Adaptive Multiview Video Streaming: Challenges and Opportunities

Jacob Chakareski, Ecole Polytechnique Fédérale de Lausanne

ABSTRACT

Delivering multiview video content over present packet networks poses multiple challenges. First, the best effort nature of the Internet exposes media packets to variable bandwidth, loss, and delay as they traverse the network. Second, the prediction dependencies employed to maximize compression efficiency make the reconstruction process at the client extremely vulnerable to missing data. Third, the heterogeneity of client devices in terms of computing power, display capabilities, and access link capacity necessitates customizing the streaming process per user. My article reviews existing opportunities for addressing these challenges from within each of the three main stages of the content delivery pipeline (i.e., encoding, transmission, and reconstruction). Concretely, I first describe adaptive source coding techniques that construct a compressed representation of the multiview video source that exhibits resilience to network bandwidth variations and client view selection uncertainty. Then I discuss intelligent methods for error protection, caching, and packet scheduling that organize the transmission of multiview data in a bandwidth-effective way. Here, I also review prospective multipath and cloud-assisted techniques for multiview video streaming. Finally, I identify robust client-side content reconstruction schemes and adaptive media playout methods that can minimize the impact of missing data and enhance the user's interactive experience. Then I proceed to describe community-driven streaming techniques for delivering interactive multiview content over a population of social peers. The article concludes with an outline of approaches for synergistic exploitation of the techniques I will present theretofore, jointly across the different layers of the network protocol stack at which they individually operate. Here, I also highlight the main deployment challenges for some of these techniques, and how their design should be addressed accordingly, to overcome them.

INTRODUCTION

The emergence of content captured simultaneously from multiple camera viewpoints, as illustrated in Fig. 1, has made possible a number of novel applications that are increasingly gaining attention. They include 3D and free-view point TV [1], immersive telepresence, and gaming and virtual worlds. In all of them, a critical component of the system is the timely delivery of the multiview content to the end user.

Unfortunately, the development of appropriate networking technologies that can meet this goal has not kept pace. Therefore, even today, video streaming is predominantly carried out over best effort networks, such as the Internet. The lack of quality of service (QoS) guarantees thereof means that the transmitted content will experience dynamic variations in data rate and congestion-induced effects, such as packet erasures and queuing delay, along its network path. This is particularly consequential in the context of multiview content due to its multi-fold bandwidth and complexity expansion, relative to its monoscopic (single camera) cousin. As the introduction of QoS networking is not forthcoming anytime soon, it is therefore required to address these relevant shortcomings from within the present multiview streaming architectures and the prospects they offer to this end. Key to many of the techniques we describe is their use of the spatial (inter-camera) correlation of the multiview video source in different contexts. This is in contrast to conventional multi-stream scenarios (e.g., scalable coding or multiple descriptions), where only temporal dependency exists among the data units comprising the content.

MULTIVIEW VIDEO STREAMING SYSTEM

The main components of a multiview video streaming system are illustrated in Fig. 2. The content is first captured via multiple cameras and then encoded into a collection of data units that are placed into the transmission buffer of the server. The packetized content is transmitted over the network to the client, where it is buffered initially. After a given period of time, called the playback delay, the client starts decoding the content data from the buffer in order to display it. The decoder buffer in conjunction with the playback delay allow the client to compensate for network latency variations and recover lost packets via acknowledgments (ACKs). Furthermore, the client may employ the backward channel to the server to request view

The research of the author is supported by the Swiss NSF Ambizione Career Award PZ00P2-126416.

0163-6804/13/$25.00 © 2013 IEEE

switching periodically in the case of interactive streaming applications.

We denote as *non-interactive* streaming the scenario where the complete ensemble of encoded views is delivered to the client. Only afterward can the client choose to interact with the content in terms of selecting a viewpoint from which the 3D scene will be observed. The client cannot switch to another viewpoint until the next segment of self-contained view data is received.

The alternative of delivering select views exclusively based on the client's dynamic viewpoint trajectory is denoted as *interactive* streaming henceforth. The former approach is analogous to cable TV broadcast, where the entire set of channels is delivered to the client continuously. The latter bears parallels to IPTV where only upon a channel switching action is the actual content of the requested TV channel delivered. The disadvantage of *non-interactive* streaming is that it is much more demanding in terms of bandwidth and complexity. The disadvantage of interactive streaming is the additional latency that is introduced into the system by the virtue of delivering select view content only upon request, which may impair the interactivity of the streaming application. I discuss throughout the remainder of the article, how the specifics of each streaming architecture allow for improvements in system performance.

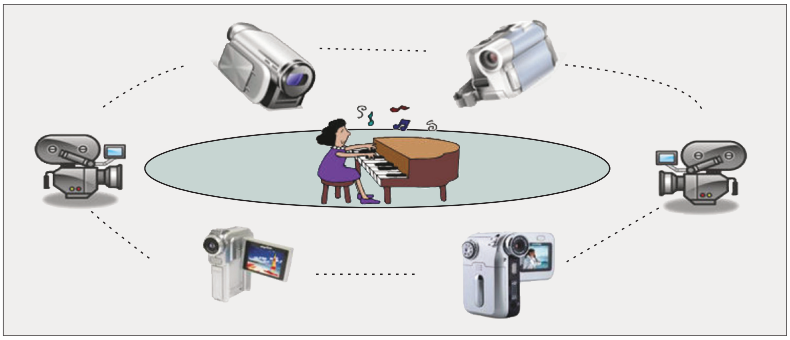

Figure 1. *Multi-camera system recording a dynamic 3D scene of interest.*

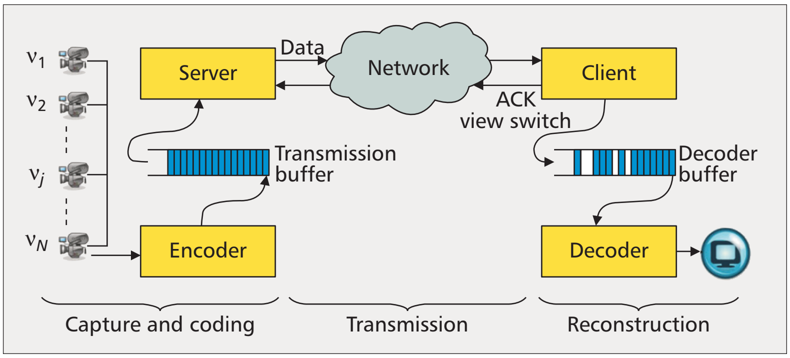

Figure 2. *Multiview content delivery: system architecture.*

Multiview Content Representation

There are two prevailing forms in which packetized multiview content can be encountered today. Historically, the multiview video coding (MVC) format appeared first and stems from the corresponding multiview extension of the H.264 encoding standard [2]. MVC data features inter- and intra-view prediction dependencies, as illustrated in Fig. 3a, which have been employed by the codec in order to maximize the compression efficiency. Streaming MVC content is generally done under the assumption of unknown client view selection actions, as it requires transmitting the complete set of captured views throughout the media session. Therefore, its natural application would be in broadcast delivery of content, although studies have explored modifying MVC content dynamically, in response to client feedback, in order to reduce its required transmission rate. The more recent "video plus depth" (VpD) multiview format [3] generally employs intra-view prediction only; however, it also encodes a depth signal for each camera, as illustrated in Fig. 3b. Such signals can be either estimated from the corresponding video signals using stereo-matching algorithms or captured directly using time-of-flight cameras.

The VpD representation is increasingly becoming popular, as it does not necessitate a network bandwidth expansion on the order of MVC, since only the data associated with the view currently being watched is transmitted. However, consistent client feedback to the server is required in order to track the user's view switching pattern over time. Another benefit of the VpD representation is that it enables synthesizing virtual views at the client via the depth signals that are sent along [4]. However, then additional data needs to be transmitted to account for the uncovered image regions (a.k.a. disocclusions) that are created in the synthesized view. Alternatively, the client may employ in-painting and interpolation methods [5] to recover the missing pixels.

Source Coding Opportunities

Tackling the problem of bandwidth variability can be achieved via scalable multiview encoding. Then, during streaming, a sufficient number of layers can be transmitted such that the available channel capacity is fully utilized. An alternative is to dynamically encode the content, on the fly, in response to network rate variations. Although this solution is appropriate for live streaming and real-time multiview communication, it is not appealing for stored content that has already been encoded, since it necessitates transcoding. That is, the content needs to be decoded and re-encoded on demand, which will augment the complexity of the system considerably. To date, source rate adaptation has been considered only in the context of stereoscopic (two-view) MVC content where left and right views are simultaneously presented to the user in order to create a perception of depth of the observed scene. These studies (e.g., [6]) take advantage of the suppression theory of human stereo perception [7] to dramatically reduce the data rate of one of the views, without causing perceptual changes in the overall (3D) video quality at the client.

For VpD content, the scalability paradigm can be extended to include the view dimension as well. That is, in addition to encoding rate

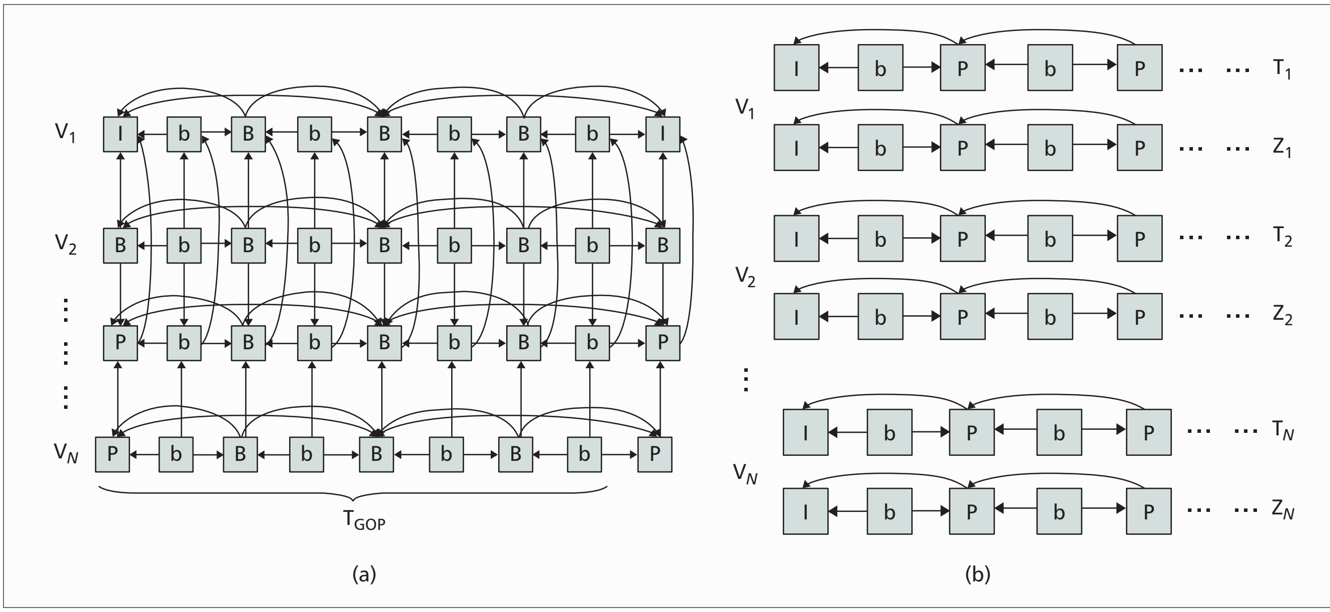

Figure 3. *Multiview content representation formats. Squares denote video frames (data units), arrows their encoding dependencies: a) MVC; b) video (T) plus depth (Z).*

adaption over the complete set of views, view omission from the compressed bitstream may also be considered as part of the scalable coding trade-off. An embedded multiview representation that is thereby created features a variable number of encoded views, depending on the decoding rate at which the content is reconstructed. Depth-signal-based view synthesis can then be used by the client to recover the views that have been omitted at that encoding rate. The goal then is to build a joint view and rate scalable encoding that maximizes the decoding (reconstruction) quality of any viewpoint at the client, over a range of prospective network rate values that may be encountered at transmission [8].

In departure from conventional MVC group of picture (GOP) encoding structures, the rate distortion efficiency of stored multiview content can be enhanced if distributed source coding (DSC) principles are employed. Specifically, another (DSC) frame type is considered at encoding that enables subsequent view switching, like conventional I-frames do, but at a lower penalty to coding efficiency [9]. Alternatively, one may consider encoding subsets of views, offline and independently, at multiple quality levels, using MVC. Then, at streaming, a suitable subset of views and quality layers is sent to match the present head position of the user, which is tracked over time [10]. This approach may also be applied online in interactive streaming, where based on the user's prior view selection actions, prospective video frames and views that the user is anticipated to subsequently select over a horizon of time are exclusively encoded and transmitted [11].

Finally, dynamic control of coding dependencies can be employed to overcome adverse channel effects such as packet loss. In particular, the inter- and intra-view prediction dependencies employed to encode the content are modified on the fly, depending on the ACK status of previously transmitted data units. This approach can be more beneficial in the case of non-interactive streaming due to the generally longer buffering employed at the client, before the data is actually passed to the decoder for reconstruction and display. The lower playout delay of interactive streaming limits the application of this approach in such a scenario, as the received data needs to be decoded quickly in order not to impair the interactive experience of the client. That is, the shortcomings of this method are longer decoding latency and lower compression efficiency, since the prediction dependencies can sometimes extend over multiple video frames.

Channel Transmission Opportunities

Error Correction and Packet Scheduling

The hierarchical dependencies of the MVC representation allow for capturing the importance of each data unit for the overall reconstruction quality of the multiview content. Thus, unequal error protection can be applied via forward error correction (FEC) coding across the data units of various views, to compensate for packet loss during transmission and maximize simultaneously the end-to-end video quality. Likewise, multiple description coding (MDC) can be applied to the content as an alternative, where the content descriptions are created such that the unequal importance of the data is again exploited. Conceptually, FEC and MDC can be applied in the same manner to the MVC view-level encoding hierarchy, which from an implementation perspective may be easier to accomplish than their packet-level counterparts.

The unequal importance of data units across views can also be exploited within an intelligent packet (re)transmission scheme that will utilize such knowledge to allocate the limited network resources accordingly. In particular, given the size of each data unit in bytes, rate-distortion optimized packet scheduling can be envisioned that maximizes the reconstruction quality of the content for the given channel capacity. Furthermore, analogous caching schemes can be designed at intermediate edge servers in order to effectively disseminate the content to end clients for the given caching resources. Designing efficient multiview video frame coding structures based on DSC principles can also facilitate the application of decentralized caching to interactive streaming. Here, the goal is to minimize the quantity of data that is sent from the central server to one of the caches, upon a cache miss by a client that interactively experiences the multiview content. In such a case, a correlated view is served instead from the local cache, and only the coded difference relative to the desired viewpoint is requested from the central server [12].

In the case of non-interactive streaming of VpD content, there is no natural inter-view hierarchy that arises among the data units comprising the content that can be exploited within the techniques described here. Still, such information can be implicitly deduced from the rate-distortion efficiency at which a view can be used to reconstruct other views via depth-signal-based synthesis. Then, the same effects of error resilience and enhanced visual quality can be achieved equally, on the basis of this knowledge.

Multiview Multicast

In the case of non-interactive streaming, multiple unicast sessions delivered over a shared path can be replaced by a single multicast one. This will enable more efficient utilization of the network bandwidth. Coupled with scalable encoding of the content, this method of content delivery can efficiently serve a heterogeneous client population, characterized by diverse access link and computing/display capability profiles. Such system architectures can also be equipped with efficient FEC methods (e.g., Reed-Solomon codes), which can take advantage of the scalable multiview representation to enhance the performance of the system even further, as described earlier. IP layer multicast is one technology that can be applied to deliver the content in this context.

The case of interactive multiview streaming needs to be addressed differently. In particular, each view is encoded and broadcast independently on a separate channel. Users switch views by simply subscribing to another multicast channel, while leaving their present one. The view switching frequency is governed by the intra coding period of the corresponding video signals. Conceptually, the operation of this system is analogous to that of IPTV. Additional interactivity features (e.g., watching a frozen moment in time across all views) can be enabled by additional multicast channels over which the data that is necessary to support them is broadcast. Depending on the desired level of interactivity, a user subscribes to one or multiple channels simultaneously, at any given time [13].

Interactive streaming can also be considered in the case of peer-to-peer (P2P) networks where each viewpoint can be associated with one dissemination tree rooted at the respective video source. In particular, here the collection of cameras capturing the content does not necessarily have to be associated with a single (server) entity. That is, different viewpoints of the same scene may be recorded independently and in a decentralized fashion by separate client devices that then share the content with their peers through application layer multicast. Alternatively, multi-mesh network P2P content delivery can be utilized as a technology for disseminating the multiview content across the client population, where each mesh a peer joins is used to distribute one camera viewpoint.

Multipath Delivery

Multipath transmission can assist multiview streaming, by providing higher bandwidth and resilience to network transients such as burst packet loss. The advent of multihoming can facilitate the deployment of multipath streaming applications. Mapping data units to network paths can be done according to the importance of each packet for the overall reconstruction quality. In particular, more important data units should be allocated to higher-quality paths, jointly in decreasing importance and quality order, until either all the data is scheduled for transmission or the aggregate multipath network capacity is reached. The same procedure applies in the case of view-level transmission decisions. There is no direct relation between the number of network paths that can be used and the number of captured viewpoints. Still, in the case of interactive streaming of VpD content, where the client is interested in synthesizing virtual viewpoints, sending the required data over two paths seems like a natural choice. That is because in such a case the video and depth signals of the two nearest captured views are generally required in order to ensure a good quality synthesized view. The application of multiple description coding to delivering free-viewpoint multiview content over burst loss network paths has been studied in [14]. Promising gains in transmission efficiency have been demonstrated over conventional single and multipath solutions that do not account for the content's specifics.

Multipath content delivery is particularly intuitive in the context of wireless networks and P2P systems, where naturally multiple multihop possibilities of reaching a destination over intermediate neighbors arise. In addition, it is quite common today to have multiple wireless adapters on a single device, such as third or fourth generation (3G or 4G) cellular and 802.11n WLAN, which can be utilized in parallel. This will add another degree of path diversity, provided by the existence of multiple orthogonal mesh networks. Typically, content delivery over wireless multihop networks is combined with advanced channel coding schemes such as network coding or rateless codes that can increase network throughput and error resilience for the data. In particular, scalable network coding applied to a scalable multiview representation can additionally provide graceful degradation in terms of video quality, as network bandwidth becomes insufficient. Similarly, conventional network coding exploiting the unequal importance of each data unit will enable the same phenomenon in the case of general (non-layered) source encoding. For two

Multipath transmission can assist multiview streaming by providing higher bandwidth and resilience to network transients such as burst packet loss. The advent of multihoming can facilitate the deployment of multipath streaming applications.

orthogonal wireless topologies spanning the same set of nodes, adaptive in-network compression via MDC for interactive streaming [15] can equally boost the rate-distortion efficiency of the multiview application.

In the context of general overlay networks, multiview clients can construct multiple application layer mesh topologies over an existing data network such as the Internet. Then, multimesh P2P content delivery can be utilized as a technology to disseminate the multiview content across the client population, where each mesh a peer joins is used to distribute one camera viewpoint. As in the context of caching, performance can be enhanced, if network rate allocation and topology construction are jointly carried out over the multi-mesh collection, taking into account the inter-view and intra-view video signal correlation.

Dynamic routing of multiview data in multihop networks can be carried out such that it takes into account the characteristics of other types of traffic sharing the same network fiber, in the case of wireless multi-service networks. In particular, routing decisions of voice, video, and data can be computed at intermediate nodes in a decentralized fashion such that the overall (end-to-end) performance of the network is maximized [16]. This is done by dynamically adapting the optimal routes for the delivery of the data associated with each source-destination pair and traffic class, in response to varying network conditions such as topology, flow rates, and battery power of the nodes. Simultaneously, the computed routes take into account the specific characteristics of each service class, in terms of resilience to delivery latency, data loss, and bandwidth variation, such that a judicious trade-off of the network's resources is achieved over them. In addition, routing decisions over the multiview data are carried out such that they exploit the inter- and intra-view video signal correlation that governs the decoding quality of the content, and the effectiveness with which one view can be exploited to reconstruct another.

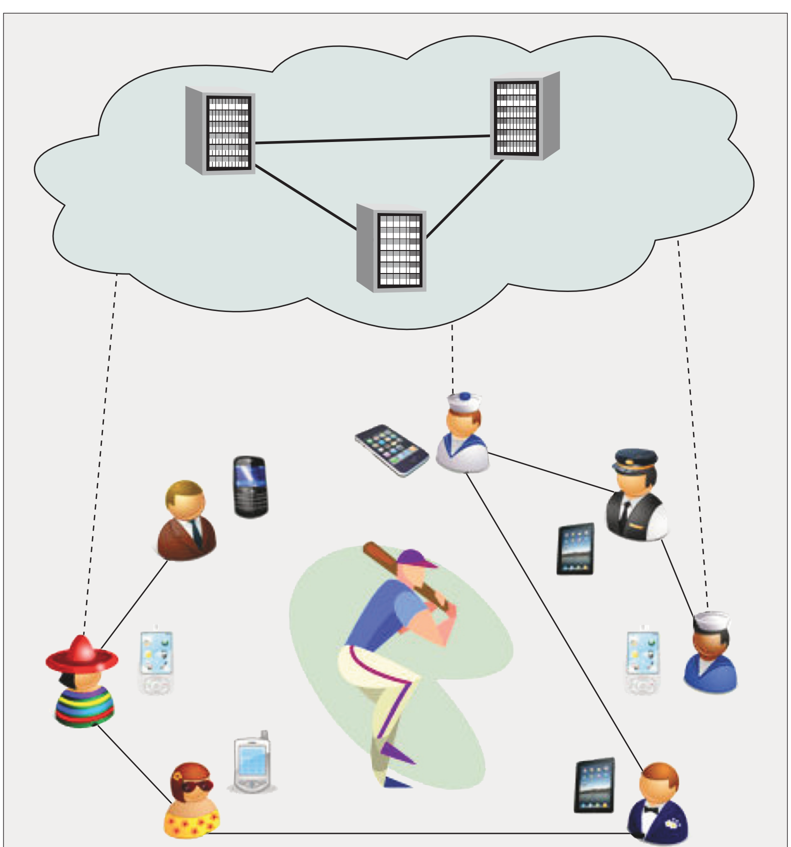

Figure 4. *Cloud-assisted live multiview streaming over mobile devices. Peer and cloud connections are denoted by solid and dashed lines, respectively.*

Cloud-Assisted Decentralized Streaming

The increasing popularity of smart portable devices, and their remarkable computing and networking capabilities open a range of interesting prospects for the development of mobile 3D video applications. For instance, a spontaneous congregation of client devices can happen to independently record video content of a sporting event from different perspectives, at the same time. The clients may be interested in sharing their respective video feeds among themselves, for an enhanced visual experience of the event. Simultaneously, they may also be interested in experiencing the 3D scene from another angle, with a perception of depth. Due to the limited battery lifetime and computing capacity of their devices, the clients may not be able to execute these tasks by themselves. However, they can leverage the availability of a cloud computing platform, to offload the necessary compute-intensive operations such as view synthesis and transcoding on it. In addition, the clients can dynamically construct local network connections among themselves on an as-needed basis, in order to further facilitate the dissemination of the content, as illustrated in Fig. 4.

The techniques for prioritized rate allocation, packet scheduling, and caching that we described earlier, can all equally be applied within the backbone network employed by the cloud operator. Although its operating costs are certainly much lower relative to the rest of the network links through which the actual content is deployed, the cloud network can still benefit in terms of efficiency and end-to-end performance if it takes advantage of such methods. This will be particularly true for large-scale cloud-assisted multiview streaming scenarios, where the client population can be exceedingly large in magnitude.

Client Side Opportunities

In the case of non-interactive streaming, error concealment techniques that jointly take advantage of the inter-view and intra-view signal correlation, across the complete ensemble of captured views, can be designed. For instance, missing data can be reconstructed by simultaneous interpolation from multiple neighboring video frames via temporal and spatial displacement estimation, as illustrated in Fig. 5. Concretely, optimization techniques can be designed that reconcile the obtained multiple predictors of the missing data in an efficient way. In the presence of VpD content, the computation of inter-view predictors via disparity estimation among the views can be conveniently replaced with depth-signal-based view synthesis. For interactive streaming with virtual view synthesis, the number of prospective neighboring view video signals that can be utilized for spatial error concealment will be limited to one [14], at most.

Adapting the playout speed of the media session (the temporal rate at which newly decoded content is displayed to the client) can also be utilized to reduce the impact of missing data. This method can be exploited in conjunction with proactive data fetching techniques that dynamically instruct the server of the present state of the client buffer and viewing angle, such that the sender can adjust accordingly its transmission rate, and coding and packet scheduling decisions. The adaptive playout takes advantage of the certain degree of audiovisual tolerance that we have to vary the refresh rate of the media presentation at the client, depending on the content's spatiotemporal dynamics. The inter-view correlation of the multiview content can be utilized to enhance the performance of this approach, by allowing smoother view switching and higher minimum playout speeds during intra-view consumption of the content.

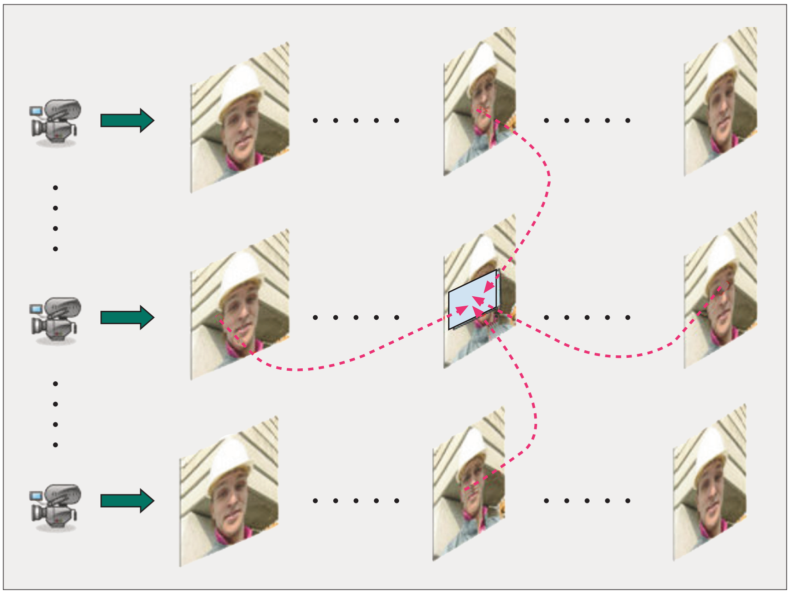

Figure 5. *Replacement of missing data via space-time concealment.*

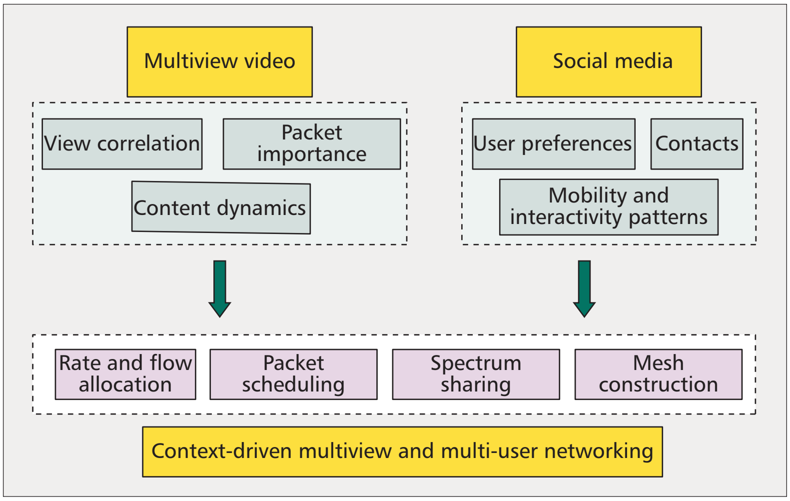

Figure 6. *Community-driven interactive multiview video communications.*

Social Multiview Video Interaction

The ubiquity of social networking applications online, and the widespread availability and facility of use of high-quality content creation tools provide a range of interesting opportunities for synergistic community-driven delivery of interactive user-generated 3D content. For instance, by tracking the content interaction patterns of the users, coding resources can be allocated such that the overall multiview video quality performance of the system is maximized [17]. Similarly, by exploiting the user's content preferences and mobility patterns, more efficient utilization of network resources can be achieved [18]. Comprehensively speaking, a variety of computer communications techniques spanning, for example, rate and flow allocation, and network topology construction can be designed to take into account the accelerating convergence of digital content and spatiotemporal user interaction, witnessed at present in the Internet. In particular, they will exploit in their operation the characteristics of the multiview content in conjunction with the specifics of the online community, as illustrated in Fig. 6, to deliver enhanced performance over multiple criteria. Furthermore, such networking techniques can pave the way for the development of novel human communication techniques that can revolutionize the ways in which we interact and collaborate at present.

Hybrid Approaches and Challenges

Some of the streaming techniques described heretofore can be applied synergistically. For instance, dynamic source rate control of multiview content can be combined with multipath transmission and routing in multihop networks. In particular, in addition to computing routing decision adaptively, based on the network's dynamics and the content's characteristics, intermediate nodes can also feed back information to the source reflecting the bandwidth availability and the path characteristics in terms of loss and latency of each route. The sender can then act upon this information accordingly, and adapt the data rate of the transmitted multiview content on a per route basis. Even more generally, control methods can be considered that turn source rate adaptation on and off, depending on the level of network dynamics, that is, the efficiency at which multiview multipath routing in the network can alone support the reliable delivery of the content. Routing and source rate control can also be combined with the multiview caching methods discussed earlier. Other cross-layer techniques that can be considered in this context include the combination of application layer packet scheduling with physical layer FEC, and joint context-driven source-rate adaptation, content caching, and physical layer network coding, *Information accuracy vs. reliable performance, and robustness to imprecise source and channel knowledge need to be considered and incorporated as part of the design of effective multiview streaming techniques that can be deployed in such a context.*

for enhanced network throughput and end-to-end video quality performance.

Many of the methods presented in this article base their operation on consistent source and/or channel information, such as unequal data unit importance, and available network bandwidth and routes. Their design must incorporate provisions for the cases when such information is not readily accessible or can only be partially acquired via estimation/approximation. This is particularly important in scenarios where a large-scale decentralized operation is considered. Therefore, information accuracy vs. reliable performance, and robustness to imprecise source and channel knowledge need to be considered and incorporated as part of the design of effective multiview streaming techniques that can be deployed in such a context.

References

[1] M. Tanimoto *et al.*, "Free-Viewpoint TV," *IEEE Sig. Proc. Mag.*, vol. 28, no. 1, Jan. 2011, pp. 67–76.

[2] A. Vetro, T. Wiegand, and G. J. Sullivan, "Overview of the Stereo and Multiview Video Coding Extensions of the H.264/MPEG-4 AVC Standard," *Proc. IEEE*, vol. 99, no. 4, Apr. 2011, pp. 626–642.

[3] P. Merkle *et al.*, "Multi-View Video Plus Depth Representation and Coding," *Proc. Int'l Conf. Image Processing*, vol. 1, San Antonio, TX, Sept. 2007, pp. 201–04.

[4] P. Merkle *et al.*, "The Effects of Multiview Depth Video Compression on Multiview Rendering," *Sig. Processing: Image Commun.*, vol. 24, no. 1–2, Jan. 2009, pp. 73–88.

[5] Z. Tauber, Z.-N. Li, and M. S. Drew, "Review and preview: Disocclusion by Inpainting for Image-based Rendering," *IEEE Trans. Sys., Man, and Cybernetics – Part C*, vol. 37, no. 4, July 2007, pp. 527–40.

[6] A. Aksay *et al.*, "End-to-End Stereoscopic Video Streaming with Content-Adaptive Rate and Format Control," *Sig. Processing: Image Commun.*, vol. 22, no. 2, Feb. 2007, pp. 157–68.

[7] L. Stelmach *et al.*, "Stereo Image Quality: Effects of Mixed Spatio-Temporal Resolution," *IEEE Trans. Circuits and Systems for Video Tech.*, vol. 10, no. 2, Mar. 2000, pp. 188–93.

[8] V. Velisavljevi c *et al.*, "View and Rate Scalable Multiview Image Coding with Depth-Image-based Rendering," *Proc. 17th Int'l Conf. Digital Signal Proc*., Corfu, Greece, July 2011.

[9] G. Cheung, A. Ortega, and N.-M. Cheung, "Interactive Streaming of Stored Multiview Video Using Redundant Frame Structures," *IEEE Trans. Image Processing*, vol. 20, no. 3, Mar. 2011, pp. 744–61.

[10] E. Kurutepe, M. R. Civanlar, and A. M. Tekalp, "Client-Driven Selective Streaming of Multiview Video for Interactive 3DTV," *IEEE Trans. Circuits and Sys. for Video Tech.*, vol. 17, no. 11, Nov. 2007, pp. 1558–65.

[11] Z. Pan *et al.*, "User Dependent Scheme for Multi-View Video Transmission," *Proc.IEEE ICC*, Kyoto, Japan, June 2011.

[12] H. Huang *et al.*, "Near-Optimal Content Replication for Interactive Multiview Video Streaming," *Proc. Int'l Packet Video Wksp*., Munich, Germany, May 2012, pp. 95–100.

[13] L. Zuo *et al.*, "Multicast of Real-Time Multi-View Video," *Proc. Int'l Conf. Multimedia and Exhibition*, Toronto, Canada, July 2006, pp. 1225–28.

[14] Z. Liu *et al.*, "Multiple Description Coding of Free Viewpoint Video for Multi-Path Network Streaming," *Proc. IEEE GLOBECOM*, Anaheim, CA, Dec. 2012.

[15] J. Sørensen *et al.*, "Multiple Description Coding with Feedback based Network Compression," *Proc. IEEE GLOBECOM*, Miami, FL, Dec. 2010.

[16] R. Matos *et al.*, "Quality of Experience based Routing in Multi-Service Wireless Mesh Networks," *Wksp. Realizing Advanced Video Optimized Wireless Networks at the Int'l Conf. Commun.*, Ottawa, Canada, June 2012.

[17] A. Fiandrotti, J. Chakareski, and P. Frossard, "Popularity-Based Rate Allocation in Multiview-Video," *Proc. Conf. Visual Commun. and Image Proc.*, Huang Shan, An Hui, China, July 2010.

[18] J. Chakareski and P. Frossard, "Context-Adaptive Information Flow Allocation and Media Delivery in Online Social Networks," *IEEE J. Selected Topics in Sig. Procesing*, vol. 4, no. 4, Aug. 2010, Special Issue on Signal and Information Processing for Social Networks, pp. 732–45.

Biography

Jacob Chakareski (jakov.cakareski@epfl.ch) completed his M.S. and Ph.D. degrees in electrical and computer engineering at Worcester Polytechnical Institute, Rice University, and Stanford University. He was a postdoctoral fellow at EPFL and a Research Scientist at Vidyo, a leading provider of Internet telepresence solutions. Presently, he is an Associate Scientist at Ecole Polytechnique Federale de Lausanne, where he conducts research, lectures, and supervises students. His research interests intersect immersive video, computer networks, and social computing. He eagerly pursues telemedicine, sensing, and biomedical applications of ultrasound. He is a member of Tau Beta Pi and Eta Kapa Nu. He is a recipient of fellowships from the Soros Foundation and from the Macedonian Ministry of Science. At Rice University, he was awarded a Texas Instruments Graduate Research Fellowship. He received the Swiss NSF Ambizione Career Development Award. He has held research positions with Microsoft and Hewlett-Packard. He has authored one monograph, three book chapters, and over 100 international publications, and has multiple pending or approved patent applications. He actively participates in technical and organizing committees of major IEEE conferences on a yearly basis. He was Publicity Chair for the Packet Video Workshop in 2007 and 2009, and for the Workshop on Emerging Technologies in Multimedia Communications and Networking at ICME 2009. He organized and chaired a special session on telemedicine at MMSP 2009. He was the Technical Program Chair of Packet Video 2012. He was the General Chair of the IEEE SPS Seasonal School on Social Media Processing in 2012. He is a Guest Editor of the Springer *PPNA Journal*'s Special Issue on P2P Cloud Systems. He won the best student paper award at the IS&T/SPIE VCIP 2004 conference. For more information, visit http://www.jakov.org.