Robust Fine-Granularity-Scalability for Wireless Video

M. van der Schaar and J. Meehan
Philips Research Laboratories, Briarcliff Manor, NY 10510, USA

ABSTRACT

Fine-Granularity-Scalability (FGS) has recently been adopted by the MPEG-4 standard for the efficient and flexible distribution of multimedia over the Internet. The FGS framework is able to perform spatio-temporal-SNR tradeoffs in real-time at transmission or (post-) processing time, depending upon the available bit-rate, sequence characteristics, priority, user preferences or receiver resources, and not at encoding time. This paper investigates the usage of FGS for enabling the Universal Media Access (UMA) paradigm for wireless networks. To cope with the detrimental channel conditions, we introduce a novel joint source-channel coding scheme that combines FGS and adaptive modulation to improve the robustness of wireless video transmission. Two schemes are presented, one uses hierarchical modulation and FGS and the other uses time division multiplexing and FGS. Both schemes are shown by computer simulation to maintain a stable video quality over a variety of multipath dynamic Rayleigh fading channels. The adaptive modulated FGS algorithms provide a graceful degradation and improvement in picture quality as the channel conditions vary.

1. INTRODUCTION

In emerging wireless communications, multimedia data is streamed over various access networks (GPRS, UMTS, WLANs etc.) to a multitude of devices having different resource capabilities (display size, processing power, hardware support, memory etc). Hence, multimedia data can be accessed by a large number of users/clients at anytime, and from anywhere. This networking access paradigm is often referred to as Universal Multimedia Access (UMA). In the UMA framework, multimedia data is streamed (accessed) from the network depending on the following three parameters: user preference, communication channel characteristics and device capabilities [1].

For the implementation of the UMA framework, “universal” scalable video coding techniques are essential components [2]. Such coding schemes enable the scalable representation of video signals in the context of networks (bandwidth scalability), terminals (complexity scalability) and the combination of networks and terminals. For instance, “universal” scalable coded content can be transmitted/stored at a different video resolution depending on the receiving device display-size. Next, a different content can be streamed depending on the various hardware/software assistant blocks at the receiver. And, based on the user preference, a different frame-rate can be selected or a different object can be enhanced. Such a “universal” scalable video coding scheme is the Fine-Granularity-Scalability (FGS) [3] recently adopted by the MPEG-4 standard [4] for the efficient and flexible distribution of multimedia over the Internet. FGS consists of a rich set of video coding tools that support various scalability structures. The FGS-framework is able to perform real-time tradeoffs between the various SNR-temporal-spatial scalabilities *at transmission or (post-)processing* time, depending on the available bit-rate, sequence characteristics, priority, user preferences or receiver resources. Another important aspect in enabling the UMA paradigm is the cooperation of the QoS management over wireless networks with the application to support adaptation rather than insulating applications from network variations. Consequently, for an efficient and robust video transmission over wireless links, joint source-channel coding tradeoffs should be made depending on the communication channel characteristics.

In this paper, the usage of FGS for robust wireless video transmission is investigated. To cope with the bandwidth variations and detrimental channel conditions over the wireless networks, a novel adaptive modulated FGS scheme is introduced that takes advantage of the FGS structure to tailor the modulation scheme to the channel conditions. The combination of FGS and Adaptive Modulation is termed Adaptive Modulated FGS (AM-FGS). The results will show that using FGS in wireless networks provides robustness to losses and adaptability to varying network conditions, while also offering additional benefits such as flexible QoS management and channel adaptation.

The paper is organized as follows. In Section 2, the FGS coding tool standardized by MPEG-4 is briefly presented and its advantages for video transmission over wireless links are highlighted. The focus is on the flexibility of the FGS-framework that allows for real-time stream processing depending on the bandwidth and channel characteristics and on its inherent complexity scalability that enables easy resource adaptation depending on the capabilities of the mobile devices. Subsequently, in Section 3, two proposals for adaptively modulating an FGS signal are presented. The first one is based on hierarchical modulation, while the second employs time division multiplex modulation. Moreover, it is shown how the AM-FGS can be easily mapped onto packet-based WLAN standards, such as 802.11a and 802.11b. The results employed using the AM-FGS transmission schemes are presented in Section 4 and the conclusions are drawn in Section 5.

2. FGS FOR WIRELESS UNIVERSAL MEDIA ACCESS (UMA)

2.1 Requirements for UMA

The networking access paradigm known as UMA refers to the way in which multimedia data can be accessed by a large number of users/clients to view any desired video stream anytime and from anywhere. UMA echoes the wireless network driving force that is "Anywhere, Anytime and Anyone". In the UMA framework, multimedia information is accessed from the network depending on the following three parameters: channel characteristics, device capabilities and user preference.

- **Varying channel characteristics.** Currently, a convergence of POTS, cable and other wired networks with various wireless networks is emerging, providing a unified heterogeneous network that exhibits large bandwidth variations. For instance, while today's deployed GPRS networks support bandwidths up to 171.2kbps per user[1], EDGE networks will carry up to 384kbps, and UMTS networks are anticipated to support up to 2Mbps[2]. As more wireless technologies, such as 4G systems emerge, the bandwidth variations will be even higher due to the increasingly heterogeneous infrastructure. Thus, the bandwidth available to a Mobile Terminal (MT) can vary between 10kbps to several Mbps. There is a large bandwidth variability associated with the different types of indoor wireless networks also. For instance with an IEEE 802.11b system it is possible to have different modes whereby the maximum theoretical data-rates are 1, 2, 5.5 and 11Mbps. An IEEE 802.11a system can provide up to 54 Mbps, and at this time a backwards compatible extension to the current IEEE 802.11b system is being standardized whereby rates of greater than 22Mbps will be achieved.

Furthermore, the prevailing channel conditions, the mobility of the terminal, the transmission conditions external to the wireless network (for instance, congestion on the Internet when trying to access multimedia data) for both scenarios will have a large impact on the efficacy of the wireless network. Also, the throughput of an in-home network will depend on the demand on the system (number of MTs) as well as the range covered by a particular node (i.e. Access Point (AP) etc.). Consequently, to serve a broad range of data rates (e.g. from a few kbit/s to several Mbit/s) on heterogeneous networks having only one universal scalable encoding, combinations of spatial/temporal/SNR scalability with fine granularity become necessary. Selecting tradeoffs among these three dimensions (spatial/temporal/quality) becomes inevitable whenever supporting a large bandwidth and complexity range with high-quality is needed.

- **Different device capabilities.** The devices remotely accessing the wireless multimedia have various capabilities, for example display size, processing power, hardware support, memory etc. Thus, video content needs to be coded in a scalable fashion to match the capabilities of a variety of devices (complexity scalability) besides providing fast and easy adaptability to a particular bandwidth (bit-rate scalability). The advantage of universal scalable coding in the context of network and terminal scalability is that the content streamed becomes not only a function of the available network bandwidth, but also of the terminal complexity.

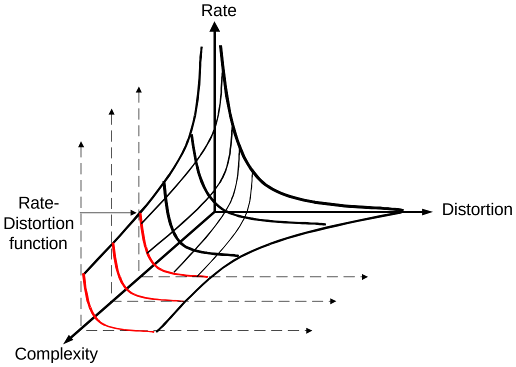

Figure 1: A typical Complexity-Rate-Distortion surface.

Bandwidth scalability enables ``bit-rate" scalable coding schemes to adapt to bandwidth variations by transmitting only the most important parts of the video stream, thereby performing tradeoffs between rate and distortion (see Fig. 1). Alternatively, complexity scalability can be defined as a transition between two operating points with different complexity coordinates. Note that complexity scalability is not the same as complexity optimization. In the latter, the goal is to decrease the complexity of a system/function with no loss of quality. Hence, complexity optimization attempts to move the operating point of a system in a rate-complexity plane, orthogonal to the distortion

---

[1] Assumes utilization of all eight available channels without error correction.

[2] 2Mbit/s for stationary devices, 144 kbit/s for moving devices and 384 kbit/s for pedestrian devices.

axis (see Fig. 1). Alternatively, complexity scalability trades distortion for complexity [5]. For illustration, in Fig. 1, a complexity scalable video coding algorithm is portrayed at various complexity levels.

- **User interactivity.** The streaming of content should depend on the user preference. For example, one user may wish to see a higher frame rate, while another prefers higher spatial resolution, and a third favors the fewest possible coding artifacts. This interactivity relates to the following tradeoffs: depending on user selection, different combinations of available scalable streams would be invoked, meaning that hybrid combinations of scalable modes should be allowed.

Summarizing, in the UMA context, the encoding process should be decoupled from the network and the terminal capabilities, since they are only known at transmission and decoding time. Furthermore, adaptive streaming of content should be allowed based on the viewer preference. With a Universal Media Codec (UMC), such as FGS, the video data is coded once for a whole range of access networks and terminals. Next, the bandwidth and complexity scalabilities of FGS are briefly highlighted.

2.2 Brief description of the FGS coding tool

FGS has recently[3] been adopted by MPEG-4 as a video-coding tool for streaming applications. Subsequently, the FGS framework is only briefly presented. (For a more detailed description of the FGS framework, the interested reader is referred to [3]). In addition to the FGS base layer, which is coded with an MPEG-4 compliant non-scalable coder, FGS consists of a single enhancement-layer coded in a progressive (fine-grained) manner (see Fig. 2). The base layer is coded with a bit-rate R_{BL}, chosen so that the available bandwidth (over the time-varying network) is higher than R_{BL} at all times ($R_{BL} \leq R_{min}$). Subsequently, the enhancement-layer is over-coded at encoding-time using a bit-rate ($R_{max}-R_{BL}$), as portrayed in Fig. 2. The enhancement-layer is then progressively (bit-plane by bit-plane[4]) coded by employing a low-complexity bit-plane embedded-DCT algorithm. As can be seen from Fig. 2, the enhancement-layer frames are intra-coded, but the coding efficiency from temporal redundancy exploitation is partially retained because the MPEG-4 motion compensated scheme is employed at the base layer.

The FGS coding scheme standardized by MPEG-4 also contains FGS temporal-scalability [6] in addition to the previously described FGS SNR-scalability. Moreover, FGS spatial-scalability[5] can also be easily added [7]. Hence, using different sub-sets of the same pre-encoded FGS scalable bitstream, tradeoffs can be made in the SNR, temporal and spatial–dimensions based on the channel characteristics, device capabilities and user preference.

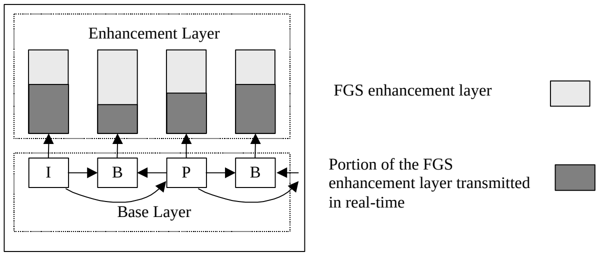

Figure 2:Example of the FGS scalability for a typical unicast streaming application.

At the streaming server, the enhancement-layer improves upon the base layer video, fully utilizing the bandwidth available at transmission-time. The FGS scalability structure allows for resilient video transmission, as long as the base layer video is reliably delivered, since the packet- or bit-losses in the enhancement-layer do not propagate.

For multicast applications, FGS also provides a flexible framework for the encoding, streaming, and decoding processes. Identical to the unicast case, the encoder compresses the content using any desired range of bandwidth [$R_{min}=R_b$, $R_{max}=R_e$]. Therefore, the same compressed streams can be used for both unicast and multicast applications. At the time of transmission, the multicast server partitions the FGS enhancement layer into any preferred number of "multicast channels" each of which can occupy any desired portion of the total bandwidth (see Fig. 3). At the decoder side, the receiver can "subscribe" to the "base layer channel" and to any number of FGS enhancement-layer channels that the receiver is capable of accessing depending on the receiver access bandwidth. Moreover, various unequal error protection strategies can be used for the different layers depending on their contribution to the overall visual quality. For instance, as will be shown in Section 3, different modulation techniques can be adopted for the various FGS layers to enable robust transmission of video over heterogeneous wireless links. Note that the various FGS enhancement-layers can also represent spatial and/or temporal quality improvements with respect to the base-layer, besides SNR

---

[3] FGS became an International Standard in March 2001.

[4] In a progressive coder, the more significant bit-planes are transmitted prior to the less significant bit-planes.

[5] Note that FGS spatial-scalability is not yet part of the MPEG-4 Standard.

improvements. Thus, in a multicasting scenario, each receiver could receive a different quality video signal, depending on the capabilities[6] of the receivers (e.g. display size, CPU, co-processors, memory, channel conditions etc.). For example, assume that two devices access the same "generically" FGS coded content. The first device is a low-power mobile phone with a CIF-resolution display and will only extract a compressed bit-stream that represents a CIF base-layer and a quality enhancement-layer compressed at 10 frames/s. Whereas the second device is a powerful fixed flat-screen DTV that will be able to stream and decode an SD-resolution video at 50 frames/s.

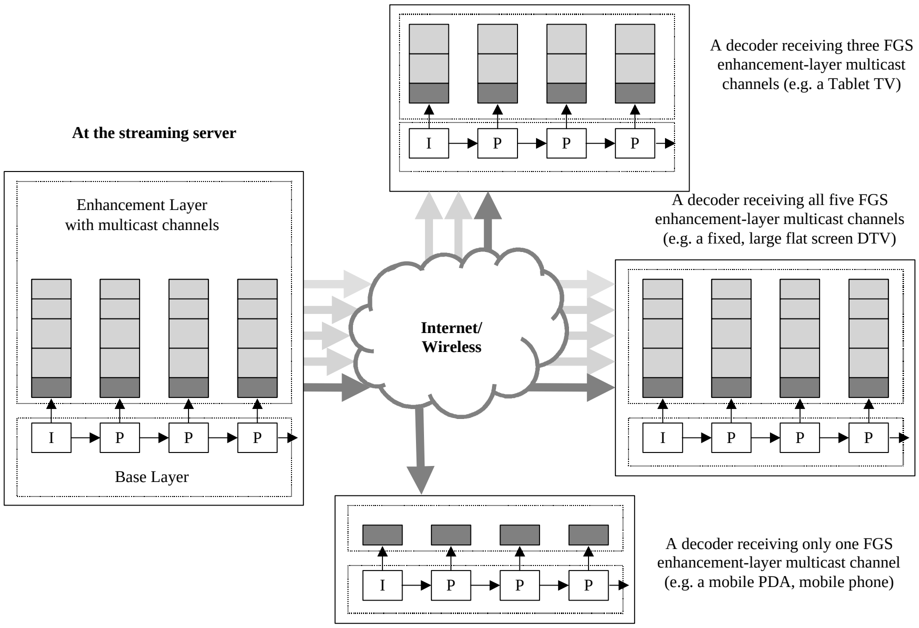

Figure 3: An example of an FGS-based multicast scenario.

Consequently, the advantage of FGS in the context of network and terminal scalability is that the content streamed becomes not only a function of the available network bandwidth, but also of the terminal complexity. In this context, while the bandwidth scalability property of FGS and its performance have been highlighted in numerous studies, the FGS complexity scalability property has not been studied. To illustrate the complexity scalability of FGS, the decoding time on a Pentium III at 500 MHz is depicted in Fig. 4 for the MPEG-4 *Foreman*-sequence at CIF-resolution 10Hz at a variety of transmission bit-rates. Fig. 4 highlights FGS property to be both complexity[8] and bandwidth scalable, as required by UMA (see Figure 1).

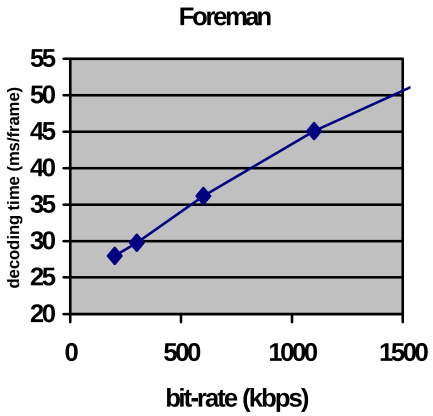

Figure 4: FGS complexity scalability as a function of the transmission bit-rate.

Summary of FGS benefits for wireless streaming. While FGS has been originally designed for Internet video

---

[6] The device capabilities can be signaled to the server by a dedicated communication protocol such as those currently defined in MPEG-21.

[8] Note however that a more in-depth study is necessary to establish FGS complexity scalability for a variety of software and hardware platforms, as required by UMA.

transmission, it has also many advantages for wireless video streaming:

- FGS enables a streaming server to perform minimal real-time processing and rate control when outputting a very large number of simultaneous unicast (on-demand) streams (unlike transcoding or simulcast approaches).
- FGS is highly adaptable to unpredictable bandwidth variations due to heterogeneous access-technologies (e.g., GPRS, UMTS, 802.11) of the receivers or due to dynamic changes in network conditions.
- FGS is able to support both multicast and unicast applications. This, in general, eliminates the need for coding content in different formats to serve different types of applications. Moreover, for multicast applications, the scalable coded streams require less bandwidth for transmission.
- FGS is resilient to packet and bit-error losses, which are quite common over the Internet and wireless networks.

FGS allows scalable-complexity decoding, thereby providing low-power devices (e.g., mobile phones) and everyday receivers (e.g., set-top-boxes and digital televisions), in addition to powerful computers, the opportunity to stream and decode any desired video content with different quality.

2.3 Adaptive QoS management

For video applications like real-time video streaming and video-conferencing, a very strict set of QoS requirements needs to be satisfied. Hence, a certain QoS level needs to be negotiated for each individual session (connection), called a Service Level Agreement (SLA). A certain SLA can define service parameters like precedence, reliability, delay and throughput [8]. Using these SLA parameters, the MT negotiates a specific SLA with the network, or chooses from a set of available SLAs when initiating the session. However, any number of conditions can cause the network to be unable to meet the QoS agreement. These conditions are primarily the prevailing channel conditions, mobility and transmission conditions that are external to the wireless network (i.e. congestion on Internet when trying to access multimedia data). This is in sharp contrast with the traditional distributed applications, where a stable presence and a consistently high network quality are possible. Hence, an unqualified QoS guarantee is impossible and it is thus the *variation in QoS* that forms the difference between wireless and wired networks. This implies that for wireless networks, an adaptive QoS management is more suitable, that specifies a range of acceptable QoS (SLAs) rather than trying to guarantee specific values.

The previously mentioned QoS functions can be supported very easily by the FGS framework that allows for prioritized transmission, different levels of resilience, different throughputs etc. For instance, given a certain channel condition, tradeoffs can be made between higher transmission bit-rates (i.e. more data transmitted through less reliable channels) and a more resilient transmission (at lower bit-rates), by switching between the baseline FGS structure (that does not employ any motion-compensation within the enhancement-layer) and the Motion Compensation FGS (MC-FGS) structure in [9] (that does employ motion-compensation in the enhancement-layer).

Furthermore, in cellular transmission, a different number of channels can be allocated to the same MT, depending for instance, on the number of MTs within the wireless cell. FGS bitstreams can be easily partitioned in prioritized sub-streams (classes) that contribute incrementally to the obtained visual quality. These sub-streams can then be transmitted through various “traffic” channels, with various QoS guarantees. As the number of channels allocated to an MT varies dynamically, the number of sub-streams transmitted to it is smoothly and instantaneously adapted. Such an adaptation cannot be achieved with non-scalable coding schemes, since the data cannot be easily partitioned in a number of discrete sub-streams whose transmission is decided “on the fly” depending on the available number of traffic channels. Employing a scalable coding technique like FGS, allows for easy admission control whenever a new MT enters a cell, while preserving the SLAs of the existing MTs. In this case, easy joint-quality control can be performed among the various video data streams that are transmitted to the various MTs within a cell, such that the quality of the previously existing MTs is degraded gracefully as a result of admitting a new data flow. Moreover, in this manner, various priorities can be assigned to different application classes. Thus, the employed video coding scheme for multimedia streaming needs to be able to adapt to on-the-fly bandwidth variations that are necessary in order to accommodate the bandwidth required for the higher priority applications. Hence, FGS is able to cooperate with the QoS management over wireless networks to support adaptation, thereby achieving a higher transmission quality of video over wireless channels.

Another important aspect in the UMA context is that future wireless video applications will have to work over an open, layered, Internet-style network with a wired backbone and wireless extensions. Therefore, common protocols like the Internet Protocol (IP) will have to be used for the transmission across both the wired and wireless portions of the network to provide seamless roaming over various heterogeneous networks. However, the standard IP-based networks provide “best-effort” data delivery that is not suited for delay sensitive transmission like video streaming. Hence, to provide QoS for specific applications, the IP services must be enhanced with QoS protocols to distinguish between traffic with strict timing requirements (real-time video) from those that can tolerate delay, jitter and loss (e.g. file transfers). Thus, the IP-based QoS protocols (e.g. RSVP, DiffServ) do not create or guarantee bandwidth, but manage it more effectively to meet the wide range of application requirements. It is important to mention that the IP-based QoS is end-to-end, between the server (that can be located on the Internet) and the receiver (that can be located on an in-home wireless network), as depicted in Fig. 5.

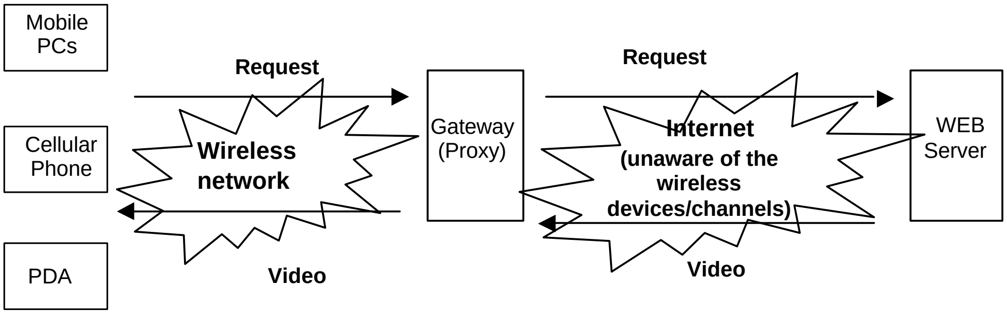

Figure 5: Post-processing of the scalable-coded content at the wireless gateway based on the wireless devices characteristics and channel conditions.

For an efficient transmission, the scalable-coded content should be handled at a “wireless gateway” or a “wireless proxy” based on the knowledge of the MT capabilities (see Fig. 5). For example, the bandwidth through the Internet backbone or core cellular network is large but only a limited bandwidth is available in a specific wireless sub-network or to a certain MT. Therefore, manipulation of the streams at the proxy will be necessary to meet the QoS requirements. These “wireless proxies” or “wireless gateways” can be located in a variety of different positions in the network. For instance when looking at the cellular environment, they could be located at the Base Stations (BSs), or at other points in the delivery network. A second location where one can easily envisage the existence of a “wireless proxy” is at the residential gateway, where it can process the content according to the capabilities of the various devices present in the home. The advantage of locating this proxy in the residential gateway is that proprietary algorithms can then be added and the capabilities of certain devices can be enhanced, by exploiting a common interface. Processing can also be distributed across several proxies, both within and outside the home. The methods proposed in the subsequent section for the robust transmission of FGS are located at these wireless gateways.

Furthermore, to ensure QoS for specific applications, each OSI layer from the application down must support QoS to ensure prioritized sending and receiving requests. For instance, certain data-link layer W-LAN technologies, are already QoS enabled, such that higher priority packets receive higher priority treatment as they traverse the network media. Thus, to provide seamless and robust network connectivity to the various wireless devices within the home, the various network layers (e.g. Application, IP, Link, MAC, Physical) need to cooperate together.

3. FGS AND ADAPTIVE MODULATION

In this section, we will show that FGS is well suited to joint source-channel coding schemes and can be successfully applied to provide robust transmission of video over wireless networks. The FGS video data is split at encoding time into multi-priority streams (Base Layer (BL) and multiple Enhanced Layers (EL)) that can cover various channel characteristics and terminal capabilities. The BL has the highest importance, followed by the Most-Significant-Bitplane (MSB) of the FGS EL, MSB-1 of the FGS EL etc. This prioritization of the FGS video data can be used by the physical layer to adapt the modulation scheme. Two proposals for adaptively modulating an FGS signal are to use hierarchical modulation or to use time division multiplex modulation and these are outlined in the following sections. The proposed methods can be applied in both cellular and home networking domains to provide a very reliable, robust, gracefully degradable wireless video stream for unicasting and multicasting.

3.1 Hierarchically Modulated FGS (HFGS)

The joint consideration of source-coding schemes, channel coding and modulation has already been proposed several decades ago for the transmission of images [10]. Also, hierarchical modulation has been applied to High Definition TV (HDTV) broadcasting in [11]. More recently, adaptive modulation schemes have been presented for simultaneous voice and data transmission over fading channels [12] and in [13], adaptive modulation has been employed for the multi-layered transmission of H.263+ scalable coded streams over wireless channels.

In our scheme, the Hierarchically Modulated FGS (HFGS) takes advantage of the FGS structure to tailor the modulation scheme to the challenges of wireless video transmission. The FGS layers are transmitted simultaneously onto a hierarchical modulated constellation. The example shown here will be for a Quadrature Amplitude Modulation

(QAM) constellation, but other constellations an also be employed. The HFGS scheme is shown in Fig. 6. The video data is coded using FGS, into a BL and a set of ELs. Each of these layers are independently error correction coded (ECC). Subsequently, within the Symbol Mapper module, the BL data can be mapped to the MSBs of the transmitted signal and the EL can be mapped to the LSBs. If feedback (such as the SNR or BER for the system or for each FGS layer) from the transmitter or client exists, this can be used to modify dynamically the Symbol Mapper, the rate of error correction coding, or the source coding rate/division.

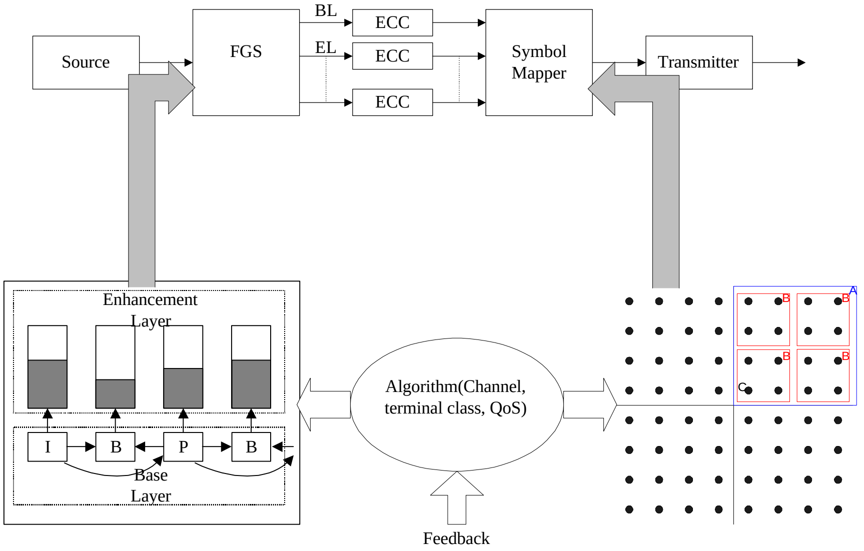

Figure 6: Graceful Wireless Video Transmission using Hierarchically-modulated FGS.

In Fig. 6, a multi-sized constellation is shown. This QAM constellation can be subdivided: the 64 QAM constellation in Fig. 6 is subdivided into a 16 QAM constellation and a 4 QAM constellation. The MSBs of a QAM symbol define the minimum QAM constellation and can be detected at the receiver more robustly than higher level QAM constellations. This is also known as hierarchical modulation. The number of bits required and the constellation sizes are listed in Table 1.

Table 1: Bit-representation per constellation for HFGS.

No. Bits	Bit Representation				Constellation
	BL	EL1	EL2	EL3	
2	A1 A0				4QAM
4	A1 A0	B1 B0			16QAM
6	A1 A0	B1 B0	C1 C0		64QAM
8	A1 A0	B1 B0	C1 C0	D1 D0	256QAM

Assume for example that receiver A is only able to receive the base layer (in this case the base layer is mapped onto 2 bits (A1 A0) – 4QAM). Receiver B can only receive one enhanced layer (total number of bits is 4 (A1 A0 B1 B0) – 16QAM). And receiver C can receive 2 enhanced layers (number of bits is 6 (A1 A0 B1 B0 C1 C0) - 64QAM). In this case, each layer has a bit-rate of 2*x* bit/s, where *x* is the number of transmitted symbols per second. From Fig. 6, Receiver A would slice on A square, Receiver B on the B squares and Receiver C on each point in the constellation. Assuming that the distance between each symbol in the constellation is *d*, the average distance between each B region is 2*d* and between each A region is 4*d*. Hence, in Fig. 6, each layer is approximately 3 dB more robust than the next, thereby providing unequal error protection for the various FGS layers. Note that if the channel degrades, receiver C will receive the MSBs more reliably than the LSBs even though it is slicing on the 64QAM constellation. For example, Receiver C could slice based on constellation A first, then slice on constellation B and then finally on C, depending on the reliability of each constellation.

The presented HFGS algorithm is ideally suited for robust video transmission over cellular networks due to its spectral efficiency (i.e. HFGS is bandwidth efficient). Results for this algorithm will be shown in Section 4.1.

3.2 Time Division Multiplexing of FGS (TD-FGS)

In this mode, the FGS layers are transmitted in a Time Division Modulation (TDM) scenario, where each FGS layer is allocated an independent time slot. In Time Division FGS (TD-FGS), the FGS layers are mapped onto different constellations. The ELs can be mapped onto 64 QAM constellations for their time slots and the BL can be mapped onto a more robust QPSK constellation. Hence, in TD-FGS the bit-rate for the BL is lower than the EL bit-rate. Moreover, as for HFGS, the channel coding can differ for each FGS layer to provide unequal error protection. The BL and EL are multiplexed onto the same data stream (see Fig. 7). The receiver must know whether the transmitted signal is BL or EL, such that it can switch the receiver constellation in the decision device (slicer) to the appropriate constellation. TD-FGS is more robust than the simultaneous transmission proposal but requires more bandwidth and receiver knowledge of the data stream layout.

BL@6Mbps	EL$_1$@12Mbps	EL$_2$@24Mbps	- - - - - - -	BL@6Mbps	EL$_1$@12Mbps

Figure 7: Mapping TD-FGS onto 802.11a Packets.

The proposed TD-FGS can be easily mapped onto the current packet-based WLAN standards (i.e. 802.11a and 802.11b standards). These standards employ packet headers indicating the modulation scheme and rate of the packet data. For example, the 802.11a standard consists of 8 Physical layer modes (6Mbps, 9Mbps, 12Mbps, 18Mbps, 24Mbps, 36Mbps, 48Mbps, and 54Mbps). Thus, the BL can be mapped into the most robust 6Mbps packets and ELs can be mapped onto higher rate packets. This gives increased flexibility to the TD-FGS scheme over other schemes, since the robustness of the ELs can be easily traded off for higher EL bit-rates. Furthermore, it should be noted that the FGS-framework can produce an enhancement-layer that can allow for even more discrete prioritization than only between BL and EL or within the SNR-scalable EL itself. For instance, dependent on the channel characteristics or terminal capabilities, a lower or higher priority can be assigned to the FGS temporal frames. Hence, by exploiting the flexibility of the FGS-framework to optimize the interaction between the FGS-layer and the physical-layer, a more efficient and robust transmission over wireless transmission channels can be realized as will be shown in the results section.

3.3 Example of AM-FGS application

A typical application area for the previously proposed AM-FGS is in home wireless video. In Fig. 8, a set-top-box (STB) connected to a rooftop antenna re-broadcasts the video stream inside the house. Assuming that the signal to the rooftop antenna is a line-of-sight (LOS) signal and assuming a high SNR channel, many bits per symbol can be used, resulting in a high quality video transmission to the home. On the other hand, the in-home environment is typically heterogeneous - the physical wireless channel can vary and also the various devices have different capabilities requiring different video qualities (such as a PDA vs. a TV in the kitchen/bed room vs. a TV in the living area). Consequently, the incoming constellation can be easily mapped using AM-FGS to a more robust constellation for in-home transmission. Subsequently, a single FGS video stream can be transmitted to multiple devices that have different robustness and qualities. For instance, if the same video stream is transmitted to both a large screen TV (at a fixed location) and a portable small screen TV, the portable TV will only receive the robust BL video stream (but the user can move within the home), while the fixed TV will receive the BL and a larger portion of the EL (but its location is fixed in the proximity of the STB). Note that all terminals in this scenario use the same AM-FGS algorithm. This is a major advantage that shows the scalability and flexibility of AM-FGS. This re-broadcasting scenario (ad-hoc network) is applicable to both cellular and Wireless LAN systems.

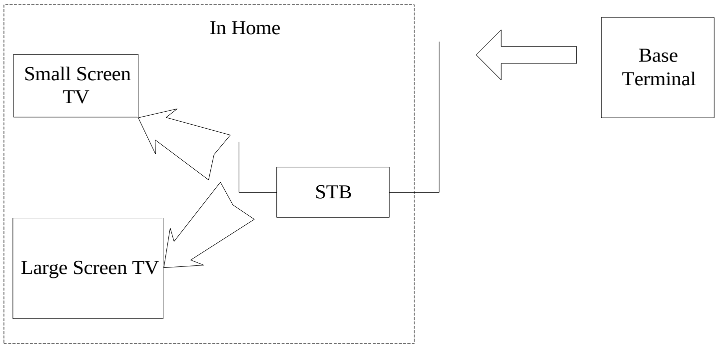

Figure 8: A typical adaptive re-broadcasting application.

4. RESULTS

4.1 Results HFGS

The HFGS technique has been implemented for a BL bit-rate R_{BL} equal to half the EL bit-rate R_{EL}. In our simulations, R_{BL} equals 100kbit/s and the adopted packet sizes equal 500 bytes. The modulation scheme resulted in a Packet Loss Rate (PLR) of PLR_{BL} for the BL packets and a higher $PLR_{EL} = 50\ PLR_{BL}$ for the EL packets. (Note that in this case the entire FGS EL is treated the same, independent of its visual importance. However, improved results can be obtained by employing different modulation techniques for the various ELs, thereby providing fine-grained loss protection for FGS.) The results obtained with and without HFGS are given in Table 2 for two well-known MPEG-4 sequences at CIF-resolution, 10 Hz, for various values of PLR_{BL} and corresponding PLR_{EL} values respectively. From this table the following conclusions can be drawn:

- The newly introduced HFGS technique results in considerably better visual quality performance compared with the non-hierarchically modulated scheme.
- The improvement introduced by HFGS is especially visible for higher PLR.
- The quality of the FGS scheme under packet-losses decreases abruptly if HFGS is not used, while in the HFGS case, the visual quality is gracefully degraded.

Table 2: PSNR performance for various sequences under different channel conditions with and without HFGS.

Sequence	PLR_{BL}	PSNR (dB) FGS – no HFGS	PSNR (dB) HFGS
Foreman	2×10^{-4}	31.0	30.7
	10^{-3}	25.2	30.4
	2×10^{-3}	22.3	29.9
Coastguard	2×10^{-4}	28.2	27.9
	10^{-3}	25.9	27.7
	2×10^{-3}	23.2	27.0

4.2 Results TD-FGS

The previously introduced TD-FGS has been implemented and the visual quality results (expressed by the PSNR) obtained by using different modulation techniques for the FGS layers are given in Table 4 for two sequences with various characteristics – *Football* and *Woman* - at SD resolution (i.e. 720x576 pixels), 50 Hz. The results are given for two different channel conditions with SNR of 18dB and 25dB. In Table 3, the PLR for the two channel conditions are given for several constellations providing different levels of robustness for the various layers. From Table 4, the following conclusions can be drawn:

- As expected, the performance of FGS improves with the channel conditions. However, the performance improvement is especially visible for high enhancement-layer transmission bit-rates (e.g. R_{EL}=9Mbit/s). At low transmission bit-rates, the various FGS layers can be protected very well, and thus the difference in visual quality for the two channel conditions is limited.
- Dependent on the channel conditions a different joint source-channel coding tradeoff should be made to achieve the best visual quality under losses. For the poor channel conditions (SNR = 18dB), the best visual result is obtained by sending two layers, the base-layer of 1.5Mbit/s and the enhancement-layer of 3Mbit/s. Alternatively, for the better channel conditions (SNR = 25dB), the best performance is obtained for a base-layer of 1.5Mbit/s and an enhancement-layer of 9Mbit/s.
- The best tradeoff between source and channel coding for the different channel conditions gives relatively similar results for the various investigated sequences independent on their characteristics (texture and motion activity). This result differs from those obtained for non-scalable coding where the sequence characteristics considerably influence this tradeoff and it has the advantage that the joint source-channel coding tradeoffs can be made offline for a specific set of channel conditions, adaptive modulation techniques, transmission bit-rates and sequence resolutions and frame-rates. This desirable FGS property is due to its lack of motion compensation within the enhancement-layer that prevents error propagation and due to its fine-granularity, that allows to easily prioritized the importance of the various quality layers.

Table 3: Various layers bit-rates and their corresponding PLR for the different TD-FGS scenarios.

Scenarios	Bit-rates (Mbps)	PLR1 (18db)	PLR2 (25db)
1	BL: 1.5	10^{-4}	10^{-5}
2	EL: 3	10^{-3}	$2*10^{-5}$
3	EL: 9	$3*10^{-1}$	$3*10^{-2}$

Table 4: PSNR performance under various channel conditions using various TD-FGS techniques.

Sequence	Scenarios	PSNR (dB) for PRL1	PSNR (dB) for PRL2
Football	1	30.90	30.95
	2	33.04	33.26
	3	31.30	33.88
Woman	1	34.75	34.80
	2	37.94	38.25
	3	35.58	39.36

5. CONCLUSIONS

In this paper, we investigated the usage of FGS as a Universal Media Codec that enables the UMA wireless transmission paradigm. In this context, we highlighted the complexity scalability of FGS besides its well-known bandwidth scalability. Furthermore, we concluded that while most known scalable video coding schemes address scalability in only three dimensions - temporal, spatial and SNR, enabling scalability in a fourth dimension, complexity, is necessary for UMA. Subsequently, we proposed a novel joint source-channel coding scheme that combines FGS and adaptive modulation for wireless transmission. AM-FGS exploits the FGS property of compressing video data in a set of multi-priority streams based on the channel characteristics and terminal capabilities by adapting the modulation scheme in the physical layer. AM-FGS is equally applicable to the cellular and home networking domains. We also highlight that additional benefits exist for using FGS for wireless video such as QoS management and channel adaptation and allocation.

REFERENCES

[1] “Study on MPEG-21 (Digital Audiovisual Framework)”, Part 1, version 2.0, MPEG output document N4040, March 2001.

[2] J. Ohm, W. Li, M. van der Schaar *et al*, "Summary of Discussions on Advanced Scalable Video Coding", Contribution to MPEG - M7016, March 2001.

[3] H. Radha, M. van der Schaar, Y. Chen, “The MPEG-4 Fine-Grained Scalable Video Coding Method for Multimedia Streaming over IP”, *IEEE Trans. on Multimedia*, vol. 3, no. 1, March 2001.

[4] “Study of ISO/IEC 14496-2:1999/FPDAM4”, MPEG output document N3670, La Baule, France, USA, Oct. 2000.

[5] D. Sow, ``Computational Resource Scalability - A theoretical framework", Philips Technical Report, Nov. 1999.

[6] M. van der Schaar, H. Radha, “A hybrid temporal-SNR Fine-Granular Scalability for Internet Video”, *IEEE Trans. of Circuits and Systems for Video Technology*, vol. 11, no. 3, March 2001.

[7] M. van der Schaar, ``All FGS temporal-SNR-spatial scalability", Contrib. to MPEG - M6490, La Baule, France, Oct. 2000.

[8] ETSI, “Universal Mobile Telecommunications System (UMTS); QoS Concept and Architecture”, ETSI TS 123 107 V3.5.0, Release 1999, Dec. 2000.

[9] M. van der Schaar, H. Radha, ``Motion-Compensation based Fine-Granular Scalability (MC-FGS)", M6475, France, Oct. 2000.

[10] J.W. Modestino, D.G. Daut, A.L. Vickers, “Combined source-channel coding of images using the block cosine transform”, *IEEE Trans. on Comm*., COM-29, 1981.

[11] K. Ramchandran, A. Ortega, K. M. Uz, and M. Vetterli, “Multiresolution broadcast for digital HDTV using joint source-channel coding,” *IEEE J. Select. Areas Commun.*, vol. 11, pp. 6–23, Jan. 1993.

[12] M. Alouini, X. Tang, A. Goldsmith, “An Adaptive Modulation Scheme for Simultaneous Voice and Data Transmission over Fading Channels”, *IEEE Journal on Sel. Areas in Comm. (JSAC)*, vol. 17, no. 5, May 1999.

[13] Y. Pei, J.W. Modestino, “Multi-layered video transmission over wireless channels using an adaptive modulation and coding scheme”, *Proc. of ICIP 2001*, vol. 2, Oct. 2001.