The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC'07)

A DEVICE AND NETWORK-AWARE SCALING FRAMEWORK FOR EFFICIENT DELIVERY OF SCALABLE VIDEO OVER WIRELESS NETWORKS

Naomi Ramos and Sujit Dey
Dept. of Electrical and Computer Engineering
University of California, San Diego
{naomir, dey}@ece.ucsd.edu

ABSTRACT

In recent years, the number of devices that can support multimedia-content over wireless networks has increased significantly. With devices ranging from cellular phones to TVs, there is a challenge faced by content service providers to support multimedia delivery for a diverse set of devices. Scalable video coding (SVC) may provide a solution by enabling devices with techniques to extract video streams scaled to match their capabilities. However, SVC can be significantly bandwidth inefficient. By sending the highest possible quality stream over the network, irrespective of the capabilities of the end-devices and the network, SVC can limit the number of devices being served and reduce the network's average quality satisfaction. In this paper, we present a Device and Network-Aware Scaling (DeNAS) framework to address these problems. Based on device capability information, network capacity, and the desired service objective, DeNAS finds the suitable scaling level for each video stream prior to its delivery over the wireless network. We demonstrate the effectiveness of DeNAS through simulation and show that DeNAS is able to improve quality satisfaction, increase bandwidth efficiency, and satisfy a greater number of clients simultaneously.

I INTRODUCTION

Fueled by the improvements in wireless technology and the proliferation of multimedia content, there is a growing demand for multimedia over wireless networks. Traditional wireless devices, such as cellular phones and PDAs, are being transformed into multimedia-capable platforms. Additionally, traditional multimedia terminals, such as PCs and TVs, are now being designed to support wireless connectivity. With these developments, the devices requesting multimedia data over wireless networks vastly differ in terms of capabilities, such as screen resolution, supported frame rate, as well as supported bit rate. Content service providers are now faced with the challenging problem of enabling seamless and efficient delivery of multimedia content to viewers using different devices.

To provide appropriate video content to varying devices and networks, video systems have been reliant on two device-centric schemes. The first is to encode and store various versions of the video at the server *a priori* and send the prepared video stream based on device capabilities. Another method is to use transcoding techniques [1] to decode and re-encode the video before delivery. While both of these schemes can be currently supported, the increasing availability of video content and the greater diversity in device capabilities will make it difficult to continue using these methods in future video systems. The first method of preparing and storing multiple versions suffers from the major drawbacks of needing significant content management and storage space, while the second method of transcoding is costly in terms of the number of servers needed to support quick decoding and re-encoding of each video stream.

An alternative to these delivery systems is the recent work in scalable video coding. Over the past few years, significant strides have been made in this area in improving coding efficiency and reducing decoder complexity. These developments have provided the motivation for the Joint Video Team (JVT) to develop and standardize a more powerful scalable video codec, called the H.264/AVC Scalable Video Coding (SVC) Extension [2]. The proposed SVC is a layered video codec that provides bit-stream scalability and allows video to be encoded once at the server in such a way that various scaled streams can be extracted and decoded from the single stream, without the need for re-encoding the video. Building on scaling techniques in previous standards, SVC makes it possible to extract video scaled in multiple dimensions, *e.g.* spatial, temporal, and quality. Future systems using SVC will allow devices to extract a video stream scaled to match their capabilities.

Although SVC provides a notable advantage over previous video standards, the penalty for this scalability comes in terms of the network delivery costs. With SVC, the entire stream, allowing for video extraction at the highest possible quality, will be sent over the network without regard for the capabilities of each receiving device. This is a significant problem because it causes bandwidth inefficiency and limits the overall number of devices that can be supported simultaneously. Additionally, there is a cost for processing and receiving the additional scalability layers unnecessarily, in terms of energy, processing power, and memory.

To address the above problem, we describe a Device and Network-Aware Scaling framework, called DeNAS, that will provide efficient delivery of scalable video over wireless networks. Based on device capabilities, network capacity, and the desired service objective, DeNAS determines the suitable scaling level for each stream. By performing the appropriate scaling prior to video delivery over the wireless network, DeNAS is able to retain the advantages of SVC of using a single-encoded stream, but is also able to increase bandwidth efficiency, support a greater number of simultaneous users, and improve quality satisfaction.

DeNAS can be used in conjunction with most of the recent SVC-related work, which includes techniques for supporting SVC in the lower layers, such as providing transport layer support [3] and prioritization schemes and error protection at the

1-4244-1144-0/07/$25.00

link layer [4], as well as methods of layer switching at the device decoder with little change to the encoder [5]. While this set of related work looks at supporting full SVC streams, DeNAS differs in that it focuses on scaling these full streams prior to delivery over wireless networks.

II MOTIVATION

We now provide motivation for a device and network-aware scaling framework. As stated earlier, SVC provides many advantages in terms of scalability; however, it can be significantly bandwidth inefficient. As an example, consider a scenario where a video stream (*BUS*) is being requested over a wireless network by a set of devices. To keep this example simple, we consider an equal distribution of three possible devices; a laptop, a PDA, and a cellular phone with capabilities that limit viewing quality to a resolution, frame rate (frames per second), and bit rate (kb/s) of CIF 30-512, CIF 15-256, and QCIF 15-128 respectively.

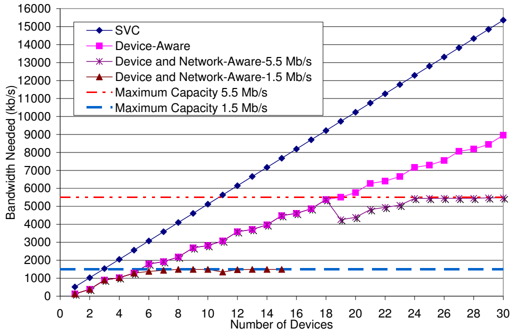

Figure 1: Bandwidth Needed vs. Number of Total Devices

With SVC, the video stream is encoded once at the server, such that the maximum quality level can be extracted from the stream by the receiving device. In this scenario, the *BUS* video is encoded and sent to each device so that the maximum scalability level (CIF 30-512) can be extracted, irrespective of the device's capabilities. Fig. 1 shows the bandwidth usage of this method. Note that the bandwidth needed with SVC increases linearly with the number of devices in the network. If a basic admission control scheme is used to limit the number of devices based on the required bandwidth, the network can only support a limited set of devices with SVC. For example, in a network with an application layer throughput limit of 5.5 Mb/s, such as in an 802.11b network, only 10 devices can be supported with SVC. Similarly, in a cellular network with a 1.5 Mb/s application layer throughput limit, only 2 devices can be supported simultaneously in a cellular site. Additionally, the average quality satisfaction with SVC, shown later in Fig. 5 and discussed in Section IV, is at a low value of 0.33 in the case of 30 devices, since only 10 devices are able to extract a stream matching their maximum capabilities.

Next, we look at what can be achieved using the proposed Device and Network-Aware Scaling (DeNAS) framework. Fig. 1 illustrates the results of a device-aware algorithm. Prior to sending the video over the wireless network, this algorithm extracts a scaled video stream matching the maximum capabilities of each device. In this scenario, the laptop would be sent the full CIF 30 video, while the phone would be sent the QCIF 15 video. As seen in the figure, a device-aware scheme is a significant improvement over SVC, with up to a 42% reduction of bandwidth usage in the case of 30 devices. Additionally, the device-aware scaling algorithm increases the number of devices that can be supported, from 10 devices with SVC to 18 devices given the application throughput limit of 5.5 Mb/s. In the case of 1.5 Mb/s, the device-aware scheme is able to increase the number of supported devices from 2 to 5 devices. Quality satisfaction has also improved with a device-aware scheme, with more devices being able to extract the suitable stream, increasing the quality satisfaction from 0.33 with SVC to 0.60 in the case of 30 devices.

With the above algorithm, the scaling is done based solely on device limitations. However, the DeNAS framework can also use network capacity information to further increase the number of clients supported and improve the average quality satisfaction across the network. The device and network-aware scaling algorithm, which is described in Section III, focuses on maximizing quality satisfaction, while improving bandwidth efficiency and supporting a greater number of users. The results with this proposed scaling algorithm can be seen in Fig. 1. Bandwidth usage is reduced by 64% in the case of 30 devices compared to SVC. Additionally, the device and network-aware scaling algorithm increases the number of simultaneous video users by 3x, from 10 devices with SVC to 30 given a 5.5 Mb/s application throughput limit. With a 1.5 Mb/s limit, we can increase the number of users from 2 to 15 devices. The algorithm is also able to maintain an acceptable quality level for each device, using scalability properties of the video. Quality satisfaction improves significantly with the device and network-aware algorithm. With 30 devices, the average quality satisfaction improves from 0.33 with SVC to 0.95.

III DeNAS FRAMEWORK AND SCALING ALGORITHM

Having shown the drawbacks of SVC and the potential for efficiently delivering video to a diverse set of devices, we now describe the proposed Device and Network-Aware Scaling framework. We first provide a discussion of the framework components and next describe the video scaling algorithms in detail.

In order to perform the scaling prior to the wireless network, the DeNAS framework must be located such that it can access information about the network, device, and application. The DeNAS framework can be placed in a proxy-type device on the edge of the wireless network or on a high-end Access Point (AP), as shown in Fig. 2.

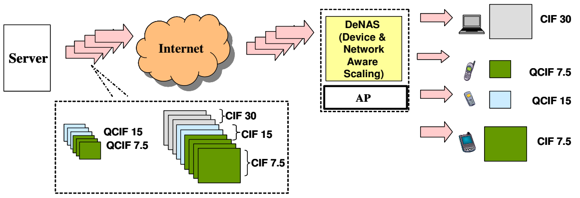

Figure 2: DeNAS Framework in a Wireless LAN Network

Clients make requests to the server for a given video and DeNAS intercepts these requests and obtains the scalable video stream from the appropriate server. DeNAS uses the SVC

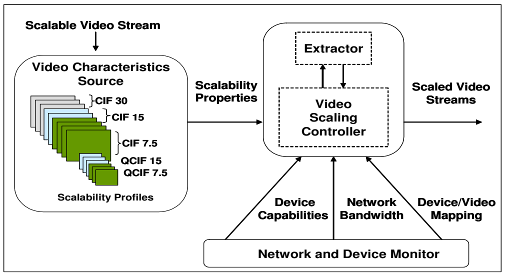

Figure 3: DeNAS Components and Interactions

video stream, scalability property information, and network bandwidth limitations to determine the appropriate scaling for each of the devices.

A DeNAS Components

DeNAS consists of three major components: Network and device monitor, Video characteristics source, and a Video scaling controller and extractor. Fig. 3 illustrates the components and their interactions with one another.

1) Network and Device Monitor

The first component is the Network and Device Monitor. This component gathers information about the capabilities of the devices, determines which videos are being requested by each device, and obtains the available bandwidth of the network. The network monitor either queries the AP to determine the number of devices or directly obtains current channel condition information from individual devices to estimate the current network bandwidth available. For the device information, the monitor's first task is to determine what type of device is requesting the video stream. The device type can be discovered during the RTP/RTSP handshake. The next step is to perform a device lookup in a Wireless Universal Resource File ($WURFL$) [6] database and obtain information about the device capabilities, *e.g.* maximum supported resolution, frame rate, and bit rate.

2) Video Characteristics Source

The Video Characteristics Source component receives the video stream and provides information to the DeNAS controller. For each video stream in the network, the component determines a set of enumerated scalability profiles. Each of these profiles represent a scalability level determined by a spatial, temporal, and bit rate combination. For example, the video stream shown in Fig. 3 has scalability profiles given by a combination of different spatial (CIF/QCIF), temporal (7.5/15/30 fps) and associated bit rate levels. The next step is to determine the attainable quality and associated cost for each profile. The quality information can be calculated offline by the server during the encoding stage and sent with the video stream. Alternatively, a dynamic video quality profiler can be used to determine the attainable quality of the scalable profiles. For the associated cost of the scalability profile, we make use of the profile's bit rate requirements.

3) Video Scaling Controller and Extractor

The final component is the Video Scaling Controller and Extractor. For a new set of incoming devices, the received and monitored inputs are used to determine the scaling level for each of the devices. The scaling levels chosen are dependent on the algorithm's desired service objective. For example, with a device-aware scaling algorithm, the controller bases the scaling solely on the capabilities of individual devices. On the other hand, for a device and network-aware algorithm, the controller considers information about the devices, as well as the available network bandwidth. We present algorithms to be used in the controller in the following subsection. Once a scaling level is chosen, DeNAS extracts the appropriate scalable video stream from the original stream. The extractor can support various scaling input parameters, including desired spatial, temporal, and quality levels.

B Proposed Video Scaling Algorithms

We now present two video scaling algorithms for the Video Scaling Controller. For each, we provide a description of the service objective and details about the algorithm. In our service objective formulations, we consider a wireless network with N clients each requesting a unique video stream. For each video stream, the set of scalability profiles is given by S_j:$[S_{j1}, S_{j2}, ...S_{jL}]$, where j indicates the video stream index and L represents the maximum number of scalable profiles available for the stream. The chosen profile for all clients in the network is given by $P : [P_i \mid i = 1 : N, P_i \in S_{v(i)}]$, where i is the client index and $v(i)$ is the video for client i.

1) Device-Aware

The first algorithm we consider is a device-aware scaling algorithm. The service objective of this algorithm is to maximize individual quality given device capability constraints. In other words, we want to find a set P such that $\forall i \in N$ $P_i.qual \equiv maxQual_i$, where $P_i.qual$ is the quality for the selected scalability profile and $maxQual_i$ is maximum attainable quality of the video stream given the device's capabilities. With this algorithm, the suitable scaling profile is determined based on the device capabilities and extracted prior to delivery over the wireless network. If the aggregate bandwidth demand of the devices exceeds the available network capacity, this algorithm selects a subset of devices in the network to support. While this algorithm improves bandwidth efficiency, as was shown in Section II, we can further increase the number of supported devices if we consider the algorithm described in the next subsection.

2) Device and Network-Aware

The service objective of this algorithm is to maximize the overall quality satisfaction given network and device constraints. We consider a quality satisfaction function Q that relates experienced quality to required quality and can also be weighted by an individual device's priority or experienced channel conditions. The constraints of the above objective are to maintain a minimum quality requirement, ensure that the bandwidth allocated does not exceed the overall network bandwidth and choose a profile that does not exceed what can be supported by the device. The objective can be expressed by the following expression, where i is the client index, P_i is the selected profile index with $P_i.bw$, $P_i.qual$, and $P_i.req$ representing the bandwidth needed, the quality achievable, and the requirements of the given profile respectively, Q is a function of quality satisfaction, $bwAvail$ is the total bandwidth available, $minQual_i$ is the minimum required quality guaranteed in the network, and $maxCap_i$ is the maximum capabilities of user i.

$$\begin{array}{lll} \max\limits_{P} & \sum_{i=1}^{N} Q(P_i) & \\ subject\ to & \forall i \in N & P_i.qual > minQual_i\ \& \\ & & P_i.req \leq maxCap_i \\ & and & \sum_{i=1}^{N} P_i.bw < bwAvail \end{array}$$

Fig. 4 presents a flow chart of the proposed scaling algorithm to meet the above objective and constraints. Briefly, we can summarize the flow diagram in the following steps. The assumed inputs to this algorithm are the video stream, information about scalability properties, network bandwidth availability, and device information, which are gathered by the other components.

- *Step 1:* Using device capability information, determine if the maximum quality obtainable ($maxQual_i$) for each device can be supported. If so, extract the device-appropriate stream but if not, select a set of devices in the network such that the minimum quality requirements ($minQual_i$) of each device can be met.
- *Step 2:* If the network is not overloaded, determine potential change in aggregate quality ($\Delta aggQual$) and aggregate bandwidth cost ($\Delta aggBW$) if all video streams are upgraded to a higher scalability profile. The upgrade can be based on simple bandwidth allocation or on expected quality impact given video scalability properties. Based on device capability information, limit the maximum profile obtainable with these upgrades.
- *Step 3:* Check for the following criteria: (1) the potential change in aggregate quality is significant and over a set threshold ($upQLimit$) and (2) the aggregate bandwidth cost is within the available network bandwidth. If the above criteria are met, upgrade the profiles for each device and repeat Step 2, considering additional network-wide upgrades.
- *Step 4:* When network-wide upgrades are no longer possible, check to see if it can improve the quality satisfaction of individual devices. Select an individual device (j) based on weights to upgrade if bandwidth is still available and calculate the potential impact on individual quality ($\Delta P_i.qual$) and cost of the upgrade. If the upgrade is possible given the device capabilities and network constraints, upgrade the scalability level and repeat until no additional upgrades are possible.

By the end of the process, the proposed device and network-aware scaling algorithm is able to select streams that maximize the aggregate quality satisfaction of the network. The algorithm insures a minimum quality for each video stream, performs network-wide upgrades for fairness, maintains bandwidth efficiency by not exceeding device limitations, and stays within the network bandwidth limitations.

IV EXPERIMENTAL RESULTS

In this section, we present the experimental evaluation of the DeNAS framework and video scaling algorithms. For our experiments, we implemented the two video scaling algorithms for the Video Scaling Controller. For the scalable video encoding, decoding, and extraction of streams, we incorporated the JSVM6 Scalable Video Coding Software [2].

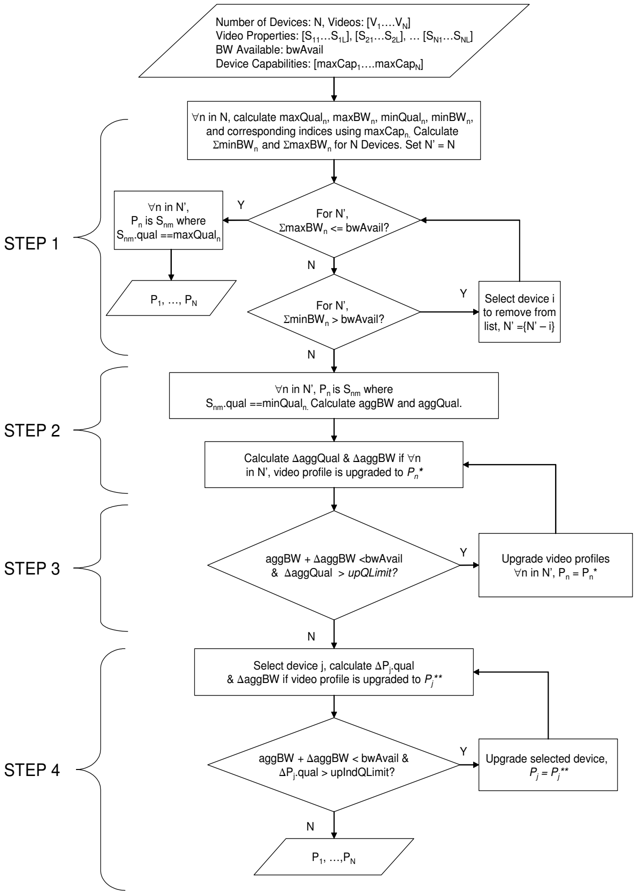

Figure 4: Algorithm Flowchart

Three video sequences are used in our simulations: *BUS*, *FOREMAN*, and *AKIYO*. We assume the scalable video streams have been encoded allowing for two layers of spatial scalability, CIF and QCIF, temporal scalability of 15 and 30 fps, and quality scalability appropriate for the given temporal-spatial selection. Quality information is given using a discrete set of scalable levels for a particular video. Scalability properties of the profiles are assumed to be available for the DeNAS controller. For our quality metric, we use the NTIA Video Quality Metric (VQM) [7]. To make the quality metric more intuitive, we set the quality metric for the experiments to be $1 - VQM$, where 1 signifies no distortion and 0 signifies maximum distortion. The minimum acceptable quality for each device is 0.5. To capture quality satisfaction, we consider the ratio between achieved quality and maximum achievable quality for each device. We look at the average quality satisfaction ratio for the devices in the network expressed by the following: $QS = \sum_{i=1}^{N} Q(P_i)$ where $Q = (1/N) * (P_i.qual/maxQual_i)$.

A Simple Scenario

We first revisit the scenario used in Section II for a network that supports 5.5 Mb/s application layer throughput. In this

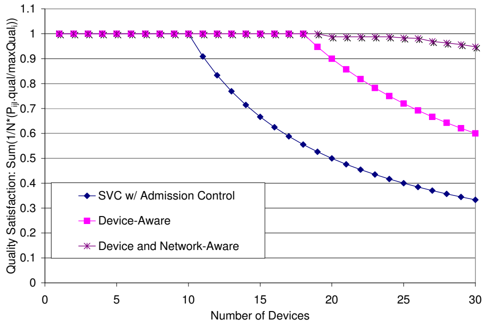

Figure 5: Quality Satisfaction vs. Number of Total Devices

scenario, we consider three set of devices in the network with the viewing capability translating to CIF 30, CIF 15, and QCIF 15. We assume that each device is requesting one *BUS* video stream and a basic admission control is used to limit the number of devices in the network based on the bit rate needs of the stream.

Fig. 1 in Section II showed the benefits of the DeNAS framework in terms of bandwidth efficiency. In comparison to SVC, the device-aware algorithm reduces bandwidth usage by 42%, while the device and the network-aware algorithm reduces usage by 64%. We now look at the impact of DeNAS on the average quality satisfaction ratio QS. Recall that QS reflects the average ratio of achieved quality versus the maximum achievable quality of the devices. As seen in Fig. 5, QS is high when the number of devices in the network can be fully supported. However, as the number of nodes increase beyond the maximum number of devices supported, the QS decreases significantly with SVC and the device-aware scheme because some devices are left unserved. The device-aware scheme, however, fares better than SVC with the quality satisfaction at 0.60 with 30 devices versus the 0.33 achieved with SVC. On the other hand, the proposed algorithm is able to maintain a quality satisfaction of 0.95 with 30 devices. Not only is the device and network-aware algorithm able to scale the streams to satisfy a greater number of clients, but it also is able to maintain quality satisfaction.

B Random Device Configuration Scenarios

After looking at a simple scenario, we evaluate the proposed device and network-aware scaling algorithm over multiple test cases where both the number of devices and their capabilities varied. We look at three different video sequences and for each video sequence, we vary the number of devices in the network. The capabilities of each of the devices are randomly chosen using equal probabilities.

Table 1: Experiments with Varying Device Capabilities

Video Name- # of Devices	# of Devices Srvd.		Avg. Quality Exp.		Quality Sat.	
	SVC	DeNAS	SVC	DeNAS	SVC	DeNAS
Bus-10	10	10	0.670	0.670	1.00	1.00
Bus-20	10	20	0.383	0.619	0.500	0.920
Bus-30	10	30	0.263	0.617	0.333	0.927
Foreman-20	20	20	0.638	0.638	1.00	1.00
Foreman-30	21	30	0.462	0.645	0.700	1.00
Foreman-40	21	40	0.353	0.619	0.525	0.965
Akiyo-20	20	20	0.882	0.882	1.00	1.00
Akiyo-30	21	30	0.625	0.854	0.700	0.964
Akiyo-40	21	40	0.477	0.834	0.525	0.944

Table 1 summarizes the results of these experiments. Each row represents the average of five test cases with a given video stream and number of devices combination. This table reports the number of devices supported in the network, the average VQM score, and the average quality satisfaction score over the five runs. In the experiments with SVC, we assume a favorable admission control policy that serves the devices with the highest possible achievable quality. This table shows that our proposed scaling algorithm is able to support a greater number of devices in the network, up to 3x in the case of the BUS video stream with 30 nodes. Additionally, the average VQM score and satisfaction score calculated across all devices in the network is significantly higher with our video scaling algorithm versus SVC. While the quality satisfaction and achieved quality for SVC quickly deteriorates with more devices, DeNAS scales well with increasing number of devices, with minimum degradation in quality satisfaction and average quality experienced. DeNAS is able to achieve a quality satisfaction score in the range of 0.92-1, while the quality satisfaction using SVC is as low as 0.33 with the *BUS* stream. In all cases, DeNAS is able to achieve high quality satisfaction, while maintaining the minimum quality requirement.

V CONCLUSION AND FUTURE WORK

In this paper, we presented the DeNAS framework. By finding the appropriate scaling level for each video stream prior to its delivery over the wireless network, DeNAS is able to provide a high quality stream to each device, while increasing network efficiency and satisfying a greater number of clients than SVC. Our future directions include extending our scaling algorithm to support device prioritization and incorporating monitored channel conditions experienced by the devices.

ACKNOWLEDGEMENTS

The authors would like to acknowlege the support received from the Center for Wireless Communications, University of California at San Diego and the UC-Discovery Grant (com04-10175).

REFERENCES

[1] A. Vetro, C. Christopoulos, and H. Sun, "An overview of video transcoding architectures and techniques," *IEEE Signal Processing Magazine*, vol. 20, no. 2, pp. 18–29, 2003.

[2] J. Reichel, H. Schwarz, and M. W. (eds), "Scalable video coding - Joint draft 6, JVT-S201," Geneva, Switzerland, Apr. 2006.

[3] Y.-K. Wang, S. Wenger, and M. M. Hannuksela, "System and transport interface of H.264/AVC scalable extension," in *Proceedings of IEEE ICIP*, 2006, pp. 165–168.

[4] M. van der Scharr, Y. Andreopoulos, and H. Zhiping, "Optimized scalable video streaming over IEEE 802.11a/e HCCA wireless networks under delay constrants," *IEEE Trans. on Mobile Computing*, vol. 5, no. 6, pp. 755–768.

[5] G. Liebl, T. Schierl, T. Wiegand, and T. Sctockhammer, "Advanced wireless multiuser video streaming using the scalable video coding extensions of H.264/MPEG4-AVC," in *Proceedings of IEEE ICME*, 2006, pp. 625–628.

[6] "Wireless universal resource file," `http://www.wurfl.sourceforge.net`.

[7] S. Wolf and M. H. Pinson, "A new standardized method for objectively measuring video quality," *IEEE Transactions on Broadcast*, vol. 50, no. 3, pp. 312–322, 2004.