3D Media over Future Internet: Current Status and Future Research Directions

Tasos Dagiuklas[1]
[1] **CONES Research Group, Dept. of Telecommunication Systems and Networks, TEI of Mesolonghi**
Nafpaktos, 30300, Greece

Abstract

Future Media Internet has been designed to overcome current limitations and address emerging trends including: network architecture, content and service delivery across heterogeneous networks, diffusion of heterogeneous nodes and devices, mass digitisation, new forms of (3D) user centric/user generated content provisioning, emergence of software as a service and interaction with improved security, trustworthiness and privacy. This paper presents current and future research trends for 3D Video Delivery across the entire networked-media ecosystem (from the encoding/packetisation, through the transmission and up to end-user experience).

Keywords: *3D Media Future Internet*

1. Introduction

The Internet was designed for purposes that bear little resemblance to today's usage scenarios and related traffic patterns. In the longer term, the exponential increase of the user-generated multimedia content and the number of mobile users will raise many new challenges. In this respect, Future Media Internet will not simply be a faster way to go online [1]. It will be designed to overcome current limitations and to address emerging trends including: network architecture, content and service mobility, diffusion of heterogeneous nodes and devices, mass digitisation, new forms of (3D) user centric/user generated content provisioning, emergence of software as a service and interaction with improved security, trustworthiness and privacy.

There are a lot of advances in 3D media technologies in terms of capturing, representation, coding, delivering, and visualization. 3D media has been evolving in various areas, covering various market segments such as professional (e.g. scientific, medicine, education, training etc) and entertainment (3D Gaming, 3D Live Sports, 3D Live Shows etc) sectors. These new technologies give the ability to design and develop new types of applications ranging from Virtual Collaborative Environments (e.g. multi-modal interactions, etc) to Edutainment (e.g, the user is allowed to choose a service in terms of content production, searching program content, provide service according to user’s preferences, tailor education content according to user’s need).

The successful deployment of 3D Media is based on issues related to efficient transport of content, as well as, effective and inexpensive 3-D displays. However, although individual technologies are evolving fast, their integration and effective use is still limited by numerous open issues. In fact, while the production of 3D films, movies and live broadcasting events started in the early 50s, the research community has only recently paid attention at the transmission/broadcasting of 3D (rather than conventional 2D) applications and services. In the same time, latest advances on broadband network access (e.g. wireless: 3G, 4G, LTE, WiMax etc) and core (optical) technologies and digital broadcasting (e.g. DVBx) in terms of capacity, speed, ubiquity and reliability necessitates the optimization of the delivery of the aforementioned new 3D applications. These applications pose unique problems in comparison to natural 2D media coding data.

The aim of this paper is to outline the challenges associated with 3D Media Delivery across heterogeneous networks. This includes media functionalities (in terms of coding efficiency/scalability, error resilience, packetisation schemes, streaming protocols, adaptation, rate control, etc), network architectures to deliver 3D Video, use of rate adaptation in order to maximize QoS under network/user constraints with guarantee 3D viewing, wireless 3D Video Delivery and se of and error concealment for 3D Video

The paper is structured as follows: Section II present the current status in terms 3D Video Coding Standards, Transport Protocols and Wireless Media Delivery, Section II outlines future research challenges, issues and directions, Section IV presents the conclusions.

2. Current Status in 3D Video Communications

2.1 3D Video Coding

Research on video coding techniques for 3D is an increasing research topic around the world. A few coding techniques exist, and the most important ones are: 2D+Depth, as specified by ISO/IEC 23002-3 (and also referred to as MPEG-C Part 3) [2], and Multiview Video Coding (MVC), as specified by ISO/IEC 14496-10 | ITU-T Recommendation H.264 [3]. 2D+ Depth supports the inclusion of depth for generation of an increased number of views. While it has the advantage of being backward compatible with legacy devices and is agnostic of coding formats, it is only capable of rendering a limited depth range since it does not directly handle occlusions. Depth information can be also included as a layer in Scalable Video Coding (SVC).

Multiview Video Coding (MVC) supports the direct coding of multiple views and exploits inter-camera redundancy to reduce the bit rate. MVC gives very good 3D rendering capability, but the bit-rate of MVC encoded video is proportional to the number of views [4]. Powerful algorithms and open international standards for MVC and coding of video plus depth data are available and under development, which will provide the basis for introduction of various 3DTV systems and services in the near future [5]. The research area is relatively young when compared to 2D video coding. Therefore there is a lot of room for improvement and development of new algorithms and coding methods.

1. Enabling stereo devices to cope with varying display types and sizes, and different viewing preferences. This includes the ability to vary the baseline distance for stereo video to adjust the depth perception, which could help to avoid fatigue and other viewing discomforts.
2. MPEG also envisions that high-quality auto-stereoscopic displays will enter the consumer market in the next few years. Since it is difficult to directly provide all the necessary views due to production and transmission constraints, a new format is needed to enable the generation of many high-quality views from a limited amount of input data, e.g. stereo and depth. The 3DV format is expected to have several advantages in terms of bit rate and 3D rendering capabilities.
    - 2D+Depth, as specified by ISO/IEC 23002-3, supports the inclusion of depth for increasing the number of view. It exhibits backward compatibility with legacy devices and is agnostic of coding formats. It is only capable of rendering a limited depth range since it does not directly handle occlusions. The 3DV format expects to enhance the 3D rendering capabilities beyond this format.
    - Multiview Video Coding (MVC), as specified by ISO/IEC 14496-10. ITU-T Recommendation H.264, supports the direct coding of multiple views and exploits inter-camera redundancy to reduce the bit rate. Although MVC is more efficient than simulcast, the rate of MVC encoded video is proportional to the number of views. The 3DV format expects to significantly reduce the bit rate needed to generate the required views at the receiver.
    - Multiview plus depth (MVD): It regards rendering techniques at the receiver providing great adaptation to varied depth experienced from different 3D Displays. This involves view synthesis by interpolating color information from multiple views. MVD can be extended to cope with large changes in view conditions by sending multiple MVD streams [6].
    - Layered depth Video (LDV): It is proposed in order to further reduce the bit rate from MVD. It projects the central camera view into other neighbouring views and then determine the difference between projection and neighbouring camera result as residual information [7], [8].

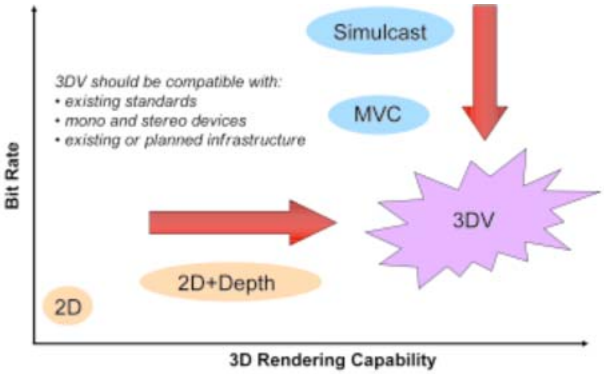

Figure 1: Towards 3DV Evolution (Source : http://multimediacommunication.blogspot.com)

2.2 Transport Protocols for 3D Video

IETF has adopted MVC as the 3D video coding standard to convey 3D video packets. MVC exhibits packetization similarities with H.264 SVC [9]. There are three transmission modes for multiple-view video, referring to Single Session Transmission (SST), Multi-Session Transmission (MST) and Media-Aware Network Element (MANE)-based transmission [10]. In SST, all

MVC packets are carried in a single RTP session utilizing only a single transport address/port. MANE generally resides in the path between the server and the clients. The server keeps using MST transmission. However in this case, MANE collects all RTP sessions and de-packetize them. It then customizes NAL Units according to the client's needs through adaptation decision taking engine (ADTE) and aggregates new packets for SST to clients through a single transport address using in a single RTP session. MANE is able to take important information using either signaling or RTP and NAL Unit Headers, as illustrated in the following figure.

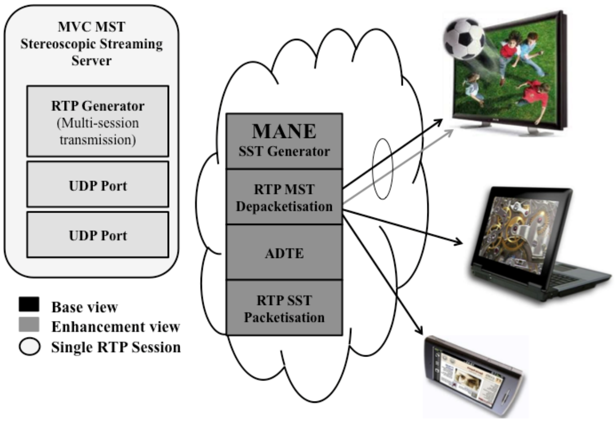

Figure 2: MVC Packetization Modes for 3D Video

The state-of-the-art video streaming protocols are RTP/UDP/IP, while the next generation protocol is expected to be RTP/DCCP/IP [11]. In particular, RTP/UDP does not contain any congestion control mechanism and, therefore, can lead to congestion collapse when large volumes of multi-view video are delivered. On the other hand, the datagram congestion control protocol (DCCP) is designed as a replacement for UDP for media delivery, running directly over the IP to provide congestion control without reliability [12]. DCCP is a transport protocol that implements bi-directional unicast connections of congestion-controlled, unreliable datagrams. DCCP provides reliable handshakes for connection setup/tear down and reliable negotiation of options. Besides handshakes and feature negotiation, DCCP also accommodates a choice of modular congestion control mechanisms. There exist two congestion control schemes defined in DCCP currently, one of which is to be selected at connection start-up time. These are TCP-like Congestion Control [13] and TCP-Friendly Rate Control (TFRC) [14]. TCP-like Congestion Control, identified by Congestion Control Identifier 2 (CCID2) in DCCP, behaves similar to TCP' s Additive Increase Multiplicative Decrease (AIMD) congestion control, halving the congestion window in response to a packet drop. Applications using this congestion control mechanism will respond quickly to changes in available bandwidth, but must tolerate the abrupt changes in the congestion window size typical of TCP. On the other hand, TFRC, which is identified by CCID3, is a form of equation-based flow control that minimizes abrupt changes in the sending rate while maintaining longer-term fairness with TCP. It is hence appropriate for applications that would prefer a rather smooth sending-rate, including streaming media applications with a small or moderate receiver buffer. In its operation, CCID3/TFRC calculates an allowed sending rate, called the TFRC rate, by using the TCP throughput equation, which is provided to the sender application upon request. The sender may use this rate information to adjust its transmission rate in order to get better results.

2.3 3D Video Networking Architectures

Transmission of video over the Internet is currently an active research and development area, where significant results have already been achieved. There are already video-on-demand services, both for news and entertainment applications, offered over the Internet. Also, 2.5G and 3G mobile network operators started to use IP successfully to offer wireless video services. Looking at these advances, the transport of 3DTV signals over IP packet networks seems to be a natural choice. The IP itself leaves many aspects of the transmission to be defined by other layers of the protocol stack and, thus, offers flexibility in designing the optimal communications system for various 3D data representations and encoding schemes. 3DTV streaming architectures can be classified as:

1. server unicasting to a single client:
2. server multicasting to several clients.
3. P2P multicasting, where each peer forwards packets to several other peers: Peer-to-peer (P2P) technology has the potential to provide a more cost effective and flexible delivery solution for future 3D entertainment services [15]. In terms of architecture definitions, two main architectures have been used, namely:
    - Mesh-based approaches: They derive from file sharing applications, peers organize themselves into a mesh, independently requesting pieces, or chunks, of the video content from neighbors, without any regard for the structure of the distribution path
    - Tree-based approaches: In tree-based approaches, video packets are forwarded along a pre-determined path forming multicast trees. The advantage of combining more than one multicast trees is that the robustness of the system is increased, due to path diversity.

2.4 Mobile 3D Media Delivery

The term ‘handover’ (HO) refers to the association of a Mobile Node (MN) from an old Point of Access (PoA) towards a new PoA wither belonging to the same access network or to another one. The mobility concept is defined in different ways by the relevant standardization fora [16]:

1. Nomadism: It is the ability of the user to change his network point of attachment while the end-user is on the move. When the network point of attachment is changed, the user's service session is completely stopped and it can be resumed later on.
2. Session Continuity: It refers to the ability that the end-user’s terminal can switch to a new network point of attachment while maintaining the on-going session from the old point of attachment to the new one. This may include a session break and resume, or a certain degree of service interruption or loss of data while changing to the new access point.”
3. Seamless handoff: the handoff algorithm should minimize the packet loss. It is sometimes referred to as smooth handoff. Transparent migration of on-going data flows between two access points belonging to independent heterogeneous technologies is achievable, and tools and mechanisms for supporting this type of mobility should be placed within the next generation networking architectures.

The co-existence of multiple wireless mobile clients that can support different networking technologies and multimedia services with heterogeneous requirements is becoming a common trend in future wireless era [17]. However, these wireless networks bear a higher level of randomness than their wired counterparts. Handling Mobility in an IP environment can be supported by Mobile IP and its extensions (Hierarchical Mobile IP, Mobile IPv6 and Fast Mobile IP with Handoffs) [18]-[21]. However, in a mobility environment comprising heterogeneous networks the support and maintenance of on-going sessions with strict QoS requirements may be a challenging task.

This necessitates the use of frameworks where handover can be either by the MT or the network through a cooperative synergy. This synergy may require the capturing of information/statistics from Physical, Network and Application Layer. This concept has been introduced under the Media Independent Handover Framework within IEEE 802.21 [22], [23]. When a node/client detects a deterioration of the received signal, it may initiate a procedure for registering itself on a different RAT, based on the available infrastructure. Such a Handover decision can be triggered by functions such as the Media Independent Handover (MIH) entity which decides on where to direct the client based on information coming from network discovery (e.g. best candidate AP, SNR), or based on information from network and the application server (e.g. QoE). Next, it applies this information to a pre-defined handover algorithm in order to decide which AP is the optimal solution for the handover [24].

Handover across heterogeneous RATs (Radio Access Technologies networks) of mobile devices with on-going 3D Video sessions is a challenging task. This is due to the fact of the number of views required by the end-user and the bit rate requirements for each view. Therefore in order to support and maintain QoS of a going-session the seamless mobility necessitates the transfer of each view towards a new wireless link. Due to the bandwidth constraints that are imposed by the different wireless networks different mobility strategies options will be considered:

1. Exploit mobility with path diversity so that each view may be rerouted to a different radio access network, combine handover with rate adaptation/transcoding in case of network bandwidth limitations.
2. Priority handling will be given to the base view (selecting the network that can accommodate the base view rate without deteriorating QoS and violating fairness of other connections).

The following figure illustrates an architecture how to support seamless mobility for 3D Video.

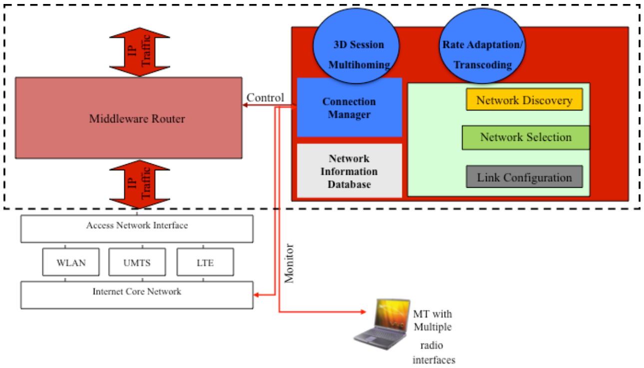

Figure 3: Seamless Mobility Architecture for 3D Video

2.3 Rate Control for 3D Video

Rate adaptation of multiview video may be achieved at constant perceived 3D video quality by adaptation of the spatial, temporal and/or Signal-to-Noise (SNR) resolution of one of the views while encoding and transmitting the other view at full rate. Several open loop and closed loop rate adaptation strategies for stereo and multiview video at the server and the client side are studied for UDP and DCCP protocols. In the closed loop rate adaptation, each client estimates some function of the received signal and feeds it back to the transmitter. The transmitter determines an optimized rate for the next transmission based on the received feedback. On the other hand, in the open loop rate adaptation, the transmitter does not use any feedback from the receiver.

Adequate perceptual metrics are necessary for 3D video quality in adaptation scenarios where different types of underlying communication channels might be used [25], [26]. The temporal effect of errors in depth perception is not yet fully known which leaves open questions about the subjective impact of freezing depth perception at the receiver during limited adaptation periods. The behavior of a 3D video system and the user perception is different when switching either to still frame or monoscopic view in case one stream is lost [27], [28]. An adaptation engine based on the subjective impact of such behavior would certainly yield better perceptual quality.

2.4 Error Concealment for 3D Video

3D Video Artifacts can be classified into the following categories: structure, color, motion and binocular [30]. These artifacts such as blurriness, distortion, aliasing, geometry distortion and cross-talk affect human perception on image structure such as contours and textures. The errors can be classified as error in shape and error in appearance [31].

Streaming media applications often suffer from packet losses at both wireline and wireless networks. Such losses may be due either to congestion or physical layer impairments. Error Concealment can be used in order to recover from errors and improve the perceived video quality. Such mechanisms may be grouped at the following categories:

1. Using Error Resilience mechanisms at the encoder part. There are several approaches how error resilience can be applied at the 3D Video that depend on 3D technology being used:
   - In MVC, it is exploited the fact that certain number of views are needed to each user at any time [11], [32], [33], [34], [35].
   - In 2D+V, 3D Video can be encoded with 10%-20% extra bit rate by exploiting the correlation between the video and the depth information [5], [6]
   - Use MDC, where the video is transmitted as several independent descriptors. Each descriptor may correspond to each view [36], [37]. Upon the reception of only one description, the bit stream can be decoded to construct a lower-quality representation.
2. Forward Error Correction: Using Unequal Error Protection (UEP), different portions of the 3D video bit stream are protected, depending on the relative importance of the views [38]. Flexible Marcoblock Ordering (FMO) could be used in order to classify the importance of bits that are generated by the encoder so that Unequal Error Protection Techniques can be applied to the encoder.
3. Exploit Temporal Correlations in 3D Domain: For stereo videos, temporal correlations within each view, in addition to inter-view correlations, can be used to conceal lost blocks or frames more effectively. [39], [40], and [41] propose error concealment algorithms for stereo videos can be employed to conceal erroneous regions in multi-view video sequences

3. Future Research Challenges

3.1MVC Coding

In the area of MVC coding, future enhancements regards MVC extensions to achieve highly flexible and scalable compressed representation. The overhead scalability should be kept to as low as possible as compared to non-scalable compressed 3D multi-view video. Similar to 2D SVC, scalable coding can offer flexibility in terms of scheduling to diverse networks. Each description will be scalable so that all descriptions can be efficiently adapted to the available rate of each link for effective congestion control. This necessitates the adoption of a framework for selecting the best encoding configuration for MD-SMVD, which will strike the best balance between minimizing the average end-to-end rate-distortion performance of each description given a set of packet loss probabilities and minimizing overall redundancy and maximizing the range of extraction points of each scalable description. The optimization variables will be some SMVD encoding parameters (e.g. layer QP values, macroblock modes, inter-layer prediction modes) and MD generation alternatives that result in different levels of redundancy at a fixed total rate for all descriptions

3.2 Novel 3D Video Networking

Network Coding is a promising technique that could be used for network content distribution [42], [43]. The concept behind network coding relies on the following remark. Communication networks today share the same fundamental principle of operation: whether it is packets over the Internet or signals in a phone network, information is transported in the same way as cars share a highway or fluids share pipes. That is, independent data streams may share network resources, but the information itself is separate. Routing, data storage, error control, and generally all network functions are based on this assumption. Network coding breaks this assumption. Instead of simply forwarding data, nodes may recombine several input packets into one or several output packets. It can be proven that the theoretical throughput within the network by applying coding (linear combinations of different pieces from the original content). In a large distributed cooperative system finding an optimal packet propagation scheme that minimizes the client download time is very difficult. This is especially the case in practical systems that cannot rely on a central scheduler and, instead, allow nodes to make local decisions. The scheduling problem becomes increasingly difficult as the number of nodes increases. A key question is the selection of network code (XOR operation among the packets) in order to optimize throughput [44], [45]. One important constraint with network coding is that as bandwidth efficiency increases, longer delays may be applied to some packets due to longer queuing time of packets, to allow packet losses occurrence. There is a lot of interest within the research community how network coding could be used in multimedia communications. This necessitates considering coding distortion conveyed in a video packet o construct the network information flows [46], [47]. Given that there are K downstream nodes with different packet loss rates, a key question is to design optimal scheduling algorithms that determine which packets (pure and mixed packets) are transmitted at certain times in order to maximize video quality. The key question that is raised regards the network code (XOR operation among the packets) so that both video distortion and throughput is optimised. The distortion of each packet can be determined by the source and the view type and communicated to the distribution nodes in order to get transmission in a rate-distortion optimised manner [47].

The employment of network coding paradigm will provide efficiency in terms of throughput within the core/distribution network, distributing large chunk of MVC 3D video packets across different edges where heterogeneous wireless networks (e.g. 3G, WLAN, WiMAX and LTE) are located. It is important to consider network coding algorithms that are video aware in terms of the importance of the packets that would significant improve throughput. It is important to consider how network coding could be used in an efficient manner for Scalable MVC. The advantage of scalable video coding necessitates the use of strategies that combine packets of the same important to the network information flow. This necessitates exploiting video packets levels of importance (e.g. in MVC packets from base view are important in order to reconstruct non-base view) and how this information can be correlated with network codes so that several packets from different views can be merged together increasing throughput. Two scenarios can be exploited. In the first one, packets from one view can construct a network flow, whereas in the second case packets from the same layer from multiple views can construct an information flows. These ideas are presented in the following figure, by considering two layers and two views.

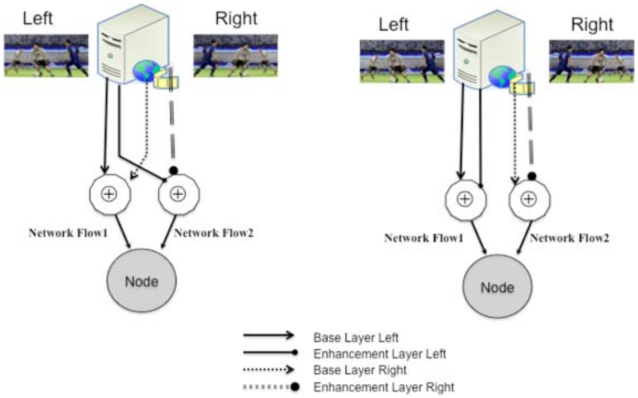

Figure 4: Network-Aware Network Coding for Scalable 3D MVC Video

3.3 Error Concealment for 3D Video

Error Concealment Schemes for 3D Video could fall in the following categories:

1. Network Assisted Mechanisms where out-band signalling protocols may be used in order to send parameter sets and important information (i.e. GOP size, Number of Views, NAL Unit Headers, View id) that are important for the reconstruction of the 3D View.
2. Concealment methods for missing depth information due to either transmission errors or packet loss. Existing spatial and temporal interpolation techniques will be subject to further research and development in order to achieve efficient perceptual concealment of depth artifacts in 3D video. The additional level of dependency between views also brings a new dimension into the problem of error propagation, which must be taken into account.
3. Methods for restoration of compressed depth maps affected by blocking artifacts and reconstructed missing areas. The 3D perceptual quality will be evaluated using both objective metrics and subjective testing.

4. Algorithms for rendering occluded areas and inpainting will be investigated, in order to provide reconstructed views with a higher fidelity.
5. UEP mechanisms based on dynamic combination of FEC and depth information will be devised to increase the quality of 3D video delivered over error prone channels. Compressed depth maps along with light forward error correction are transmitted for increased error resilience and improved error concealment at the used terminal.
6. Error concealment methods based on structural and motion correlations of 2D plus depth video will be investigated. Moreover, the redundant information available in adjacent views will be employed to restore missing image blocks in both 2D and depth map video.

3.4 Quality of Experience

QoE is inherently a subjective concept. The ultimate way of measuring QoE is to obtain subjective measurements of the different dimensions of QoE. Subjective evaluations attempt to assess the perceived quality of audio and video multimodal networked services. Several methodologies for subjective evaluation of video quality, sound quality and multimedia quality for particular applications have been standardized within the ITU, e.g. recommendation ITU-R BT.500-11 [48], which deals with assessment of television picture quality, and ITU-T P.910 [49], which deals with video quality in multimedia applications. Moreover, ITU-T Recommendation P.911 describes subjective assessment methods for evaluation of audiovisual quality in multimedia applications, P.920 includes aspects relevant for interactive applications [51] and ITU-R Recommendation BS-1284 specifies general methods for the subjective assessment of sound quality.

On the other hand, objective measurement methods analyse and measure characteristics of the source and/or the reconstructed processed media signal by predicting the perceived quality of audio, video and speech. An extensive coverage of existing models for predicting video quality in both a double-ended and single-ended fashion is provided by Winkler, in his book, entitled “Digital Video Quality – Vision Models and Metrics”, published by John Wiley in 2005.

Visual attention utilises a bottom-up control model of visual attention in primates, as introduced in [52] and [53]. This can be accomplished by a decomposition of the input image into a set of multi-scale neural “feature maps” which extract local spatial discontinuities in the modalities of colour, intensity and orientation. Compared to the underlying original psychological attention model of Treisman [54], the feature “stereo-distance” is still missing in [55]. Additionally, Treisman distinguishes between top-down and bottom-up features. Top-down features lead the search for salient objects by prior knowledge on context and / or object properties.

Attention Models using top-down information are rarely available in literature and mainly known from robotics. Novel visual attention models could be used to detect important regions by extracting salient image features. These regions of interest could be used for region based coding and guidance of the content adaption process.

3D immersive environments pose the unique situation of service operators having to deliver multiple multimedia types directly to the user. Interfacing with the users avatar within the environment would allow the user full control of their QoE and make the architecture truly adaptive. However this is not the only way QoE could be attained in such environments. Certain activities, which take place within the environment, can require different levels of QoE. For instance when an user’s avatar is interfacing with the 3D environment itself would require a different level of QoE in certain medias than if a user’s avatar was interacting with another user. Thus, due to the subjective quality of the interaction itself users can be satisfied in different ways. The wide range of applications used today on mobile devices means that many different levels of service are required to satisfy users.

4. Conclusions

This paper surveys the current and future research trends associated with Next generation 3D Media Applications over Future Internet. Particular emphasis has been given on the media encoding and delivery of 3D Applications across heterogeneous content-aware networks (e.g. wireless such as LTE, WiMAX, DVB and wireline such as xDSL, FTTx), in an ecosystem across the entire value chain of Networked Media (from the encoding/packetisation, through the transmission and up to end-user).

References

[1] P. Daras, F. Alvarez,: A future perspective on the 3D Media Internet. In: Tselentis, G.e.a. (ed.) Towards the Future Internet - A European Perspective, pp. 303–312. IOS Press, Amsterdam (2009)
[2] ISO/IEC JTC1/SC29/WG11, “ISO/IEC CD 23002-3: Representation of auxiliary video and supplemental information”, Doc. N8259, Klagenfurt, Austria, July 2007.

[3] ISO/IEC 14496-10:2005/FDAM 3 Scalable Video Coding, Joint Video Team (JVT) of ISO-IEC MPEG & ITU-T VCEG, Lausanne, N9197, Sep. 2007.

[4] Y. Chen, Y.-K. Wang, K. Ugur, M. Hannuksela, J. Lainema, and M. Gabbouj, "The Emerging MVC Standard for 3D Video Services", EURASIP Journal on Advances in Signal Processing, Volume 2009

[5] A. Smolic, et al, "Coding Algorithms for 3DTV: A Survey", IEEE Trans. Ob Circuits and Systems for Video Technology, Vol. 17, November 2007

[6] P. Kauff et al,"Depth map creation and image-based rendering for advanced 3DTV services providing interoperability and scalability", Signal Processing: Image Communication, Vol. 22, February 2009

[7] K. Muller et al, ""View synthesis for advanced 3D video systems", EURASIP Journal on Image and Video Processing, 2008

[8] K. Muller et al, "3D Visual content compression for communications", IEEE Multimedia Communications E-Letter, Vol. 4, July 2009

[9] IETF Draft, "RTP Payload Format for SVC Video", draft-ietf-avt-rtp-svc, October 2010

[10] K.-D. Seo, J.-S. Kim, S.-H. Jung, and J.-J. Yoo, "A Practical RTP Packetization Scheme for SVC Video Transport over IP Networks", ETRI Journal, No. 2, April 2010, DOI: 10.4218/etrij.10.1409.0031

[11] A. M. Tekalp, E. Kurutepe and M. R. Civanlar, "3DTV over IP", IEEE Signal Processing, November 2007

[12] IETF RFC 4340, "Datagram Congestion Control Protocol", March 2006

[13] IETF RFC 4341, "Profile for Datagram Congestion Control Protocol (DCCP) Congestion Control ID 2: TCP-like Congestion Control", March 2006

[14] IETF RFC 4342, "Profile for Datagram Congestion Control Protocol (DCCP) Congestion Control ID 3: TCP-Friendly Rate Control (TFRC)", March 2006

[15] Y. Liu, Y. Guo and C. Liang, "A survey on peer-to-peer video streaming systems", *Peer-To-Peer Networking and Applications*, Vol. 1, 2008

[16] N. Passas et al, "Architectures and protocols for mobility management in All-IP Mobile Networks", IEEE Wireless Magazine, April 2008

[17] Y. M. Fang, "Special Issue: Seamless Content Delivery in the Future Mobile Internet", IEEE Wireless Communications Magazine, October 2009

[18] IETF RFC, "IP Mobility for IPv4", March 2002

[19] IETF RFC 4140, "Hierarchical Mobile IPv6 Mobility Management", August 2005

[20] IETF RFC 4068, "Fast Handovers for Mobile IP", July 2005

[21] IETF RFC 4721, " Mobile IPv4 Challenge/Response Extensions", January 2007

[22] IEEE 802.21/D10.0. Draft Standard for Local and Metropolitan Area Networks: Media Independent Handover Services (Work in progress). Technical report, IEEE Draft, April 2008

[23] G. Lampropoulos, A.K. Salkintzis, and N. Passas, "Media-Independent Handover for Seamless Service Provision in Heterogeneous Networks," IEEE Communications Magazine, Vol. 46, January 2008.

[24] J. Rodriguez, M. Tsagkaropoulos, I. Politis, T. Dagiuklas and S. Kotsopoulos, "A Middleware Architecture Supporting Seamless and Secure Multimedia Services across Inter-Technology Radio Access Network", IEEE Wireless Communications Magazine, October 2009

[25] G. B. Akar, M. Tekalp, C. Fenn and R. Civanlar, "Transport methods in 3DTV- a survey", IEEE Trans. On Circuits and Systems on Video Technology, Vol. 17, November 2007.

[26] A. Aksay et al, "End-to-end stereoscopic video streaming system with content-adaptive rate and format control", Signal Processing: Image Communication, Vol. 22, February 2007

[27] S. Jolly, M Armstrong and R Salmon, The Challenges of Three-Dimensional Television, White Paper WHP 173, BBC, January 2009

[28] O. Stankiewicz, K. Wegner, M. Domański, "Error Concealment for MVC and 3D Video Coding", Picture Coding Symposium 2010,Nagoya Japan, December 2010

[29] A. Boev, D. Hollossi, A. Gotchev and K. Egiazarian, "Classification and simulation of stereoscopic artifacts in mobile 3DTV content", Proc. Of SPIE, January 2009

[30] A. Boev, D. Hollossi, A. Gotchev and K. Egiazarian, "Classification and simulation of stereoscopic artifacts in mobile 3DTV content", Proc. Of SPIE, January 2009

[31] J. Kilner, J. Starck, J. Y. Guillemaut and A. Hilton, "Objective quality assessment in free-viewpoint video production", Signal Processing: Image Communication, Vol. 24, January 2009

[32] G. B. Akar, M. Tekalp, C. Fenn and R. Civanlar, "Transport methods in 3DTV- a survey", IEEE Trans. On Circuits and Systems on Video Technology, Vol. 17, November 2007.

[33] X. Cao, Y. Liu, Q. Dai, "A flexible client-driven 3DTV system for real-time acquisition, transmission and display of dynamic scenes", EURASIP Journal on Advances in Signal Processing, Vol. 2009

[34] E. Kurutepe, and T. Sikora, "Feasibility of multi-view video streaming over P2P networks", IEEE 3DTV Conference, Instabul, Turkey, May 2008

[35] J. Lou, H. Cai and J. Li, "Interactive multiview video delivery based on IP Multicast", EURASIP Advances in Multimedia, Volume 2007

[36] A. Norkin et al, "Schemes for multiple description coding of stereoscopic video", LNCS on Multimedia Content Representation, Classificatoon and Security 2006

[37] H. A. Karim, 'Scalable Multiple Description Video Coding for Stereoscopic 3D', IEEE Transc. On Consumer Electronics, May 2008

[38] A. Tan, A. Aksay, G. Akar and E. Arikan, "Rate-distorton optimization for stereoscopic for 3D transmission", EURASIP Journal on Advances in Signal Processing, Vol. 2009

[39] K.A.M. Guenther, C. Clemens, T. Sikora, A fast displacement-estimation based approach for stereoscopic error concealment, Proc. PCS 2004, Oregon, USA, November 2004

[40] L. Pang, M. Yu, W. Yi, G. Jiang, W. Liu, Z. Jiang, Relativity analysis-based error concealment algorithm for

entire frame loss of stereo video, in: Proc. International Conference on Signal Processing, 2006.

[41] X. Xiang, D. Zhao, Q. Wang, X. Ji, W. Gao, A novel error concealment method for stereoscopic video coding, in: Proc. ICIP'07, 2007, pp. 101–104.

[42] R. Ahlswede, Ning Cai, Shuo-Yen Robert Li, and Raymond W. Yeung, "Network Information Flow", IEEE Trans. On Information Theory, Vol. 46, July 2000

[43] H. Wang, J. Liang and C.-C. J. Kuo, "Overview of Robust Video Streaming with Network Coding", *Journal of Information Hiding and Multimedia Signal Processing*, January 2010

[44] S. Katti et al, "XORs in the air: Practical network coding", *ACM SIGCOMM*, Pisa, Italy, September 2006

[45] R. Koetler and M. Medard, 'An algebraic approach to network coding', IEEE/ACM Trans. On Networking, 2003

[46] IEEE COMSOC MMTC E-Letter, "Network Coding for Multimedia Communications", March 2010

[47] H. Seferoglu and A. Markopoulou, 'Video-Aware Opportunistic Network Coding over Wireless Networks", IEEE JSAC, Vol. 27, No.5, pp. 1-16

[48] Recommendation ITU-R BT.500-11, "Methodology for the subjective assessment of the quality of television pictures," ITU-R, Geneva, 1974-1997

[49] Recommendation ITU-T P.910, "Subjective video quality assessment methods for multimedia applications," ITU-T, Geneva, 1996.

[50] Recommendation ITU-T P.911, "Subjective audiovisual quality assessment methods for multimedia applications" ITU-T, Geneva, 1998.

[51] Recommendation ITU-T P.920, "Interactive test methods for audiovisual communications," ITU-T, Geneva,1996.

[52] S. Frintrop, "VOCUS: A Visual Attention System for Object Detection and Goal-Directed Search", Lecture Notes in Artificial Intelligence, vol. 3899, 2006

[53] Yao Zhang, Chun Yuan, Yuzhuo Zhong: Implementing DRM over Peer-to-Peer Networks with Broadcast Encryption. In: Proceedings of the 8th Pacific Rim Conference on Multimedia, pp. 236-245, 2007.

[54] A. Treisman, "Features and Objects in Visual Processing," Scientific American, vol. 255, pp. 106 – 115, 1986.

[55] L. Itti, C. Koch, E. Niebur, "A Model of Saliency-based Visual Attention for Rapid Scene Analysis," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 20, pp. 1254–1259, 1998

Tasos Dagiuklas received the Engineering Degree from the University of Patras-Greece in 1989, the M.Sc. from the University of Manchester-UK in 1991 and the Ph.D. from the University of Essex-UK in 1995, all in Electrical Engineering. Currently, he is employed as Assistant Professor at the Department of Telecommunications Systems and Networks, Technological Educational Institute (TEI) of Mesolonghi, Greece.

He is the Leader of the Converged Networks and Services Research Group. He is also Senior Research Associate within the Wireless Telecommunications Laboratory of the Electrical and Computer Engineering Department at the University of Patras, Greece. Past Positions include teaching Staff at the University of Aegean, Department of Information and Communications Systems Engineering, Greece, senior posts at INTRACOM and OTE, Greece. He has been involved in several EC R&D Research Projects under FP5, FP6 and FP7 research frameworks, in the fields of All-IP network and next generation services.

He has served as TPC member to more than 30 international conferences. His research interests include Future Internet architectures and converged multimedia services over fixed-mobile networks. Dr Dagiuklas has published more than 100 papers at international journals, conferences and standardisation fora in the above fields.