3D video streaming in a centrally controlled network. In addition, we believe that this is the first work that proposes the use of a Choquet integral to characterize 3DQoE by considering the interactions among the different factors that contribute to the 3DQoE.

The rest of the paper is organized as follows. Related work is reviewed in Section II, followed by the system overview in Section III. Section IV presents the 3DQoE model based on Choquet integral. Next, the proposed 3DQoE-oriented and energy-efficient 3D video flow routing optimization framework is described in Section V. Section VI provides the experimental results and finally Section VII concludes this paper.

II. Related Work

To capture accurately the perceived quality of the provided service to a user, the concept of quality of experience (QoE) instead of conventional quality of service (QoS) has been proposed [12]. Thus, designing and optimizing a multimedia networking system with a QoE objective is necessary in practical applications [13]. In the recent literature, significant progress has been made to QoE-aware multimedia communication [13], traffic management [14], congestion control [15] and video quality adaptation [16]. These works are based on a 2D video QoE model, and are not directly applicable to 3D video due to the increased depth perception in 3DQoE. Since the third dimension is included into the 3D video,

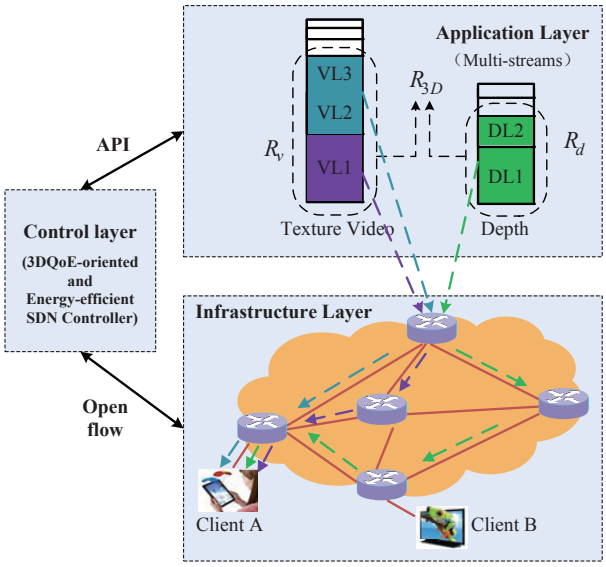


Fig. 1. OpenFlow-assisted 3D video streaming system (VL: video layer, DL: depth layer, R_v: rate of video, R_d: rate of depth, R_{3D}: the total 3D video rate).

QoE evaluation is more important for 3D video applications. Consequently, 3DQoE has been investigated in the literature [17]-[26]. 3DQoE is a complex multi-dimensional user-centric concept that characterizes the overall influence in the service processing chain on the total experience. Currently, it is very difficult to fully and accurately characterize 3DQoE in an entirely automated way [18]. However, for specific applications, one can model the service quality by only considering the dominant factors that affect it [27]. Even though 3DQoE can be modeled by considering a set of primary contributing factors, the interplay of the multiple factors that affect the overall 3DQoE is generally not fully understood.

Regarding the networks, an important evolution that took place in the last few years is the SDN [28], that allows network components to become accessible and controllable from the application layer. Hence, as an instance of centrally controlled networks, SDN provides a novel mechanism to optimize the transmission path by using a global view. Specifically, the OpenFlow protocol in a SDN has the ability to add new functionalities without modifying any network device. Consequently, the OpenFlow protocol [10], as a set of specifications for SDN, has been used in many applications [29]-[32]. To fully exploit the advantages of OpenFlow in flow controlling and managing, SDN was first used to optimize traffic engineering [29] and traffic policing [30]. For a 3D tele-immersion application [31], SDN was adopted to manage the streams with a cross-layer methodology. An important work in the area of SDN by Egilmez et al. [32] proposed a QoS optimization framework for adaptive video streaming. Furthermore, SDN was also used to assist network service migration to guarantee QoS requirements [33][34]. These works primarily aim at improvements in terms of classic QoS, instead of optimizing the QoE of user.

Regarding the aspect of energy consumption, energy saving technologies with intelligent device sleeping methods have been studied for flow routing [35]-[37]. However, these methods only consider minimizing the total energy consumption during traffic routing, and they neglect the influence of energy-efficient routing on the traffic quality. Moreover, there are no works that focus on the 3D video flow rate allocation with routing path selection over SDN to reach an optimal trade-off between the 3DQoE and the energy cost.

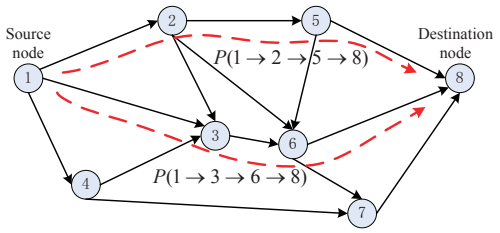


Fig. 2. Node tree graph.

III. System Overview

The OpenFlow-based 3D VoD streaming system is illustrated in Fig. 1. At the application layer, the 3D source data that consist of the video and depth are independently encoded by SVC to allow adaptation of the source coding rate. The multiple scalable video and depth streams have different importance levels and consequently different rate allocation combinations should be used between them. The controller calculates the optimal transmission path that maximizes the optimization metric which is the ratio of 3DQoE to the energy cost. The flow forwarding table is maintained by each switch at the infrastructure layer and it records the results of the optimal transmission path selection. The specific flows are dynamically switched over from the network elements by respecting the flow forwarding table. Finally, the different source flows converge at each target receiver/client. Because communication paths can be time-varying, the 3D video flow routing optimization is also dynamic. We assume that only one controller is centrally responsible for managing all the SDN network forwarding elements. When multiple controllers are used in large scale SDN applications [29][38], the routing computations can be performed by multiple controllers that need to reach a consensus.

At the infrastructure layer, the network is modeled as a set of V_N nodes which are interconnected by a set of links ES_N. The network is a directed graph $G(V_N, ES_N)$. In the graph, the link between two adjacent nodes is called the edge segment. We assume that a total of N edge segments are included in the graph. As an example, Fig. 2 shows a node tree graph with six intermediate nodes and two end nodes managed by the SDN controller. In this graph, Node 1 and Node 8 are the source node and the destination node, respectively. From the source node to the destination node there are many paths and each path consists of many links and intermediate nodes. $P(V_1 \rightarrow V_2 \rightarrow \cdots \rightarrow V_{dn})$ denotes one arbitrary transmission path from the source node to the destination node, where $V_1, V_2, ..., V_{dn}$ denote the nodes in the path, and $V_{dn-1} \rightarrow V_{dn}$ denotes the edge ES_{dn-1} in the path which connects to V_{dn-1} and V_{dn}. For example, $P(1 \rightarrow 2 \rightarrow 5 \rightarrow 8)$ and $P(1 \rightarrow 3 \rightarrow 6 \rightarrow 8)$ are both different paths that can be utilized to deliver 3D video streams.

In the network, each path is characterized by its channel bandwidth and quality. To adapt the source data rate to the

time-varying channel, the available channel bandwidth should be communicated to the SDN controller periodically. The SDN controller can interact with the network infrastructure layer and obtain through probing the available bandwidth, bit error rate (BER), and delay of different paths in real-time [39]. Furthermore, each path may consist of many connected edge segments, and each segment is characterized by a specific BER, bandwidth and transmission delay. Thus, the segment with the lowest bandwidth among all segments in the path is the bottleneck of the entire path. Regarding the delay of the path, in the case of video transmission, it can be converted to packet loss events after considering the maximum allowable playout delay.

To model the quality of the routing paths in the SDN we proceed as follows. For video transmission, the packet loss rate (PLR) is usually a typical indicator of the path quality. Packet loss is caused by link errors or violations of the playout deadline. For packet-based video transmission, the bit errors on the link usually lead to the application-layer packet loss. To capture the PLR introduced by link errors, let us denote a path that includes M (with $M < N$) segments as $(ES_1, ES_2, ..., ES_n, ..., ES_M)$. In the assumption of independent bit errors [40] [41], the PLR introduced by link errors over the edge segment ES_n can be estimated as $\rho_{e,ES_n} = 1-(1-b)^S$ by using the physical layer BER b and packet length S. However in real networks, the bit errors usually occur in bursts. Thus for a S-bit packet, because $S \cdot b$ is far less than 1, ρ_{e,ES_n} can be approximated as [40]

$$\rho_{e,ES_n} \approx S \cdot b. \tag{1}$$

Thus, the packet loss probability introduced by link errors for one packet over the entire path is computed as [41]

$$\rho_e = 1 - \prod_{n=1}^{M} (1 - \rho_{e,ES_n}). \tag{2}$$

Besides the PLR introduced by link errors, the delay-induced PLR can be estimated according to the transmitted data rate and the available bandwidth. By assuming that an M/M/1 queue [42] models the average queuing delay T_d at a node, it can then be expressed as

$$T_d = \frac{S}{R_l - R_p}, \tag{3}$$

where R_p denotes the allocated source coding bit-rate for path p and R_l denotes the lowest available link capacity among all segments in the path. The probability that the queuing delay of a packet exceeds the playout deadline $T_{\max}$, can be written as [43]

$$\rho_d = e^{-(R_l - R_p)\cdot T_{\max}/S}. \tag{4}$$

Hence, together with the PLR introduced by link errors, the total PLR ρ_p for the transmission path p is computed as

$$\rho_p = \rho_e + (1-\rho_e)\cdot\rho_d \tag{5}$$

IV. 3DQoE Model

Since flow routing directly affects the user-perceived visual experience, there is a need to build a network-dependent 3DQoE estimation model to guide the flow routing optimization. This model has to include several factors that affect the 3D visual experience, such as the content, source/video coding algorithm, network statistics, device characteristics, and of course the human factor [17]. In this paper we are concerned with the development of a 3DQoE model that includes factors related to the centrally managed network and the source content. Therefore, we focus our attention on modeling the impact of the source content characteristics and the network on 3DQoE.

Recent research has shown that 3DQoE is related to the natural characteristics of video content [17][44]. Thus, we consider the content type into the proposed 3DQoE model. Generally, the camera motion and content motion activity level is used to characterize the categories of content type. In our model we define a discrete content description parameter ct to indicate the level of motion. The adopted motion levels include lower motion (slight and smooth object movement in temporal domain), medium motion (medium object movement in temporal domain and slight camera movement in spatial viewpoint domain), and higher motion (fast object movement in temporal domain and gentle camera movement in spatial viewpoint domain). To obtain the content type for a given 3D video clip, the mean absolute difference (MAD) between successive frames is computed. In particular we compare the average MAD $\bar{d}$ over all frames in the clip with two predefined empirical threshold values (lower values d_l and upper values d_u) to determine the content type of the clip. The content type is at the low motion level on condition of $\bar{d} < d_l$ and is at the high motion level on condition of $\bar{d} > d_u$. Otherwise, the content type is at medium motion level.

Regarding the network, we consider the impact of bandwidth and packet loss on the 3DQoE. It is true that during the video transmission, besides the available bandwidth and packet losses, the playback jitter and the transmission delay both contribute significantly on the total 3DQoE. However, problems related to jitter are alleviated with the help of a playback buffer that is typically used at the receiver side. Regarding the second problem, we adopt the maximal playback deadline limitation for the video transmission which means that if a packet has not arrived by the prescribed deadline, it is discarded for playback whether it is received later or not. Thus, the impact of delay on 3DQoE is converted to the equivalent of packet loss events. For 2D plus depth based 3D video streaming, the available bandwidth and packet loss directly influence the quality of the transmitted video and depth, that eventually affect the total 3DQoE. Hence, the impact of the network on the 3DQoE can be indirectly transformed into those of video and depth qualities on 3DQoE.

We must also note that the overall 3DQoE [45] depends also on the 3D visual comfort. The improper disparity and the unnaturalness of the synthesized virtual view both lead to eye fatigue [46]. The 3D visual fatigue can be attributed to the quality of the 3D video data because the visual discomfort is mostly introduced by the improper binocular view fusion in the human brain. Hence, for 2D plus depth based 3D video data, the contribution of the visual comfort to the overall 3DQoE can be decomposed into the contributions of the video and depth qualities on 3DQoE. Correspondingly, the visual comfort is not taken as an independent term in our proposed 3DQoE model

but implicitly contained in the model.

Based on the previous discussion, we can see that the video and depth qualities are the central elements of 3DQoE in the 2D plus depth based 3D video streaming. Nevertheless, how the video and depth qualities affect 3DQoE is still an open problem. Humans perceive the 3D experience in the brain via cyclopean view and depth perception [47]. However, how the cyclopean view and depth perception are fused into the final 3D perception is still unclear. In the literature, the energy of Gabor filter bank responses on the left and right view images was used to simulate the selection of cyclopean image quality [47], and the disparity information was only used to form the cyclopean view. Thus, they can only predict the cyclopean view quality to obtain the total 3D visual quality. Even though this method can predict the 3D visual quality, it also does not give an explicit mapping relationship between the cyclopean view quality and depth perception to the entire 3D perception. In [48], the mapping from the cyclopean view vision and depth perception to the 3DQoE was approximated as a linear relationship.

In the formation of the 3D perception, each factor has its own impact on the 3DQoE besides their relative interplay. The 3DQoE formation is not a simple additive process and it involves the interdependency among several factors. Thus, the conventional linearly additive measurement cannot fully characterize this process. We capture this complex 3D perception formation process in the brain by leveraging the advantages of fuzzy measurements in non-linear multi-attribute and target-oriented decisions. We characterize quantitatively the interdependency among different factors contributing to 3DQoE by proposing the use of Choquet-integral-based fuzzy measurements to determine 3DQoE. We consider the cyclopean view image quality, depth quality and content type as the main attributes affecting 3DQoE, and then transform this non-additive relationship between them and 3DQoE into an additive form in terms of the Choquet integral.

A. Choquet Integral

Let $X = \{x_1, x_2, \cdots, x_q\}$ be a set of feature attributes (such as the contributing factors on 3DQoE). Assume the attribute data consist of l observations $f_1, f_2, \cdots, f_q$ of feature attributes $x_1, x_2, \cdots, x_q$ and the objective attribute y (such as 3DQoE), and we have the form

$$\begin{array}{cccccc} x_1 & x_2 & \cdots & x_q & y \\ \hline f_{11} & f_{12} & \cdots & f_{1q} & y_1 \\ f_{21} & f_{22} & \cdots & f_{2q} & y_2 \\ \vdots & \vdots & & \vdots & \vdots \\ f_{l1} & f_{l2} & \cdots & f_{lq} & y_l \end{array}$$

where each row is an observation of feature attributes $x_1, x_2, \cdots, x_q$ and objective attribute y, and q is the number of feature attributes. Each observation of $x_1, x_2, \cdots, x_q$ can be regarded as a transform function $f : X \to (-\infty, +\infty)$. Correspondingly, the j^{th} observation of $x_1, x_2, \cdots, x_q$ is denoted by f_j, and we write $f_{ji} = f_j(x_i)$ where $1 \le i \le q$ and $1 \le j \le l$.

The inter-dependency among the set of feature attributes X contributing to the objective attribute y is characterized by a set function $\mu : P(X) \to \Re$ with $\mu(\emptyset) = 0$, where $P(X)$ is the power set of X and $\Re$ is the set of real numbers. The Choquet integral of function $f : X \to (-\infty, +\infty)$ with the fuzzy measure can be defined as [49]

$$\int_{(c)} f d\mu = \int_{-\infty}^{0} [\mu(F_a) - \mu(X)]da + \int_{0}^{+\infty} \mu(F_a)da, \quad (6)$$

where $F_a = \{x \mid f(x) \ge a\}$ for any $a \in (-\infty, +\infty)$ is called an *a-cut set* of f. Based on the Choquet integral model, the interaction among X towards y can be expressed as a non-linear regression form [50],

$$y = e + \int_{(c)} f d\mu + \epsilon, \quad (7)$$

where e is a constant (For some cases, e is zero) and ϵ is an error term that follows the normal distribution $N(0, \delta^2)$ with expectation 0 and variance δ^2.

As for the calculation of Choquet integral in (7) for the given set function μ and function f, Wang et al. [50] further proposed a genetic algorithm as

$$\int_{(c)} f d\mu = \sum_{s=1}^{2^q-1} (z_s \cdot \mu_s), \quad (8)$$

where μ_s denotes the fuzzy measurement and z_s denotes the corresponding contributing coefficient of μ_s in the whole integral.

Let $H_s = \min_{i:\mathrm{frc}\left(\frac{s}{2^i}\right) \in \left[\frac{1}{2}, 1\right)} (f(x_i)) - \max_{i:\mathrm{frc}\left(\frac{s}{2^i}\right) \in \left[0, \frac{1}{2}\right)} (f(x_i))$, and then in the above,

$$z_s = \begin{cases} H_s, & \text{if } H_s > 0 \text{ or } s = 2^q - 1 \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

In the last equation, $frc(\frac{s}{2^i})$ is the fractional part of $\frac{s}{2^i}$ and the maximum operation on the empty set is zero. When s is expressed in the binary form $s_q, s_{q-1}, ..., s_1$, $\{i \,|frc(\frac{s}{2^i}) \in [\frac{1}{2}, 1)\} = \{i \,|s_i = 1\}$ and $\{i \,|frc(\frac{s}{2^i}) \in [0, \frac{1}{2})\} = \{i \,|s_i = 0\}$.

B. 3DQoE Estimation

In 3D video transmission the cyclopean view is generally difficult to be calculated. The reason is that the true cyclopean view formation process in the human brain is currently a process not well understood. However, what is known is that the cyclopean view and depth perceptions rely on both the left view video and left view depth. Hence, we will show that the calculation of the cyclopean view perception in the 3DQoE model can be equivalently calculated by the Choquet integral of left view video quality, left view depth quality and 3D content type. This can minimize the necessary calculations for the 3DQoE model. Thus, in the remaining of this subsection we present the derivation of this proof.

Lemma 1 *The 3DQoE originally evaluated in terms of cyclopean view quality, depth quality, and video content type can be equivalently evaluated by the Choquet integral of the left*

view video quality Q_v, left view depth quality Q_d, and video content type ct as

$$\begin{aligned} 3DQoE = &\hat{z}_{Q_v} \cdot \mu(Q_v) + \hat{z}_{Q_d} \cdot \mu(Q_d) \\ &+\hat{z}_{ct} \cdot \mu(ct) + \hat{z}_{Q_v,Q_d} \cdot \mu(Q_v, Q_d) \\ &+\hat{z}_{Q_v,ct} \cdot \mu(Q_v, ct) + \hat{z}_{Q_d,ct} \cdot \mu(Q_d, ct) \\ &+\hat{z}_{Q_v,Q_d,ct} \cdot \mu(Q_v, Q_d, ct), \end{aligned} \tag{10}$$

where $\hat{z}_{Q_v}$, $\hat{z}_{Q_d}$, $\hat{z}_{ct}$, $\hat{z}_{Q_v,Q_d}$, $\hat{z}_{Q_v,ct}$, $\hat{z}_{Q_d,ct}$ and $\hat{z}_{Q_v,Q_d,ct}$ denote the fuzzy measurement coefficients corresponding to the fuzzy measurements of $\mu(Q_v)$, $\mu(Q_d)$, $\mu(ct)$, $\mu(Q_v, Q_d)$, $\mu(Q_v, ct)$, $\mu(Q_d, ct)$ and $\mu(Q_v, Q_d, ct)$, respectively.

Proof: By using the Choquet integral, the 3DQoE can be characterized as an integral of the cyclopean view quality Q_c, depth quality Q_d and video content type ct, that is

$$\begin{aligned} 3DQoE &= \int_{(c)} f \, d\mu = \sum_{s=1}^{2^q-1} (z_s \cdot \mu_s) \\ &= z_{Q_c} \cdot \mu(Q_c) + z_{Q_d} \cdot \mu(Q_d) + z_{ct} \cdot \mu(ct) \\ &+z_{Q_c,Q_d} \cdot \mu(Q_c, Q_d) + z_{Q_c,ct} \cdot \mu(Q_c, ct) \\ &+z_{Q_d,ct} \cdot \mu(Q_d, ct) + z_{Q_c,Q_d,ct} \cdot \mu(Q_c, Q_d, ct), \end{aligned} \tag{11}$$

where $\mu(x)$ denotes the fuzzy measurement of x and z_s denotes the coefficient for the corresponding fuzzy measurement s. In (11), $\mu(Q_c)$ can be further characterized as a Choquet integral of the left view video quality Q_v and and left view depth quality Q_d,

$$\mu(Q_c) = z'_{Q_v}\mu(Q_v) + z'_{Q_d}\mu(Q_d) + z'_{Q_v,Q_d}\mu(Q_v, Q_d). \tag{12}$$

Correspondingly, $\mu(Q_c, Q_d)$ can also be characterized as a Choquet integral of left view video quality Q_v and left view depth quality Q_d,

$$\mu(Q_c, Q_d) = z''_{Q_v}\mu(Q_v) + z''_{Q_d}\mu(Q_d) + z''_{Q_v,Q_d}\mu(Q_v, Q_d). \tag{13}$$

Similarly, $\mu(Q_c, ct)$ can be expressed as a Choquet integral of Q_v, Q_d and ct,

$$\begin{aligned} \mu(Q_c, ct) = &z'''_{Q_v} \cdot \mu(Q_v) + z'''_{Q_d} \cdot \mu(Q_d) + z'''_{ct} \cdot \mu(ct) \\ &+z'''_{Q_v,Q_d} \cdot \mu(Q_v, Q_d) + z'''_{Q_v,ct} \cdot \mu(Q_v, ct) \\ &+z'''_{Q_d,ct} \cdot \mu(Q_d, ct) + z'''_{Q_v,Q_d,ct} \cdot \mu(Q_v, Q_d, ct). \end{aligned} \tag{14}$$

And also,

$$\begin{aligned} \mu(Q_c, Q_d, ct) = &z''''_{Q_v} \cdot \mu(Q_v) + z''''_{Q_d} \cdot \mu(Q_d) \\ &+z''''_{ct} \cdot \mu(ct) + z''''_{Q_v,Q_d} \cdot \mu(Q_v, Q_d) \\ &+z''''_{Q_v,ct} \cdot \mu(Q_v, ct) + z''''_{Q_d,ct} \cdot \mu(Q_d, ct) \\ &+z''''_{Q_v,Q_d,ct} \cdot \mu(Q_v, Q_d, ct). \end{aligned} \tag{15}$$

Then, combining (11)-(15) with the fuzzy measurements, (11) can be further expressed as (10). ■

As discussed in the previous paragraphs of this section, the source coding rates and transmission PLRs for video and depth directly affect the video quality, depth quality, and finally the 3DQoE. Thus, the network-dependent 3DQoE model based on (10) can be rewritten as

$$\begin{aligned} 3DQoE(R_v, R_d, \rho_v, \rho_d) = &\hat{z}_{Q_v} \cdot \mu(Q_v(R_v, \rho_v)) \\ &+\hat{z}_{Q_d} \cdot \mu(Q_d(R_d, \rho_d)) + \hat{z}_{ct} \cdot \mu(ct) \\ &+\hat{z}_{Q_v,Q_d} \cdot \mu(Q_v(R_v, \rho_v), Q_d(R_d, \rho_d)) \\ &+\hat{z}_{Q_v,ct} \cdot \mu(Q_v(R_v, \rho_v), ct) \\ &+\hat{z}_{Q_d,ct} \cdot \mu(Q_d(R_d, \rho_d), ct) \\ &+\hat{z}_{Q_v,Q_d,ct} \cdot \mu(Q_v(R_v, \rho_v), Q_d(R_d, \rho_d), ct), \end{aligned} \tag{16}$$

where R_v, R_d, ρ_v and ρ_d denote the coding rates and transmission PLRs for video and depth, respectively.

Even though we have a closed-form 3DQoE model in (16), the specific values of the coefficients of the fuzzy measurements need to be obtained through off-line training. To do that one can construct the desired video and depth transmission patterns for different bit-rate levels and different packet loss situations. By using the video qualities, depth qualities and the content types under different 3D video transmission patterns, we can obtain the corresponding subjective 3DQoE values under different transmission environments.

The range of value for each factor (the content type, depth quality or video quality) is different. To distinguish fairly the contribution of each factor in 3DQoE, their values need to be normalized into the same dimension. Specifically, a linear normalization is performed independently for each factor as $\chi' = (\chi - \chi_{\min}) \cdot \frac{(\chi'_{\max} - \chi'_{\min})}{(\chi_{\max} - \chi_{\min})} + \chi'_{\min}$ [51], where χ' and χ are the normalized and original values of the factor, respectively. Also, $\chi'_{\min}$ and $\chi'_{\max}$ are set to 0 and 1, respectively, and $\chi_{\max}$ and $\chi_{\min}$ are the original maximum and minimum values of the factor. After that, the Choquet transforms are performed with these normalized values for different transmission patterns and then the general fuzzy measurement is calculated by multi-regression.

Based on these fuzzy measurements, we estimate the 3DQoE by predicting the video and depth qualities during the 3D video transmission. At the 3D video sender, the distortion of video and depth can be estimated in terms of the predicted PLRs of different layers by simulating the error-corrupting decoding process [52]. Specifically, the PLRs for different layers that are transmitted over different paths are estimated first by (5). By considering the layer-dependency in the SVC decoding, the PLR τ_{k_v} for that only the first k layers are decodable in the total L_v layers for video can be expressed as

$$\tau_{k_v} = \begin{cases} \rho_{k_v} \times \prod_{j=1}^{k-1} (1 - \rho_{j_v}), 1 < k < L_v \\ \rho_{k_v}, \qquad k = 1 \\ \prod_{j=1}^{L_v} (1 - \rho_{j_v})), \quad k = L_v \end{cases} \tag{17}$$

where ρ_{k_v} and ρ_{j_v} denote the PLR for the k^{th} and j^{th} layers of video, respectively. Similarly, the PLR τ_{kd} for that only the first k layers are decodable in the total L_d layers for depth can be computed. Thus, the video and depth qualities in terms of PSNR can be obtained independently and then the 3DQoE value can be estimated via (16) by substituting τ_{k_v} and τ_{k_d} for ρ_v and ρ_d. The specific 3DQoE estimation flowchart is shown in Fig. 3.

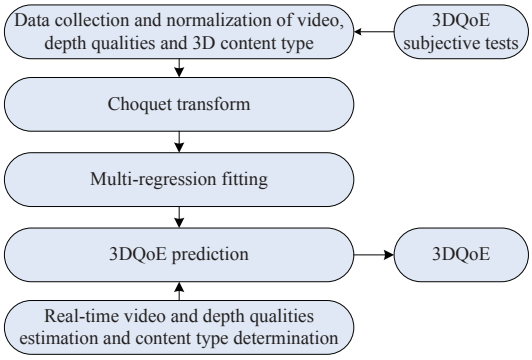


Fig. 3. The 3DQoE estimation flow.

TABLE I
TESTED VIDEO CONTENT.

Sequences	SR	Content type
Balloons	1024×768	medium motion level
Kendo	1024 ×768	medium motion level
Book_arrival	1024×768	low motion level
Newspaper	1024×768	low motion level
Dancer	1920×1080	high motion level
Poznan_hall	1920×1080	medium motion level
GhosttownFly	1920×1080	high motion level

C. 3DQoE Model Evaluation

In current stage, the online 3D videos distributed over networks can be easily displayed on mobile 3D screens. Aiming at mobile 3D visual experience, we performed extensive subjective tests with 15.6 inch lenticular lenses based on a stereoscopic PC notebook display for different 3D video contents.

The test video content types that correspond to different spatial resolutions (SR) are summarized in Table I and their sample images are shown in Fig. 4. In Table I, the d_l and d_u are set to 2.6 and 4.2, respectively for determining the content type. The combination of quality levels (compression bit-rate levels for video and depth) and PLR are shown in Table II. In the subjective testing, one left view video and another right view video synthesized by the left view video and depth were used as the inputs of 3D video display. The view synthesis reference software (VSRS) in MPEG [53] was used for virtual view video generation. The proposed 3DQoE model was trained with the 3DQoE patterns which were combined with different rate levels, different PLRs and different content types. In the tests, we randomly selected 80% patterns as the stimuli to train the 3DQoE model and then utilized the remaining 20% of the patterns to verify the model accuracy.

The subjective test adopts the SSIS (Single Stimulus Impairment Scale) method described in ITU-R BT. 500 [54] and ITU-R BT.2021 [55]. In the subjective tests, the volunteer viewers sat in front of the screen with comfortable distance and the field of view was about 15^o. The measured environmental illumination was 210 lux. The specific viewing conditions were consistent with the requirements of the ITU-R BT.2021 standard. A total of 30 subjects participated in the tests with a maximum age of 41 years and a minimum age of 20 years, consisting of 23 males and 7 females. Before the subjective testing, the subjects were all required to carefully read the viewing instructions. More specifically, the procedures of Single Stimulus measurement in ITU-R BT.2021 were performed. The tests were carried out with a series of judgment sessions which include 3D video sequence viewing duration of 10s and voting duration of 5s. In each session, only one test sequence with one 3DQoE pattern was presented on the stereoscopic display. The subjects received break every 30 minutes of 3DQoE evaluation, as suggested in ITU-R BT.2021.

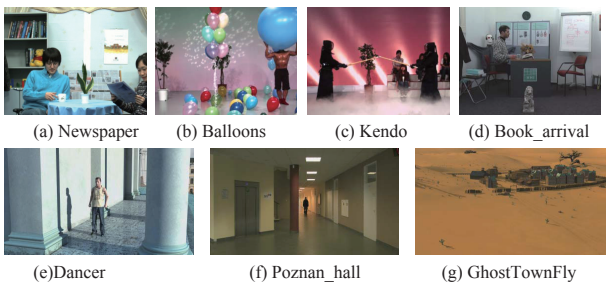


Fig. 4. The left view images for subjective tests.

The opinion score that each viewer reported was continuously mapped to a value between 0 (bad perceptual quality) to 10 (excellent perceptual quality). During verification the rating result obtained by the 3DQoE model was not the direct mean opinion score (MOS) value, but it was converted to the MOS values. The non-linear logistics function in [56] was used to map the results rated by 3DQoE model to the MOS values.

The extensive subjective test results for the 3DQoE model are shown in Fig. 5, and they illustrate the strong interdependency among left view video quality Q_v, depth quality Q_d and video content type ct towards forming the overall 3DQoE. The non-additive interaction measurement can be observed from Fig. 5, where $\mu(Q_v) + \mu(Q_d) \neq \mu(\{Q_v, Q_d\})$, which means that the joint contribution of $\{Q_v$,$Q_d\}$ to 3DQoE is not equal to the sum of the individual contributions made by Q_v and Q_d.

Fig. 6 shows the consistency of the mapped MOS values from the proposed 3DQoE model and actual human subjective evaluation. According to Fig. 6, it is evident that there is a statistically high correlation between the quality ratings from 3DQoE model and the real measurements under different network conditions. The quantitative prediction errors of 3DQoE model are compared with the real measurements and are reported in Table III. The root mean square error (RMSE), Pearson correlation coefficient (PC) and Spearman rank order correlation coefficient (SROC) between the predicted MOS values and the measured MOS values were computed. A low

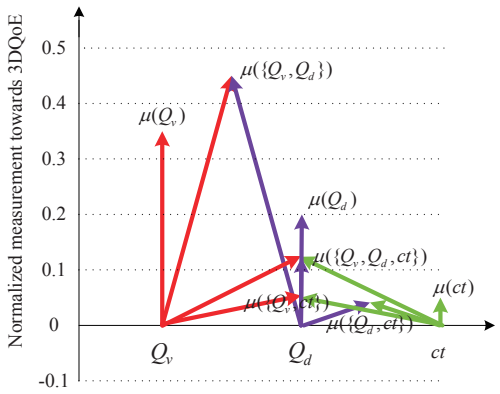


Fig. 5. Interaction measurement for three contributing factors Q_v, Q_d and ct to the overall 3DQoE.

TABLE II
3DQOE PATTERNS WITH DIFFERENT RATE LEVELS AND DIFFERENT PLRS

SR	Bit-rate level (kbps)		PLR	Content type
1024 × 768	Video	Depth	0%,1%,3%,5%,7%	Low and medium motion levels
	1000,1500,2000,2500,3000	500,1000,1500,2000,2500	0%,1%,3%,5%,7%	
1920 × 1080	Video	Depth	0%,1%,3%,5%,7%	Low and high motion levels
	3000,3500,4000,4500,5000,6000,7000	1500,2000,2500,3000,3500,4000,5000	0%,1%,3%,5%,7%	

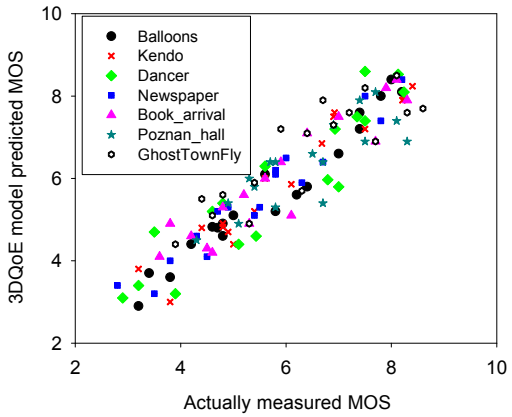


Fig. 6. The predicted MOS values from the 3DQoE model versus actually measured MOS values.

RMSE, a high PC and a high SROC all suggest reliable prediction accuracy.

From Table III, it is obvious that the 3D video viewing experience predicted by the proposed 3DQoE model is highly correlated with the actual result of the evaluation. In the experiments, we compared our model to the state-of-the-art 3D QoE or 3D quality prediction models, such as the automatic QoE model [57], depth perception evaluation model [22] and the 3D quality analysis model [23]. When compared to the 3DQoE model in [57] that offers a prediction accuracy of 95% with R^2 evaluation (which is equal to a PC 0.91), the prediction accuracy of our proposed model is 94% corresponding to a PC 0.88. Similarly, when compared to [22] that provides the evaluation accuracy with a PC 0.8 and an RMSE 0.38, and [23] that provides a PC 0.84, and an RMSE 0.078 in the grade range from 0 to 100, our model obtains a PC 0.88, a SROC 0.87 and an RMSE 0.36 in grade range from 0 to 10. Hence, the comparison results indicate that our proposed model can obtain a prediction accuracy close to that in [57] at lower complexity and an improved prediction accuracy to those in [22] and [23]. Compared with the complex and accurate content clustering approach in [57], a simple motion type classification approach was adopted in our model, therefore the model in [57] can provide more accurate prediction for the impact of 3D video content on the overall 3DQoE than our model. In the future work, we shall consider a more accurate content analysis approach to further improve the model accuracy.

V. 3DQOE-ORIENTED AND ENERGY-EFFICIENT 3D VIDEO FLOW ROUTING OPTIMIZATION FRAMEWORK

With the presence of the SDN controller, the multiple scalable 3D streams can be dynamically forwarded to different paths. Besides 3DQoE, the energy cost is also needed so as to perform the forwarding decisions. In this section we present

TABLE III
THE ACCURACY OF THE PROPOSED 3DQOE MODEL.

Sequences	RMSE	PC	SROC
Balloons	0.3601	0.8935	0.9123
Kendo	0.3523	0.8867	0.8931
Book_arrival	0.3364	0.9159	0.9053
Newspaper	0.3876	0.8514	0.8661
Dancer	0.3271	0.9304	0.9225
Poznan_hall	0.4037	0.8434	0.8325
GhosttownFly	0.3838	0.8763	0.8887
Average	0.3644	0.8811	0.8797

first the energy cost model, and then we formulate and solve the complete 3DQoE-oriented and energy-efficient 3D video path selection problem.

A. Energy Cost Model

The energy cost of the active network elements in each path are important for deciding the optimal routing path. In a wireline network, the energy cost is dominated by the energy of the network equipment that includes the routers, repeaters and Ethernet cards. The consumed energy in these devices is related to the volume of transmitted data, the state of the device and the particular type/model of the device. One transmission path includes one source node, several routers and one destination node. The energy consumption for one transmission path mainly comes from the ports and Ethernet cards that are used, besides the chassis energy consumption of devices. The source and destination nodes are also characterized by the same behavior.

The energy consumption of a port is related to the volume of data that it transports. Generally, the energy consumption of the port is an approximately linear function of transmission payload [36]. Similarly, the port capacity may be increased with the rate of the communicated traffic. For example, the capacity of a 10Gbps port can also be set to 10Mbps and 100Mbps. Thus, the energy cost for a port also varies dynamically with the used transmission rate. When a transmission rate of 20Mbps is used for a port, the energy cost is less than that of 100Mbps and more than that of 10Mbps. Hence, the piecewise energy cost [36] for a port is modeled as

$$E_{port}(R) = \begin{cases} E_0 & if \;\; R = 0 \\ k_1 R + E_1, & if \;\; R \in [0, C_1] \\ \vdots & \\ k_i R + (E_i - k_i C_{i-1}) & if \;\; R \in [C_{i-1}, C_i] \\ \vdots & \\ k_\tau R + (E_\tau - k_\tau C_{\tau-1}) & if \;\; R \in [C_{\tau-1}, C_\tau] \end{cases} \tag{18}$$

In the above k_i denotes the slope of each capacity level of the port, E_i denotes the starting energy cost of each level, C_i is the piecewise bandwidth capacity of the port and τ is the number of the piecewise bandwidth capacities that the

port can provide. The parameters in (18) can be obtained experimentally for each specific device model.

For each Ethernet card, the energy cost is denoted as E_{Eth}, which is generally a constant for a specific model. Consequently, the energy cost for the j^{th} edge segment (link) in one routing path is

$$E_{ES_j}(R) = \sum_{i=1}^{N_p} E_{port}(R) + \sum_{i=1}^{N_{Eth}} E_{Eth} + E_{chassis} \quad (19)$$

where N_p is the number of ports that are used for transmitting rate R in the current edge segment and N_{Eth} is the number of Ethernet cards for transmitting at rate R. Also, $E_{chassis}$ is the chassis energy cost of the devices in the j^{th} edge segment.

To obtain an even more realistic model that matches modern networks, we consider that the last hop is wireless. In this case the wireless network energy cost is primarily related to the transmission power of the wireless access point (AP). The energy consumption model in [58] is used to calculate the wireless energy as

$$E_{wireless}(R) = \begin{cases} \alpha(s) \cdot R + \beta, & if\ 0 \le R \le h(s) \\ \gamma, & if\ R > h(s) \end{cases}, \quad (20)$$

where $\alpha(s)$ is the energy cost per bit for datagrams with size s bytes, β is the amount of energy consumed by the wireless AP in idle mode, γ is the maximum amount of energy consumed by the wireless AP, R is the transmission rate of the wireless gateway and $h(s)$ is a threshold value that can be set to a different value for different types of access gateways.

To get the energy cost for the entire path, we assume that the energy cost of the input and output ports are the same. Consequently, for the transmission path p that consists of M segments $(ES_1, ES_2, ..., ES_M)$ (the last segment is the wireless link), its energy cost with transmission rate R in one path routing session over SDN can be computed as

$$E_p(R) = \sum_{j=1}^{M-1} E_{ES_j}(R) + E_{wireless}(R) \quad (21)$$

B. Joint Rate Allocation and Flow Routing Optimization

Given $G(V_N, ES_N)$, a number of different candidate paths m_p from source node H0 to the destination node H1 can be selected. The value of m_p as well as the actual paths can be calculated by a depth-first search algorithm. For each of these paths, the BER and and bandwidth can be collected by the SDN controller in real-time. Depending on the aggregate bandwidth of the available transmission paths we can select n_p candidate data layers (i.e., flows including n_v video layers and n_d depth layers with $n_v + n_d = n_p$) to be distributed over n_p different paths from the available m_p ($n_p < m_p$). With the total available rate constraint R_a, the video and depth layers n_v, n_d must have rates R_v and R_d that satisfy $R_v + R_d < R_a$.

Assume that a number of n_R video/depth rate allocation scenarios can be found. For each rate allocation scenario, each layer has different size and all of them can be ranked with a decreasing order in size that is denoted as $\vec{R}$. When these streams are distributed over the different n_p paths among the m_p, there are $A_{m_p}^{n_p}$ path selection scenarios. Thus, we can obtain $n_R \cdot A_{m_P}^{n_P}$ combinations of 3D stream rate allocation and path selection options. We denote the set of 3D stream rate allocation and path selection scenarios as U_{RP}. If each path selection option is denoted as $\vec{P}$, the integrated 3D stream rate allocation and path selection options can be expressed as $\vec{P}(\vec{R})$ with $\vec{P}(\vec{R}) \in U_{RP}$.

The decisions for R_v and R_d in (16) can be characterized by those of $\vec{R}$. Similarly, the decisions for ρ_v and ρ_d in (16) can be characterized by $\vec{P}$. Thus, the predicted 3DQoE $3DQoE(\vec{P}(\vec{R}))$ can be computed using (16) by substituting $Q_v(\vec{P}(\vec{R}))$ and $Q_d(\vec{P}(\vec{R}))$ with $Q_v(R_v, \rho_v)$ and $Q_d(R_d, \rho_d)$. In the same way, the energy cost $E(\vec{P}(\vec{R}))$ for each rate allocation and path selection scenario can also be computed by using (21). Note that $E(\vec{P}(\vec{R}))$ denotes the energy cost for transmitting multiple scalable 3D video streams over a group of selected paths $\vec{P}$, while the energy cost for transmitting one scalable stream over each path can be computed by (21).

With the previous analysis, the optimal rate allocation and the optimal routing path can be determined by solving the following optimization problem

$$\vec{P}^{opt}(\vec{R}^{opt}) = \underset{\vec{P}(\vec{R}) \in U_{RP}}{\arg\max} \frac{3DQoE(\vec{P}(\vec{R}))}{E(\vec{P}(\vec{R}))}, \quad \text{s.t. } R_v + R_d < R_a \quad (22)$$

where $\vec{P}^{opt}(\vec{R}^{opt})$ denotes the selected optimal rates and the optimal paths for video and depth.

Unfortunately (22) is a non-linear optimization problem. There is a non-linear fractional objective, plus discrete combinatorial decisions. The combinatorial nature makes this non-linear optimization problem NP-hard which means that no existing algorithms can provide an exact solution in polynomial time. To solve such an optimization problem an exhaustive search is needed to ensure optimality at the cost of exponential complexity.

To reduce the computational complexity, we propose a low-complexity algorithm based on the branch and bound concept. Specifically, we propose to prune the search space of the candidate video/depth rate allocation options, i.e., n_R. This can be done by removing options whose video/depth rate allocation ratio is less than a threshold value that indicates a certain quality ratio between video and depth. We experimentally calculated this threshold value to be 1.7. After that, the range of the candidate paths can be reduced by removing from the search space paths with high PLR. We also use a PLR threshold to do the above. For the remaining candidate transmission paths, the possible rate allocations between video and depth are determined afterwards. For each possible video/depth rate allocation corresponding to these paths, the 3DQoE and energy cost can be estimated by using (16) and (21), respectively. When the 3DQoE values and energy costs for all candidate rate allocation scenarios are obtained, the scenario with the maximal ratio of 3DQoE to energy cost is selected. The branch and bound algorithm for the joint video/depth rate allocation and routing path selection that we briefly described is presented in Algorithm 1. By periodically executing this

algorithm, we can have a dynamic 3DQoE-oriented and and energy-efficient 3D video flow routing optimization.

Regarding the computational complexity, the depth-first search for determining m_p requires $O(n[V+e])$ time, where V and e are the numbers of the nodes and edges of the graph. Thus, the overall computational complexity of Algorithm 1 is $O(n[V+e]+n[n_R \cdot A_{m_p}^{n_p} \cdot n_p])$. When reducing n_R, V will be reduced. As stated above, via appropriately reducing n_R and n_p, the significantly lower complexity of the proposed algorithm will lead to fast solutions.

Algorithm 1 Joint rate allocation and routing path selection

Input: R_a, $PLR_{threshold}$, λ_{th}, MAX
Input: layer sizes information of video S_v and layer sizes information of depth S_d
Output: the optimal rate allocation and path selection scenario $\vec{P}^{opt}(\vec{R}^{opt})$

Determine m_p using depth first search
Find the number n_R of the rate allocation scenarios which satisfy $R_{3D}=R_v+R_d<R_a$ and $R_v/R_d>\lambda_{th}$ in terms of layer size information S_v and S_d;
for each rate allocation scenario $\vec{R}$ in n_R **do**
  $3DQoERopt \leftarrow MAX$;
  $\vec{P}^{opt}(\vec{R}^{opt}) \leftarrow null$;
  Find $A_{m_p}^{n_p}$ paths set $\vec{P}(\vec{R})$;
  for each path p in $\vec{P}(\vec{R})$ **do**
    Compute the packet loss rate PLR_p of the path with corresponding rate;
  end for
  Compute the average packet loss rate for $PLR_a(\vec{P}(\vec{R}))$;
  if $PLR_a(\vec{P}(\vec{R})) > PLR_{threshold}$ **then**
    $3DQoE(\vec{P}(\vec{R})) \leftarrow 0$;
  else
    Compute $3DQoE(\vec{P}(\vec{R}))$;
    for each path p in $\vec{P}(\vec{R})$ **do**
      Compute the energy cost $E(p)$;
    end for
    Compute the total energy cost $E(\vec{P}(\vec{R}))=\sum_{p\in\vec{P}(\vec{R})} E(p)$;
    Compute $3DQoE(\vec{P}(\vec{R}))\Big/E(\vec{P}(\vec{R}))$;
  end if
  if $3DQoERopt < \{3DQoE(\vec{P}(\vec{R}))\Big/E(\vec{P}(\vec{R}))\}$ **then**
    $3DQoERopt = 3DQoE(\vec{P}(\vec{R}))\Big/E(\vec{P}(\vec{R}))$;
    $\vec{P}^{opt}(\vec{R}^{opt}) = \vec{P}(\vec{R})$;
  end if
end for
Select the optimal rate allocation and path selection scenario $\vec{P}^{opt}(\vec{R}^{opt})$ to transmit 3D video;

VI. Experimental Results

Based on the network emulator mininet [59], we implemented the 3DQoE-oriented and energy-efficient scalable 3D video streaming system. In Fig. 7 we illustrate the network topology that consists of 23 switches (S0, S1, $\cdots$, S22) and 45 links. In this figure the dotted lines indicate an edge segment (link) between the nodes, and the solid arrow line denotes the direction of link channel. We modeled three congestion levels for $G(V_N, ES_N)$ by assuming a Poisson random process model [60] with mean available transmission rate of 3Mbps, 4Mbps and 5Mbps. The BER and data rate for each link were randomly generated. The maximal and minimal BER for the links were set to 5×10^{-6} and 10^{-8}, respectively. Our optimization algorithm was periodically performed for each group of pictures (GOP) of the transmitted video.

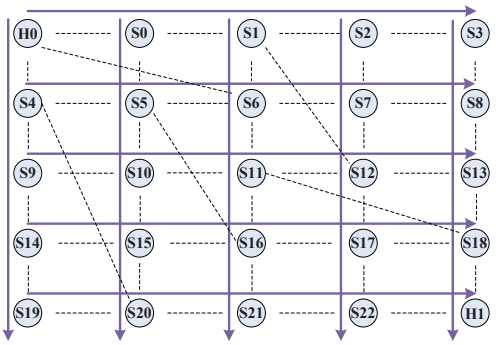


Fig. 7. Experimental network topology.

To match the random properties of the different network devices in real networks, we assume in our simulation that these nodes are equipped with a random number of cards, ports, and transmission capacities. The specific power consumption of different types of cards reported in [61] were used to compute the energy consumption. For the wireless part, the energy cost model of WiFi with the parameters reported in [58] was used. In the experiments, the 3DQoE values were limited in the range of 0 to 10 and the energy cost was limited in the range of 0 to ∞. Since the energy cost is considerably larger than the 3DQoE, the energy cost values are scaled by $E_s=\frac{E}{\upsilon_s}$, where E_s and E denote the scaled and original energy cost values, respectively. In our experiments, the scaling coefficient υ_s was empirically set equal to 800 in terms of the valid range of the energy cost E.

TABLE IV
THE 3D VIDEO SOURCE BIT-RATES FOR DIFFERENT LAYERS (Q1, Q2, Q3, Q4 DENOTE THE QUALITY LAYERS FROM LOW TO HIGH)

Sequences	Balloons bit-rate(kbps)		Newspaper bit-rate(kbps)	
Video/depth	Video	Depth	Video	Depth
Q1	1134.30	580.48	1256.23	467.38
Q1+Q2	1934.96	1046.75	1868.26	847.43
Q1+Q2+Q3	2452.68	1427.36	2263.20	1120.34
Q1+Q2+Q3+Q4	2748.03	1656.92	2556.87	1428.21

Regarding the configuration of the 3D video source data subsystem, the depth map was generated by the depth estimation software provided by Tanimoto Laboratory of Nagoya University [62]. The video and depth data of the 3D videos Balloons and Newspaper with spatial resolutions of 1024×768, were independently encoded with the H.264/SVC reference software JSVM9.19. The 3D video sequences consist of 2500 frames and they were encoded at four video quality

layers and four depth quality layers to provide different levels of source rates that range from 1.5Mbps to 5Mbps, as shown in Table IV. The GOP size was set to 8 frames. To provide an error-resilient decoding, the openSVC decoder [63] was used by enabling the frame-copying error concealment option.

To compare the benefits of the proposed 3DQoE-oriented and energy-efficient 3D video streaming system against the state-of-the-art solutions we examined the following systems. First we tested the non-scalable 3D video streaming system. For this system, the fixed video and depth rates were encoded by H.264/AVC and the path routing was optimized by achieving an optimal ratio of 3DQoE to energy cost. Second, the fixed path H.264/SVC-based 3D video streaming system was also tested. In this system the optimal paths were selected so as to minimize the network congestion. Specifically, the appropriate transmission paths for video and depth streams were calculated first by matching the source data with the available bandwidths of different paths (to the extent that it was possible) and then the 3D video transmission over the selected paths was simulated in our streaming platform. Third, the QoS-enabled adaptive video streaming system for 3D video reported in [32] was utilized as the reference system. In this system the weighted packet loss measurement and delay variation was used as the QoS cost to determine the routing path in order to ensure a fair comparison.

A. Video/Depth Rate Allocation Performance

In this subsection, we evaluate the influence of bandwidth-adaptive video/depth rate allocation on the system performance. Fig. 8 shows the performance comparisons between the fixed-rate non-scalable 3D video streaming and the proposed 3DQoE-oriented and energy-efficient (denoted as 3DQoE-EE) 3D video streaming (that includes the optimal path routing) for different congestion levels. It can be seen from Fig. 8 that the proposed system can provide a higher ratio of the 3DQoE to energy cost than the non-scalable 3D video streaming system. For the 3DQoE-EE system, the ratio of 3DQoE to energy cost is decreased when congestion level changes from 3Mbps to 4Mbps for both video sequences. This is because 3DQoE is improved due to the higher available bandwidth but the energy cost is also increased at the same time. For the non-scalable 3D video streaming system, the ratio of 3DQoE to energy cost is gradually increased with the increased available transmission

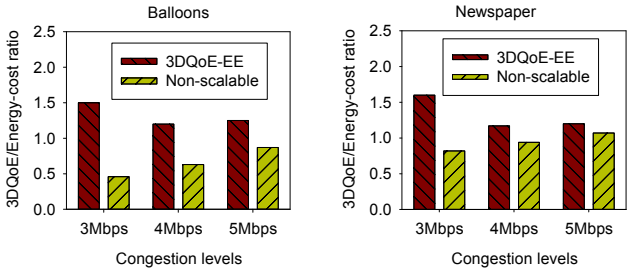


Fig. 8. Performance comparison in terms of the objective (3DQoE/Energy Cost), between the non-scalable system and the proposed 3DQoE-EE streaming system for different levels of network congestion.

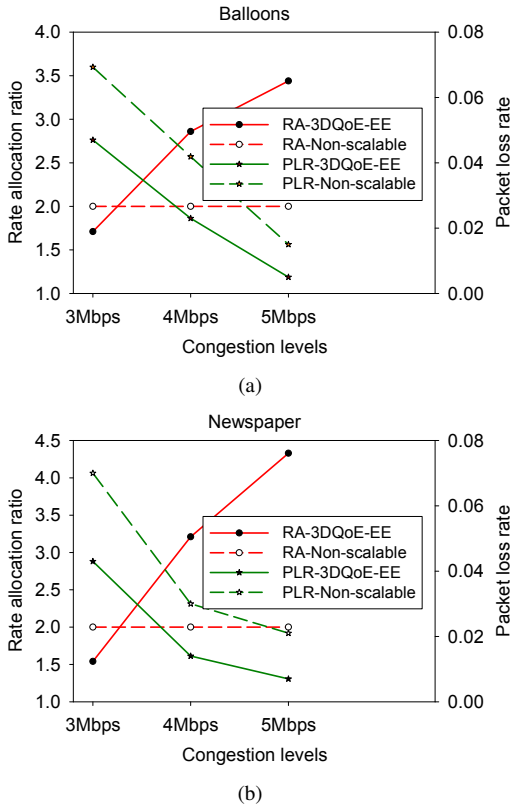


Fig. 9. The video/depth rate allocation ratios and packet loss rates for different congestion levels: (a) is for Balloons sequence and (b) is for Newspaper sequence.

bandwidth. The time-varying video/depth rate allocation (RA) ratios and PLRs for the two systems are shown in Fig. 9. In Fig. 9, "RA-3DQoE-EE" and "RA-Non-scalable" denote the rate allocation ratios of the 3DQoE-EE and non-scalable systems, respectively. Likewise, the "PLR-3DQoE-EE" and "PLR-Non-scalable" denote the PLRs of the 3DQoE-EE and non-scalable systems, respectively. It can be seen from both Fig. 9(a) and Fig. 9(b) that the video/depth rate allocation ratio is adaptively regulated for different congestion levels and it results in lower PLR compared to the non-scalable 3D video streaming system since it can reduce congestion.

Fig. 10 presents the energy cost comparison between non-scalable 3D video streaming system and the proposed 3DQoE-EE 3D video streaming system. In this figure we observe that the energy cost for the proposed system is slightly higher than the non-scalable 3D video streaming. This is because the proposed system involves several network flows for a single 3D video flow (scalable video and depth flows) that consume more energy than the non-salable counterpart that consists of only two flows (video and depth).

B. Path Routing Performance

Table V depicts the time-varying routing path changes of the video and depth base layers for the Balloons sequence at different congestion levels under the proposed system. It can be seen from Table V that the proposed system can dynamically select the routing paths that are enforced by the SDN controller. Moreover, as a result of the optimization,

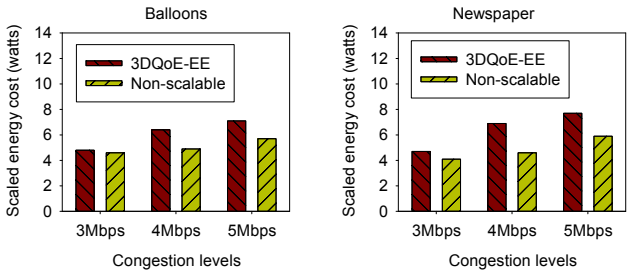


Fig. 10. Energy cost comparison between non-scalable and the proposed 3DQoE-EE streaming systems.

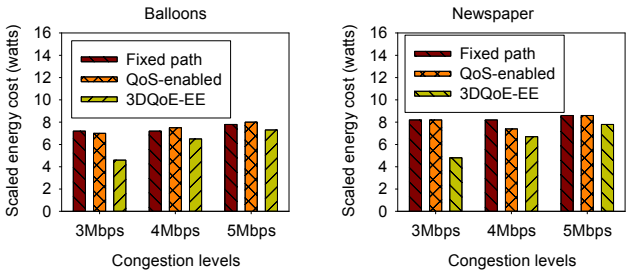


Fig. 11. The energy cost (scaled) comparison between the proposed system and QoS-enabled adaptive streaming system.

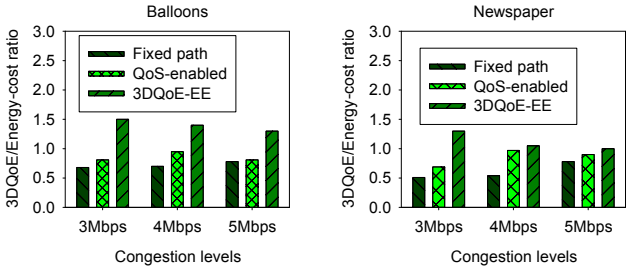


Fig. 12. The ratio of 3DQoE to energy cost for the three streaming systems.

the corresponding average PLRs for video and depth base-layers are significantly reduced when the congestion level is increased.

Next, Fig. 11 shows a comparison between the energy costs of three systems. We notice that the proposed system typically consumes lower energy than the other two systems for the same congestion level. If we look into the results more carefully, we notice that the QoS-enabled adaptive streaming system consumes more energy than the proposed system since it does not consider the energy cost in the path routing algorithm. In the experiment, we also recorded the ratios of 3DQoE to energy cost for the three streaming systems and we present the related results in Fig. 12. It can be seen from this figure that the proposed system can provide a higher ratio than the other two systems. It is important to recall that the QoS-enabled adaptive streaming system also allows routing path changes. However, it uses a QoS and not a 3DQoE metric for deriving the streaming routing policy which does not lead to the optimal 3DQoE performance. We also note that for the considerably different video sequences that we used (the medium motion video Balloons and the low motion video Newspaper), the impact of congestion on 3DQoE is also different. In both cases our system can improve 3DQoE by taking into account the content characteristics. Overall, the 3DQoE-EE system can achieve superior performance when compared to the QoS-enabled system since the content type is embedded into optimization metric.

C. Dynamic Streaming Quality

Congestion is a dynamic phenomenon in the networks we study and so we evaluate the ability of our system to deal with this type of events. The proposed system can dynamically configure the routing paths by adapting to the level of congestion. We introduced the cross-traffic with time-varying dynamic bandwidth to change the congestion level in the experiment. Fig. 13 illustrates the evolution over time of the 3DQoE for the proposed 3DQoE-EE and the QoS-enabled adaptive streaming systems. It can be seen that the proposed system can always obtain superior performance to the QoS-enabled adaptive streaming system. During the time period between 70-90 seconds, where the congestion is severe, the proposed system can provide a rapid configuration of transmission paths so that it can improve performance.

TABLE V
ROUTING PATH CHANGES AT DIFFERENT CONGESTION LEVELS FOR BALLOONS SEQUENCE

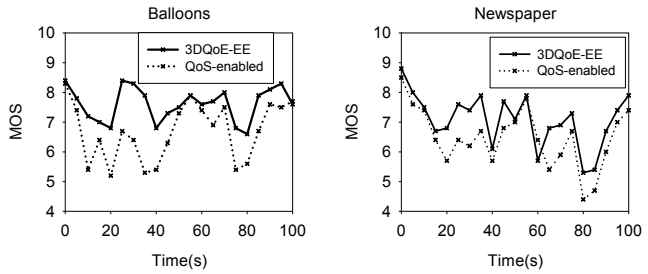
Congestion levels	Layer stream	Traffic paths	Average PLR
3Mbps	Video	H0→S4→S5→S16→S17→S18→H1	0.031
	Depth	H0→S0→S1→S12→S13→S18→H1	
4Mbps	Video	H0→S0→S5→S16→S21→S22→H1	0.016
	Depth	H0→S4→S9→S10→S11→S18→H1	
5Mbps	Video	H0→S4→S20→S21→S22→H1	0.001
	Depth	H0→S6→S7→S8→S13→S18→H1	

The comparative results for the PLR between the two systems are shown in Fig. 14 and correspond precisely to the experiment we discussed in the previous paragraph. There is an inverse relationship between MOS and PLR that can easily be identified in Figs. 13 an 14.

Finally we present in Fig. 15 the real-time PSNR values for the proposed 3DQoE-EE and the QoS-enabled adaptive streaming systems. The PSNR values correspond to the first 1000 frames and a time period of 40 seconds for the right view video generated by left view video and depth map. It can be seen that the objective PSNR values can provide similar performance to the subjective evaluations. Moreover, the shape of the PSNR curve follows approximately the PLR curve, which also supports our claim that the proposed system can outperform the QoS-enabled streaming system in terms of the practical 3D video streaming performance.

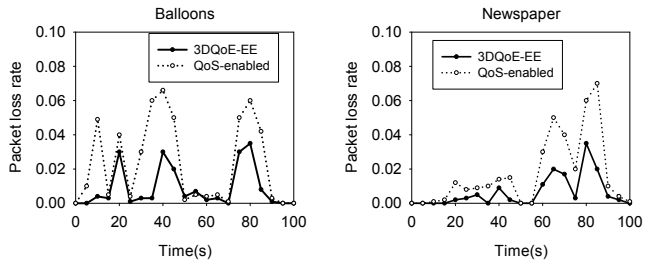
VII. CONCLUSIONS

This paper presented a 3DQoE-oriented and energy-efficient 3D video streaming framework for centrally controlled networks. Utilizing SDN as an instance of a centrally controlled



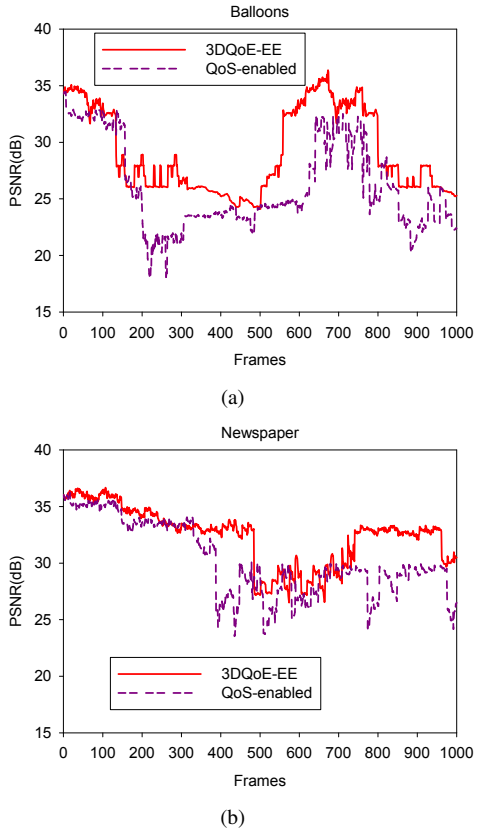
(a) (b)

Fig. 13. MOS comparisons between the proposed 3DQoE-EE and the conventional QoS-enabled streaming systems with time-varying congestion levels.



(a) (b)

Fig. 14. PLR comparisons between the proposed 3DQoE-EE and the conventional QoS-enabled streaming systems with time-varying congestion levels.



(a)

(b)

Fig. 15. PSNR comparisons between the proposed 3DQoE-EE and the QoS-enabled streaming systems for the right view video generated by left view video and depth map.

network, our streaming system exploits both the features of SDN and the properties of scalable 3D video by integrating them in a cross-layer optimization framework. To accomplish the above, the 3D visual experience formation process in the human brain was first modeled with the Choquet integral. The proposed 3DQoE model characterizes different factors that affect the entire 3D video streaming chain especially the interplay between different factors on the overall 3D experience. Next, we developed a realistic and comprehensive energy model for the SDN. Based on these two models, we formulated the problem of joint rate allocation and flow routing to maximize the ratio of the 3DQoE to the energy cost. We conducted extensive experiments and they demonstrate that the proposed optimization framework can provide superior streaming performance for 3D VoD streaming applications by ensuring a balance between visual experience and operator cost.

REFERENCES

[1] J. Thatte, J.-B. Boin, H. Lakshman, B. Girod, "Depth Augmented Stereo Panorama for Cinematic Virtual Reality with Head Motion Parallax", in *Proc. IEEE International Conference on Multimedia and Expo*, Seattle, USA, July 2016.

[2] K. Müller, H. Schwarz, D. Marpe et al., "3D high efficiency video coding for multi-view video and depth data", *IEEE Trans. Circ. Syst. Video Technol.*, vol. 22, no. 9, pp. 3366-3378, Sep. 2013.

[3] Cisco Visual Networking Index: "Forecast Highlights, 2012-2017", May 2013.

[4] YouTube Blog, "Now in 3D: Join the Experiment With Us!", 2009. [Online]Available: http://bit.ly/lZ6pd. Accessed on: May 2, 2017.

[5] A. Vetro, A.M. Tourapis, K. Müller, T. Chen, "3D-TV content storage and transmission", *IEEE Trans. Broadcast.*, vol. 57, no. 2, pp. 384-394, June 2011.

[6] L.S. Karlsson, M. Sjostrom, "Layer assignment based on depth data distribution for multiview-plus-depth scalable video coding", *IEEE Trans. Circ. Syst. Video Technol.*, vol. 21, no. 6 pp.742-754, June 2011.

[7] Y. Liu, Q. Huang, S. Ma et al., "Joint Video/Depth Rate Allocation for 3D Video Coding based on View Synthesis Distortion Model", *EURASIP Signal Process.: Image Commun.*, vol.24, no.8, pp. 666-681, Sep. 2009.

[8] J. Apostolopoulos, T. Wong, W. Tan, S. Wee, "On multiple description streaming with content delivery networks", in *Proc. IEEE Infocom*, June 2002, pp. 1736-1745.

[9] Z. Bozakov, A. Rizk, D. Bhat, and M. Zink, "Measurement-based Flow Characterization in Centrally Controlled Networks", in *Proc. IEEE Infocom*, 2016.

[10] [Online] Available: http://www.openflow.org. Accessed on: May 20, 2017.

[11] U. Lee, I.Rimac, D. Kilper and V. Hilt, "Toward Energy-Efficient Content Dissemination", *IEEE Network*, vol. 25, no.2, pp.14-19, March 2011.

[12] ITU-T Rec. G.100/P.10 Amendment 1, 2007, "New appendix I - Definition of quality of experience (QoE)".

[13] M. G. Martini, C. Chen, Z. Chen, et al., "Guest editorial - QoE-aware wireless multimedia systems", *IEEE J. Sel. Areas Commun.*, vol. 30, no.7, pp. 1153-1156, Aug. 2012.

[14] B. Fu, D. Staehle, G. Kunzmann et al., "QoE-aware priority marking and traffic management for H.264/SVC-based mobile video delivery", in *Proc. of the 8th ACM workshop on Performance monitoring and measurement of heterogeneous wireless and wired networks*, Oct. 2013, pp. 173-180.

[15] O. Habachi, Y. Hu, M. van der Schaar et al., "MOS-Based Congestion Control for Conversational Services in Wireless Environments", *IEEE J. Sel. Areas Commun.*, Vol. 30, No. 7, August 2012.

[16] A. Khan, L. Sun and E. Ifeachor, "QoE Prediction Model and its Application in Video Quality Adaptation over UMTS Networks", *IEEE Trans. Multimedia*, vol.14, no.2, pp.431-442, April 2012.

[17] C.T.E.R. Hewage and M.G. Martini, "Quality of experience for 3D video streaming", *IEEE Commun. Mag.*, vol. 51, no. 5, pp. 101 - 107, May 2013.

[18] T. Kim, J. Kang, S. Lee and A.C. Bovik, "Multimodal Interactive Continuous Scoring of Subjective 3D Video Quality of Experience", *IEEE Trans. Multimedia*, vol. 16, no. 2, pp.387-402, Feb. 2014.

[19] Q. Huynh-Thu, P. Le Callet, and M. Barkowsky, "Video quality assessment: From 2D to 3D – Challenges and future trends", in *Proc. of 17th IEEE International Conference on Image Processing (ICIP)*, Sep. 2010, pp.4025-4028.

[20] K. Wang, M. Barkowsky, Kjell Brunnström, et al., "Perceived 3D TV transmission quality assessment: multi-laboratory results using absolute category rating on quality of experience scale", *IEEE Trans. Broadcast.*, vol. 58, no. 4, pp.544-557, April 2012.

[21] C. T. E. R. Hewage, S. T. Worrall, S. Dogan at al., "Quality evaluation of color plus depth map-based stereoscopic video." *IEEE J. Sel. Topics Signal Process.*, vol. 3, no. 2, pp. 304-318, April 2009.

[22] P. Lebreton, A. Raake, M. Barkowsky, P. Le Callet, "Evaluating depth perception of 3D stereoscopic videos", *IEEE J. Sel. Topics Signal Process.*, vol.6, no.6, pp.710-720, Aug. 2012.

[23] S.L.P. Yasakethu, C.T.E.R. Hewage, W. A. C. Fernando, A. M. Kondoz, "Quality analysis for 3D video using 2D video quality models", *IEEE Trans. Consum. Electron.*, vol. 54, no. 4, pp. 1969-1976, Nov. 2008.

[24] Y. Zhao, L. Yu, "A perceptual metric for evaluating quality of synthesized sequences in 3DV system." in *Proc. Visual Communications and Image Processing* 2010.

[25] C.T.E.R. Hewage and M.G. Martini, "Edge-based reduced-reference quality metric for 3-D video compression and transmission", *IEEE J. Sel. Topics Signal Process.*, vol.6, no.5, pp.471-482, 2012.

[26] C.T.E.R. Hewage and M.G. Martini, "Reduced-reference quality assessment for 3D video compression and transmission", *IEEE Trans. Consum. Electron.*, vol.57, no.3, pp.1185-1193, 2011.

[27] I. Politis, L. Dounis, C. Tselios et al., "A model of network related QoE for 3D Video", in *Proc. of IEEE Globecom Workshops*, Dec. 2012, pp.1335-1340.

[28] N. McKeown, T. Anderson, H. Balakrishnan et al., "OpenFlow: Enabling innovation in campus networks", *SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 2, pp. 69-74, April 2008.

[29] S. Agarwal, M. Kodialam, T. V. Lakshman, "Traffic Engineering in Software Defined Networks", in *Proc. of IEEE Infocom*, pp.2211-2219, April 2013.

[30] E. Maratsolas, P. Koutsakis and A. Lazaris, "Activity-Based Video Traffic Policing: A New Paradigm", *IEEE Trans. Multimedia*, vol. 16, no. 5, pp. 1446-1459, August 2014.

[31] A. Arefin, R. Rivas, R. Tabassum and K. Nahrstedt, "OpenSession: SDN-based Cross-layer Multi-stream Management Protocol for 3D Teleimmersion", in *Proc. of International Conference on Network Protocols*, 2013.

[32] H. E. Egilmez, S. Civanlar and A. M. Tekalp, "An optimization framework for QoS-enabled adaptive video streaming over OpenFlow networks", *IEEE Trans. Multimedia*, vol. 15, no. 3, pp.710-715, April 2013.

[33] E. Keller, S. Ghorbani, M. Caesar and J.Rexford, "Live migration of entire network (and its hosts)", in *Proc. of the 11th ACM Workshop on Hot Topics in Networks (HotNets)*, pp. 109-114, 2012.

[34] N. Xue, X. Chen, S. Li et al., "Demonstration of OpenFlow-Controlled Network Orchestration for Adaptive SVC Video Manycast", *IEEE Trans. Multimedia*, vol.17, no. 9, pp. 1617-1629, Sep. 2015.

[35] E. Amaldi, A. Capone, and L.G. Gianoli, "Energy-aware IP traffic engineering with shortest path routing", *Computer Networks*, vol. 57, no. 6, pp.1505-1517, April 2013.

[36] N. Vasić and D. Kostić, "Energy-aware traffic engineering", in *Proc. of 1st ACM Int. Conf. Energy-Efficient Comput. Netw.*, pp. 169-178, 2010.

[37] S. Oda, D. Nobayashi, Y. Fukuda, T. Ikenaga, "Flow-based Routing Schemes for Minimizing Network Energy Consumption using OpenFlow", in *Proc. of ENERGY*, 2014, pp.69-72.

[38] H. E. Egilmez and A. M. Tekalp, "Distributed QoS Architectures for Multimedia Streaming over Software Defined Networks", *IEEE Trans. Multimedia*, vol. 16, no. 6, pp. 1597-1609, Oct. 2014.

[39] W. John, and C. Meirosu, "Low-overhead packet loss and one-way delay measurements in Service Provider SDN", *Open Networking Summit(ONS 2014)*, March 2014, pp.1-2.

[40] N. McKeown, "CS244a: An Introduction to Computer Networks", [Online] Available: https://web.stanford.edu/class/cs244a/handouts/H13%20Error%20Control%202008.pdf. Accessed on: Oct. 8, 2017.

[41] Y. Andreopoulos, N. Mastronarde and M. Van Der Schaar, "Cross-layer optimized video streaming over wireless multihop mesh networks", *IEEE J. Sel. Areas Commun.*, vol. 24, no. 11, pp.2104-2115, Nov. 2006.

[42] D. P. Bertsekas, R. Gallager, Data Networks, *Prentice Hall*, 1992.

[43] X. Zhu, E. Setton, and B. Girod, "Congestion-distortion optimized video transmission over ad hoc networks", *EURASIP Signal Process.: Image Commun.*, vol. 20, no. 8, pp. 773-783, Sep. 2005.

[44] Y. Liu, S. Ci, H. Tang et al., "QoE-oriented 3D Video Transcoding for mobile streaming", *ACM Transactions on Multimedia Computing, Communications and Applications (ACM TOMM)*, vol. 8, no. 3s, article 42, Sep. 2012.

[45] W. J. Tam, F. Speranza, S. Yano et al., "Stereoscopic 3D-TV: Visual comfort", *IEEE Trans. Broadcast.*, vol. 57, no. 2, pp. 335-346, Apr. 2011.

[46] M. Urvoy, M. Barkowsky, P. Le Callet, "How visual fatigue and discomfort impact 3D-TV quality of experience: a comprehensive review of technological, psychophysical, and psychological factors", *Annales des Télécommunications*, vol. 68, no.11-12 , pp. 641-655, 2013.

[47] M.-J. Chen, D.-K. Su, C.-C. Kwon et al., "Full-reference quality assessment of stereopairs accounting for rivalry", in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Nov. 2012, pp. 1-5.

[48] W. Chen, J. Fournier, M. Barkowsky, P.L. Callet, "Quality of experience model for 3DTV", *SPIE Stereoscopic Displays and Applications XXIII*, Feb. 2012, 82881P.

[49] Z. Wang and G. J. Klir, Fuzzy Measure Theory, New York, *Springer*. pp. 39-40, 1992.

[50] Z. Wang, K.-S. Leung, and G. J. Klir, "Applying fuzzy measures and nonlinear integrals in data mining", *Fuzzy Sets and Systems*, vol. 156, no.3, pp. 371-380, 2005.

[51] Y. Liu, J. Liu, Z. Xu, S. Ci, "Choquet Integral based QoS-to-QoE Mapping for Mobile VoD Applications", in *Proc. IEEE/ACM International Symposium on Quality of Service (IWQoS)* 2016, June 20-21, Beijing, China.

[52] Y. Liu, J. Liu, S. Ci, Y. Ye, "Joint video/depth/FEC rate allocation with considering 3D visual saliency for scalable 3D video streaming", in *Proc. of IEEE Visual Communications and Image Processing*, Dec. 2013, pp.1-6.

[53] ISO/IEC JTC1/SC29/WG11, "View synthesis reference software, March 2010, version 3.5.

[54] ITU-R BT.500-11, "Methodology for the Subjective Assessment of the Quality of Television Pictures".

[55] ITU-R BT.2021, "Subjective Methods for the Assessment of Stereoscopic 3DTV Systems," *International Telecommunication Union,* August 2012.

[56] VQEG,"Final report from the video quality experts group on the validation of objective models of video quality assessment, Phase I", April 2000.

[57] H. Malekmohamadi, W. A. C. Fernando, A. M. Kondoz, "Automatic QoE Prediction in Stereoscopic Videos", in *Proc. of IEEE International Conference on Multimedia and Expo Workshops*, pp. 581-586, July 2012.

[58] K. Gomez, D. Boru, R. Riggio et al., "Measurement-based modelling of power consumption at wireless access network gateways", *Computer Networks*, vol. 56, no. 10, pp. 2506-2521, July 2012.

[59] [Online] Available: http://mininet.org/. Accessed on: May 20, 2017.

[60] M. Zukerman, T. D. Neame, and R. G. Addie, "Internet traffic modeling and future technology implications", in *Proc. of the 22nd Annual Joint Conference of the IEEE Computer and Communications (INFOCOM)*, 2003, pp.587-596.

[61] J. Chabarek, J. Sommers, P. Barford et al., "Power awareness in network design and routing", in *Proc. IEEE Infocom*, 2008, pp. 457-465.

[62] [Online] Available: http://www.fujii.nuee.nagoya-u.ac.jp/multiview-data/. Accessed on: June 8, 2017.

[63] [Online] Available: http://sourceforge.net/projects/opensvcdecoder/. Accessed on: March 8, 2017.

Yanwei Liu (M'14) received the B.S. degree in applied geophysics from Jianghan Petroleum University, China, in 1998, the M.S. degree in computer science from China Petroleum University (Beijing) in 2004 and the Ph.D. degree in computer science from Institute of Computing Technology, Chinese Academy of Sciences in 2010. Currently, he is an Associate Professor with the Institute of Information Engineering, Chinese Academy of Sciences. His research interests include digital image/video processing, multiview/3D video/VR video processing, multimedia communication and wireless video communication. He has published over 50 scientific papers. He serves on the TPCs of several international conferences in the area of multimedia, communications, and networking.

Jinxia Liu received the B.S. degree and M.S. degree in physics from Harbin Normal University, China, in 1994, and 2005, respectively. In 2005, she joined the Zhejiang Wanli University. Currently, she is a Professor with Zhejiang Wanli University. Her research interests include laser imaging, digital image/video processing, multiview and 3D video coding, and wireless communication.

Antonios Argyriou (S'00-M'07-SM'15) received the Diploma degree in electrical and computer engineering from Democritus University of Thrace, Xanthi, Greece, and the M.S. and Ph.D. degrees in electrical and computer engineering (as a Fulbright Scholar) from Georgia Institute of Technology, Atlanta, GA, USA, in 2001, 2003, and 2005, respectively. Currently, he is an Assistant Professor with the Department of Electrical and Computer Engineering, University of Thessaly, Volos, Greece. From 2007 to 2010, he was a Senior Research Scientist with Philips Research, Eindhoven, The Netherlands. From 2004 to 2005, he was a Senior Engineer with Soft.Networks, Atlanta, GA, USA. He currently serves on the Editorial Board of the Journal of Communications. His research interests include wireless communication systems and networks, and video delivery. He has also served as Guest Editor for the IEEE TRANSACTIONS ON MULTIMEDIA Special Issue on Quality-Driven Cross-Layer Design, and he was also a Lead Guest Editor for the Journal of Communications, Special Issue on Network Coding and Applications. He serves on the TPCs of several international conferences and workshops in the area of communications, networking, and statistical signal processing.

Song Ci (S'98–M'02–SM'06)received the B.S. degree from the Shandong University of Technology (now Shandong University), Jinan, China, in 1992, the M.S. degree from the Chinese Academy of Sciences, Beijing, China, in 1998, and the Ph.D. degree from the University of Nebraska-Lincoln, Lincoln, NE, USA, in 2002, all in electrical engineering. He is currently an Associate Professor with the Department of Electrical and Computer Engineering at the University of Nebraska-Lincoln. His current research interests include dynamic complex system modeling and optimization, content-aware quality-driven cross-layer optimized multimedia over wireless, Energy Internet and green computing. Dr. Ci serves as Editor, Guest Editor or Associate Editor in editorial boards of many journals, including the IEEE TRANSACTIONS ON CSVT, the IEEE TRANSACTIONS ON MULTIMEDIA, the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and the IEEE ACCESS. He also serves as the TPC member for numerous conferences. He is a senior member of IEEE and a member of ACM and AAAS.