ADAPTIVE POWER SAVING ON THE RECEIVER SIDE IN DIGITAL VIDEO BROADCASTING SYSTEMS BASED ON PROGRESSIVE VIDEO CODECS

Eugeniy Belyaev[1] Timo Koski[2] Jarkko Paavola[2] Andrey Turlikov[1] Anna Ukhanova[1]
[1]Saint-Petersburg State University of Aerospace Instrumentation, Russia
[2]University of Turku, Finland

ABSTRACT

This paper discusses the utilization of progressive video codecs in digital video broadcasting. The benefit is two-fold. Progressive codec would enable the possibility to receive a single broadcast stream with variety of screen sizes ranging from HDTV televisions to cellular phone screens in mobile TV application. Another benefit is the possibility for trade-off between video quality and power saving in the receiver. This is especially important in battery powered hand-held receivers.

Here, the mobile broadcast standard DVB-H along with JPEG2000 for video encoding is used to show how to make it possible to allow receiver to control the level of power consumption. Actually, any other scalable codec could be used instead of JPEG2000, which was chosen just to illustrate this possibility. In this paper, JPEG2000 is also compared with H.264/AVC and MPEG-2 in the power consumption and shows quite favorable results. In future work, authors will analyze the scalability of the proposed system in detail.

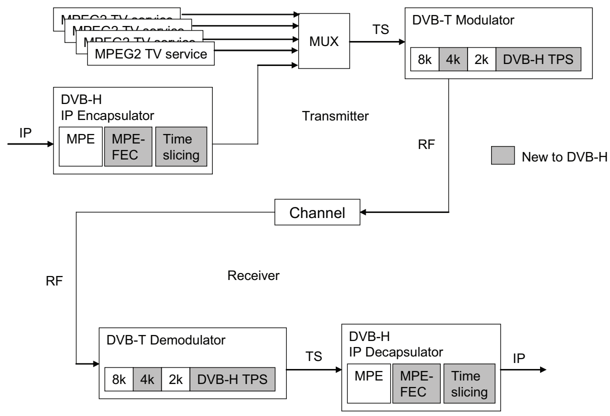

Figure 1: A conceptual description of the DVB-H system [1]

I INTRODUCTION

The era of wireless digital video broadcasting (DVB) started in the beginning of 90's when satellite (DVB-S) and terrestrial (DVB-T) systems were launched. In November 2004 the European Telecommunications Standards Institute (ETSI) ratified the DVB-H standard [1], which is an amendment of DVB-T [2] for handheld terminals.

One of the main goals in DVB-H development was the idea of minimizing receiver power consumption by time-slicing, where transmission occurs in bursts. The main idea of it is that receiver operates for a short time interval, and during it the part of the video data is received. Then the receiver powers off radio parts completely while video is constantly decoded. The received data also contains information about when to power on again to receive next part of the data. The degree of power consumption depends on the parameters such as peak bitrate, burst bitrate and service bitrate.

In the existing DVB-H systems the MPEG-2 [3] transport stream is used transmit H.264/AVC [4] encoded video content. There, fixed parameters peak bitrate and burst bitrate define the level of power consumption. Thus, it does not allow to control the level of power consumption at the receiver side.

In this paper it is proposed to use the progressive codecs to provide adaptive method to perform trade-off between power consumption and video quality in mobile handheld receiver. Another benefit would be that a single broadcast could be used to transmit video stream to screens, which have broad range of different resolutions. For example, from HDTV screens to QCIF screens utilized in mobile tv reception in handheld terminals. For future DVB systems it would be highly beneficial to have only one system that would cover terrestrial and mobile transmission instead of separate systems that must be used nowadays.

As an example, two mature technologies are utilized to demonstrate the benefit of utilizing progressive video codes in digital video broadcasting. DVB-H along with JPEG2000 standard [5] for video encoding allows receiver to control the level of power consumption depending on the priorities. In the case of H.264/AVC or MPEG-2 video stream the type of power consumption is set by the codec, but in the proposed case the receiver is responsible for the type of power consumptions. For example, either if there is almost no charge in the receiver, or the transmission of the video data will last for a long time, or the high level of power saving is needed, the receiver can switch to the energy economy work conditions. But if there is no need for power consumption, receiver could work with the highest quality of the video data.

The paper is organized as follows: next the DVB-H system presented and its power saving possibilities are discussed. Then JPEG2000 standard, and its rate control capabilities are discussed in sections III and IV. In Section V scheme for possible implementation of JPEG2000 in DVB-H is presented. Practical results are described in Section VI. Finally, conclusions about the power saving are given.

II DVB-H STANDARD DESCRIPTION

A detailed overview of DVB-H can be found from [6]. To allow mobile reception, changes to terrestrial system were needed to enable low power consumption, more flexibility in network planning, better mobile performance and compatibility with IP networks. These were achieved by adding Multi-Protocol Encapsulation (MPE), time-slicing, error correction and a new 4K FFT OFDM mode in addition to the 2K and 8K modes in DVB-T. Time-slicing operation means that transmission occurs in bursts. Power-saving feature occurs due to the fact, that receiver can switch off radio components between bursts.

Conceptual diagram of the physical and link layers of a DVB-H system is illustrate in Figure 1. The physical layer consists of the DVB-T modulator and demodulator. Operations in DVB-H link layer consist of the IP encapsulation, time slicing and optional forward error correction. In the transmitter, IP packets from the network or application layer are encapsulated to MPE-FEC frames and encoding and time slicing is performed.

The size of the MPE-FEC (Multi-Protocol Encapsulation - Forward Error Correction) frame is service independent. The IP datagrams are encapsulated column-wise into the MPE-FEC frame and the data are encoded row-wise using RS(255,191) code. Different MPE-FEC code rates are achieved with code shortening and puncturing. If forward error correction is not used, only Multi-Protocol Encapsulation is performed and the whole time-slice burst is filled with IP data carried in MPE-sections.

The frame is divided into sections so that an IP datagram forms the payload of an MPE-section and a redundancy column form the payload of an MPE-FEC-section. When the section header is attached, the CRC-32 redundancy bytes are calculated for the section. The MPE-sections are transmitted first and then the MPE-FEC-sections. The sections are transmitted in a MPEG-2 transport stream format. The more detailed descriptions of link layer operations can be found from [7, 8]. The duration of the burst carrying the MPE-FEC frame can be changed by altering the bit rate used in the transmission. Using lower bit rate introduces longer time interleaving in the MPE-FEC. On the other hand, the power saving in the terminal is reduced, since the off time between the bursts decreases if the bit rate of the service remains the same. The effect of burst duration is analyzed for example in [9].

Time slicing operation is illustrated in Figure 2. It can be described with following parameters: Burst Size B_s refers to the number of bits within a burst; Burst Bitrate B_b is the bit rate used by stream inside a burst. Constant Bitrate C_b is the average bitrate required by the stream when it is not time-sliced.

Power saving calculations are presented in [10, 11]. Burst duration B_d is calculated as $B_d = B_s/(0.96B_b)$ (assuming 4% overhead from packet and section headers [10]). Off-time O_t is the time between bursts $O_t = B_s/(0.96C_b) - B_d$. Then, taking into account synchronization time S_t and Delta-t (the time instant before the second burst cannot occur) jitter D_j, the power saving P_s can be given as

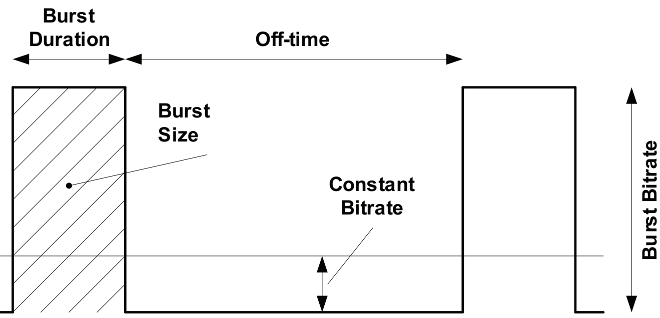

Figure 2: Time slicing in DVB-H [10]

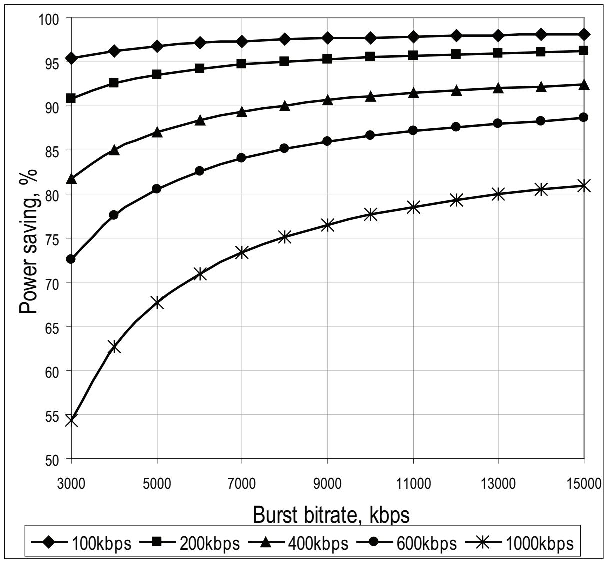

Figure 3: Relation between bit rate and power saving

$$P_s = \left(1 - \frac{C_b}{B_b} - 0.96\frac{C_b}{B_s}\left(S_t + \frac{3}{4}D_j\right)\right) \cdot 100\%. \qquad (1)$$

Relation between bit rate and power saving for different constant bit-rates is shown in Figure 3.

III SINGLE-SOURCE RATE-CONTROL ALGORITHM IN JPEG2000

The general scheme of the algorithm JPEG2000 [12] is described below. At first, color space transformation is applied to the frame-component data. Then a discrete wavelet transformation is used to decompose each frame into a hierarchy of subbands. Each component is represented as a collection of resolution levels, and each level consists of some subbands. Each subband is further partitioned into code blocks, and entropy coding is performed on each code block independently. The entropy coding scheme that is performed in so called Tier-1 part in JPEG2000 is based on the classical context-adaptive arithmetic coder. Each bitplane in each codeclock (except the most significant non-zero bitplane from which the coding starts) is encoded in the three passes. The bitstream generated by each pass is referred to as coding pass. After entropy coding a rate

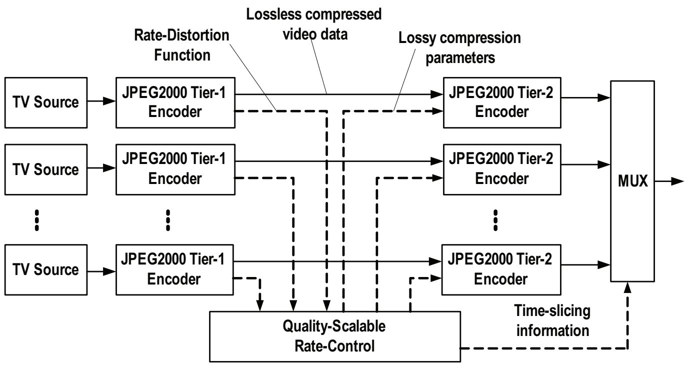

Figure 4: DVB-H Transmitter scheme based on JPEG2000 Encoder

control mechanism could be applied. Then all data is sent to Tier-2 part of the codec to be divided into packets and sent to the channel.

Rate control in JPEG2000 guarantees that the resulting frame size is not higher than the given threshold value providing image quality is as high as possible. The frame is composed of an integer number of code blocks with numbers $i = 1, 2, 3, \ldots$. Each block is subsequently composed of an integer number of coding passes. Each additional pass contributes some bytes to the block and, in turn, decreases the overall image distortion. Any block may be truncated after any integer number of passes to meet the budget constraint. The task of the rate control is to decide which passes are to be included in the resulting code stream, or, more specifically, to find the set of truncation points $\{n_i\}$ that satisfy the above requirements.

We will proceed with the mathematical formulation of the problem to find the optimal set of truncation points which is also known as the rate-distortion optimization task. Consider $\mathbf{n} = \{n_i\}$ to be *truncation vector*, where n_i signifies that i-th block is truncated in the point number n_i, and $\mathbf{N}$ is the set of truncation vectors. Denote the overall frame distortion after the truncation as $d(\mathbf{n}) = \sum_i d_i^{n_i}$, where quantity $d_i^{n_i}$ corresponds to the distortion value of i-th block truncated in the point number n_i. Similarly, denote resulting frame rate as $r(\mathbf{n}) = \sum_i r_i^{n_i}$. Vectors from the set $\mathbf{N}$ possess the following property: the number of bits for the frame decreases while the distortion value increases. The optimization task is to find $\mathbf{n}_{opt} \in \mathbf{N}$, so that

$$\begin{cases} d(\mathbf{n}_{opt}) = \min\limits_{\mathbf{n} \in \mathbf{N}} d(\mathbf{n}) \\ r(\mathbf{n}_{opt}) \leq r_{max}. \end{cases} \quad (2)$$

In JPEG2000 standard the solution of this optimization task is based on the Lagrange relaxation technique [5] that is used to find the local extreme of a function.

IV MULTI-SOURCE RATE-CONTROL ALGORITHM FOR JPEG2000

The proposed scheme of DVB-H with S video sources on the base of JPEG2000 is shown on the Figure 4. At first, each source is lossless compressed with JPEG2000 Tier-1 codec. During this process for each frame i of the video source s rate-distortion function is calculated. Truncation vector $\mathbf{v}_i^s$, that determine the point $\{r(\mathbf{v}_i^s), d(\mathbf{v}_i^s)\}$ of that curve, as in (2), is found by solving the following optimization task:

$$\begin{cases} r(\mathbf{v}_i^s) = \min\limits_{\{\mathbf{n}_i^s\}} r(\mathbf{n}_i^s) \\ d(\mathbf{v}_i^s) \leq d_{max}, \end{cases} \quad (3)$$

where d_{max} is the maximum value for that point. To get the rate-distortion curve the task (2) is solved for the some quantity of values $\{d_{max}\}$.

The set of these functions is further sent to the input of Quality-Scalable Rate-Control. The task of the Quality-Scalable Rate-Control is the dividing each video stream of each source into L scalable quality levels. To form each level the following Multi-source rate-control algorithm is proposed. For each frame with the index i, from the group of pictures (GOP) with the index k, quality level with the index l is formed so that:

$$\begin{cases} \texttt{minimize} \max\limits_{i,s} d(\mathbf{v}_i^s), i \in [k \cdot N, ..., (k+1) \cdot N - 1], \\ \sum\limits_{s=1}^{S} \sum\limits_{i=k \cdot N}^{(k+1) \cdot N - 1} r(\mathbf{v}_i^s) \leq \frac{l}{L} \cdot B_{max}, \end{cases} \quad (4)$$

where N is the number of frames in GOP, B_{max} is maximum acceptable joint bit-rate.

The stream formed with (4) will have the following properties. At first, video stream for each video source will be divided into L data levels encoded independently. To decode the video stream receiver should receive one or more data levels.

Secondly, bit resources inside the data level will be reallocated so that the visual quality of all video sources in GOP after receiving equal number of data levels will also be equal.

The above-mentioned properties allow the receiver to receive different number of data levels l depending on the energy consumption and reconstructed image quality requirements. In the case of $l = 1$ the highest energy consumption work condition will be chosen. In the case of $l = L$ the work condition with the highest level of reconstructed image quality will be chosen.

V CONSIDERATIONS ON JPEG2000 IMPLEMENTATION FOR DVB-H

The basic scheme for implementing JPEG2000 video transmission over DVB-H can be described as follows. It is designed so that only minor changes for DVB specifications would be required. First, the JPEG2000 codestream is divided into multiple RTP (Real-time Transport Protocol) [13] sessions, each RTP session carrying one quality level of JPEG2000 stream. The RTP payload format for JPEG2000 is discussed in [14] and [15]. The RTP packets are further encapsulated into UDP

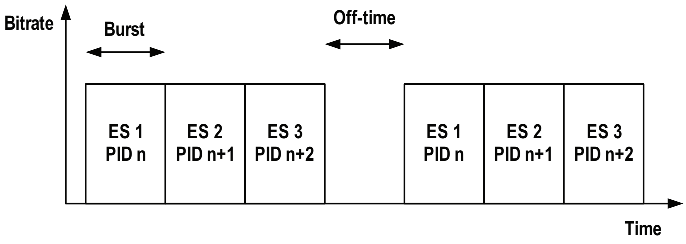

Figure 5: Considerations on JPEG2000 Implementation for DVB-H

(User Datagram Protocol) packets, which are finally encapsulated into IP packets. The RTP sessions must be identified by carrying each session in a distinct multicast IP session, identified by IP address. Identification by port number (and having the same IP address) is not sufficient, since the PSI/SI (Program Specific Information / Service Information) signaling [16] used in DVB-H does not carry port information, and the minimum requirement for IP stream signaling in PSI/SI is only IP destination address. Second, in the IP encapsulation process, the IP encapsulator creates a separate MPEG-2 Elementary Stream of each multicast IP session carrying an RTP session. The Elementary Streams in MPEG-2 Transport Stream are identified by having the same PID (Packet IDentifier) in the MPEG-2 TS packet header. One DVB-H burst carries information belonging to one Elementary Stream. When the JPEG2000 codestream is divided into different Elementary Streams, each quality level is identified by PID, and time-slicing can be applied to receive only bursts containing the desired quality levels. For example, if the video stream contains 3 levels of quality, the base level can be transmitted by ES having PID n, the first additional level can be transmitted by ES having PID n+1, the second having PID n+2. It is the task of the IP encapsulator to ensure that the bursts are created in an appropriate order, i.e. that consecutive bursts contain data related to the same service, and that the data is transmitted in the correct order. The example is illustrated in Figure 5. Third, the receiver must be capable of receiving parallel Elementary Streams (i.e. streams illustrated in Figure 5). Moreover, the receiver must be capable of determining how many Elementary Streams it must receive. Due to the fact that JPEG2000 quality levels are transmitted in consecutive bursts, the receiver must be able to adjust its on and off times accordingly, so that unnecessary power ups and power downs are avoided. For service discovery, the JPEG2000 video services are announced in the ESG (Electronic Service Guide), where each service is assigned at least one SDP (Session Description Protocol) [17] definition. The SDP contains information about the video and audio streams (encoding parameters etc.), and announces the IP addresses and ports where the stream can be accessed. When a desired service is selected, the IP addresses delivering the service are read from the SDP. The receiver must then search from PSI/SI tables the INT (IP/MAC Notification Table) for the IP address and a reference to the PMT (Program Map Table), where the PID for the Elementary Stream carrying the IP stream can be found. Currently SDP definition does not offer semantics for scalable codecs. Hence

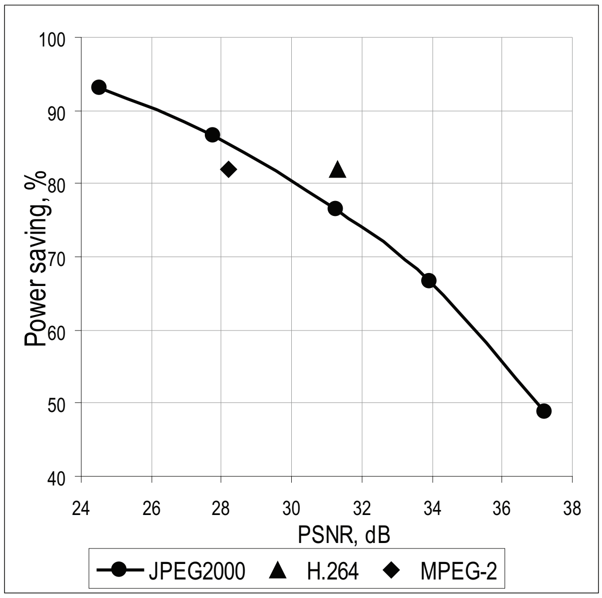

Figure 6: Relation between average PSNR and power saving for test video sequence "matrix", $B_b = 3$Mbps.

the signaling of the use of JPEG2000 codec in the ESG must be supported by the receiver. SDP signaling of JPEG2000 codec may be accomplished by announcing multiple connection addresses in the SDP Media Announcement section (see [17]), one for each IP multicast session. Further, JPEG2000 codec specific attributes can be added as media specific attributes. It is up to the receiver to correctly interpret the attributes.

VI Practical results

Practical results was obtained for DVB-H system in case three video sources ($S = 3$). For it were chosen "matrix", "phenomenon" and "robin" test video sequences with 340x240 resolution and frame rate of 15 Fps. This sequences were compressed with reference JPEG2000 codec [18] with proposed Multi-source rate-control algorithm with quantity of quality levels $L = 5$, number of frame in GOP $N = 15$ and maximum acceptable joint bit-rate $B_{max} = 3000$ kbps.

To compare power saving performance of proposed scheme test video sequences were compressed also with reference MPEG-2 codec with TM5 rate-control algorithm [19] and H.264/AVC codec with JVT rate-control algorithm [20]. Both codecs compressed with bit-rate of 400 kbps and 15 frames in GOP with one I and 14 P-frames.

Figures 6–8 shows for above-mentioned codecs relations between average peak signal-to-noise ratio (PSNR) and power saving in case of burst-rate of 3 Mbps.

VII Conclusions

JPEG2000 is considered to be the best choice for adaptive energy economy as it allows receiver to control power consumption depending on the priorities. This paper illustrates it on the example of JPEG2000 compression algorithm, although

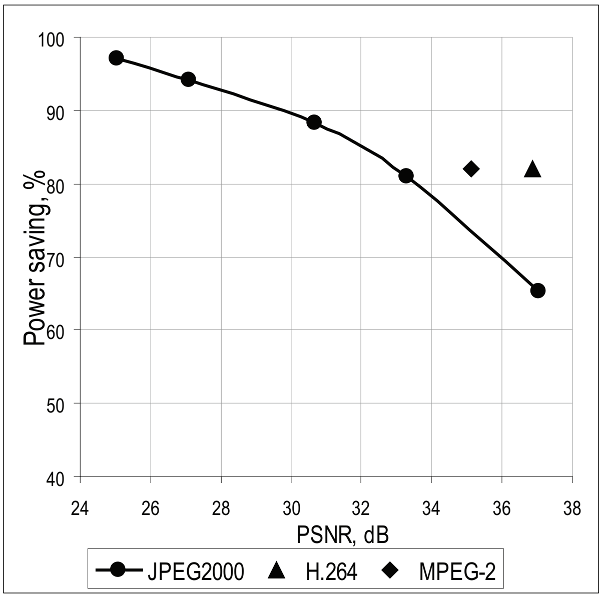

Figure 7: Relation between average PSNR and power saving for test video sequence "phenomenon", $B_b = 3$Mbps.

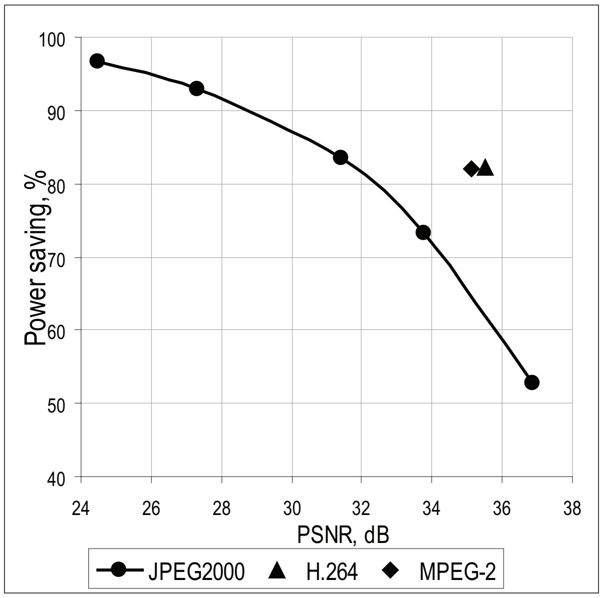

Figure 8: Relation between average PSNR and power saving for test video sequence "robin", $B_b = 3$Mbps.

it can be shown on any other scalable codec. JPEG2000 allow receiver to choose the level of power consumption from five different quality levels. JPEG2000 was also compared with MPEG-2 and H.264/AVC. This comparison shows that despite of good compression rate results, neither MPEG-2 nor H.264/AVC allow receiver to control the level of power consumption. But better result could be achieved if scalable extension of the H.264/AVC [21] is used as compression algorithm. In the future work, authors will investigate scalability issues in detail, and consider also other scalable video codecs.

ACKNOWLEDGMENT

This article is prepared within the scope of Finnish-Russian University Cooperation Program in Telecommunications (FRUCT, http://www.fruct.org) and the authors would like to thank all experts and organizers of the FRUCT program for their help and contribution.

REFERENCES

[1] ETSI EN 302 304 v1.1.1 (2004-11): "Digital Video Broadcasting (DVB): Transmission System for Handheld Terminals (DVB-H)", *European Telecommunication Standard*.

[2] ETSI EN 300 744 (2004-01): "Digital Video Broadcasting (DVB); Framing Structure, channel coding, and modulation for digital terrestrial television", *European Telecommunication Standard*.

[3] ISO/IEC 13818 (MPEG-2): Generic coding of moving pictures and associated audio information, Nov. 1994.

[4] Advanced video coding for generic audiovisual services. *ITU-T Recommendation H.264 and ISO/IEC 14496-10 (AVC)*, 2003.

[5] ITU-T and ISO/IEC JTC 1, "JPEG 2000 Image Coding System: Core Coding System, ITU-T Recommendation T.800 and ISO/IEC 15444-1" *JPEG 2000 Part 1*, 2000.

[6] G. Faria, J. A. Henriksson, E. Stare, and P. Talmola, "DVB-H: digital broadcast services to handheld devices" *Proceedings of IEEE*, vol. 94, no. 1, pp. 194-209, 2006.

[7] H. Himmanen, A. Hazmi and J. Paavola "Comparison of DVB-H Link Layer FEC Decoding Strategies in a Mobile Fading Channel", Proc. *PIMRC '06*, Helsinki, Finland, September 2006.

[8] J. Paavola, H. Himmanen, T. Jokela, J. Poikonen and V. Ipatov, "The performance analysis of MPE-FEC decoding methods at the DVB-H link layer for efficient IP packet retrieval" *IEEE Transactions on Broadcasting*, Vol. 53, No. 1, pp. 263-275, March 2007.

[9] Wing TV project deliverable D15a - Simulation report, 2006. Available on-line: http://projects.celtic-initiative.org/WING-TV/

[10] ETSI TR 102 377 V1.1.1 (2005-02): "Digital Video Broadcasting (DVB); DVB-H Implementation Guidelines", *European Telecommunication Standards Institute*.

[11] S. Loncaric, S. Grgic, B. Zovko-Cihlar, "Minimizing Power Consumption of the DVB-H Receiver" *48th International Symposium ELMAR-2006 focused on Multimedia Signal Processing and Communications*, pp.309-313, 2006.

[12] D. Taubman and M. W. Marcellin, "JPEG2000: Image Compression Fundamentals, Practice and Standards" *Massachusetts: Kluwer Academic Publishers*, 2002.

[13] IETF RFC 3550, RTP: A Transport Protocol for Real-Time. Applications. http://www.ietf.org/rfc/rfc3550.txt

[14] IETF Internet-Draft, RTP Payload Format for JPEG2000 Video Streams. http://www.ietf.org/internet-drafts/draft-ietf-avt-rtp-jpeg2000-19.txt

[15] IETF Internet-Draft, Payload Format for JPEG2000 Video: Extensions for Scalability and Main Header Recovery. http://www.ietf.org/internet-drafts/draft-ietf-avt-rtp-jpeg2000-beam-10.txt

[16] ETSI TS 102 470 V1.1.1 (2006-04): "Digital Video Broadcasting (DVB); IP Datacast over DVB-H: Program Specific Information (PSI)/Service Information (SI)", *European Telecommunication Standards Institute*.

[17] IETF RFC 2327, SDP: Session Description Protocol. http://www.ietf.org/rfc/rfc2327.txt

[18] JPEG-2000 Software, http://www.ece.uvic.ca/ mdadams/jasper/

[19] MPEG-2 Video Codec, http://www.mpeg.org/mpeg/mssg/

[20] H.264/AVC JM Reference Software, http://iphome.hhi.de/suehring/tml/

[21] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the Scalable Video Coding Extension of the H.264 / AVC Standard" *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, No. 9, pp. 1103-1120, 2007.