Controlling Joint Optimization of Wireless Video Transmission: the PHOENIX Basic Demonstration Platform

Maria G. Martini, Matteo Mazzotti, Marco Chiani, Gianmarco Panza,
Catherine Lamy-Bergot, Jyrki Huusko, Gabor Jeney, Gabor Feher, Soon X. Ng

Abstract— **A global approach for realistic network-aware joint source and channel system optimization for wireless video transmission is described in this paper. After a description of the information to be exchanged among the system component blocks, the concept of "JSCC/D controllers" is introduced and the implementation of the first "basic chain" demonstrator realized in the framework of the *PHOENIX* project is described. Simulation results obtained with such first software demonstration platform confirm the validity of the described approach.**

I. Introduction

Relying on Shannon's theory [1], the field of communications has developed under the assumption that the two basic operations of a communications system, source coding and channel coding, can be performed independently with no performance degradations relative to joint design. Regardless of that, many of the modern information technology systems, such as those involving transmission of video sources over rate constrained channels, actually violate the conditions upon which the optimality of that principle relies. For such systems, performance improvements may be achieved by moving from separate design and operation of source and channel codes to joint source-channel code (JSCC) design and operation. Unlike separation-based techniques, joint source-channel coding techniques rely on the joint or cooperative optimization of communication system components. The joint approach allows strategies where the choice of source coding parameters varies over time or across users in a manner that in some way depends on the channel or network characteristics. Likewise, joint source-channel coding techniques allow for systems in which the choice of channel code, modulation, or network parameters varies with the source characteristics. For a tutorial description of the JSCC approach see e.g. [2]. One of the main drawbacks of the joint approach is that it requires the exchange of a variety of information between the systems blocks and that these information need to be managed jointly in order to perform the system optimization.

In this paper, a quality driven approach for wireless video transmission relying on the joint source and channel coding paradigm is proposed. In particular, the management of the information to be exchanged is addressed and the logical units responsible for the system optimization, in the following referred to as JSCC/D controllers, having a key role in the system, are analyzed. The first, basic, demonstrator of the described system realized in the framework of the IST PHOENIX project is also described and simulation results are shown.

Maria G. Martini, Matteo Mazzotti, Marco Chiani are with CNIT/University of Bologna, Italy. Gianmarco Panza is with CEFRIEL/Politecnico Milano, Milan, Italy. Catherine Lamy–Bergot is with Thales Communications, Colombes, France. Jyrki Huusko is with VTT Electronics, Oulu, Finland. Gábor Jeney and Gábor Feher are with Budapest University of Technology and Economics, Hungary. Soon X. Ng is with University of Southampton, U.K.

II. Overall System Architecture

Figure 1 represents our overall system architecture developed in the framework of the PHOENIX project, including the signals for transmitting the JSCC/D control information through the system. The transmitter unit is illustrated in the upper part of the figure and the receiver unit in the lower side of the figure.

The figure shows that in addition to classical encoding and decoding blocks, the architecture includes both physical and application layer controller units. Controllers are used for supervising the (de)coders and channel modulation units and for triggering the (de)coders to adapt to changing conditions, through the management of information about the source, network and channel conditions and user requirements. For the controlling purpose, a signaling mechanism may be defined.

A. Side information exchanged in the system.

In particular, the information we take into account for the system optimization are represented by source significance information (SSI), i.e. the information on the sensitivity of the source bitstream to channel errors; channel state information (CSI); decision reliability information (DRI), i.e. soft values output by the channel decoder; source a-priori information (SRI), e.g. statistical information on the source; source a-posteriori information; (SAI), i.e. information only available after source decoding; Network state information (NSI), represented e.g. by packet loss rate and delay and finally the video quality measure, output from the source decoder and used as feedback information for system optimization.

SSI is strictly related to the data stream and needs to be synchronized with the stream. Its knowledge allows e.g. performing unequal error protection of the source bitstream. Due to the strict relation to the media stream, the transmission of the SSI signals provides a significant

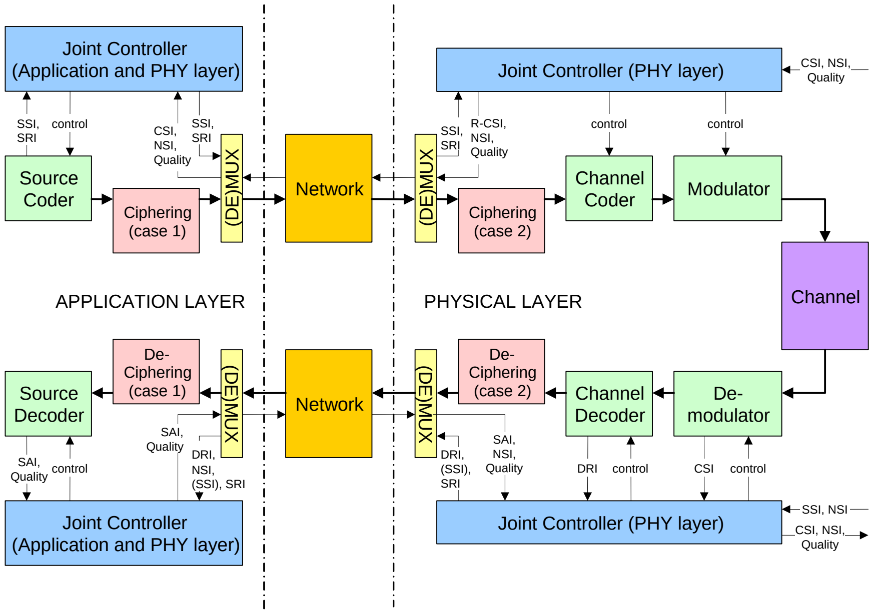

Fig. 1. System Architecture

challenge to interlaying network and protocol communications. Example techniques to transmit SSI and to reduce SSI overhead have been proposed by the authors in [3], [4].

CSI represents the actual conditions of each wireless channel through which the media stream is directed. CSI is generated by the radio receiver node and effectively exploited by all the JSCC/D protocol levels on the transmitting side, at the radio interface (for the channel coders and modulators) and at the source coder. A different level of detail can be required to CSI at the different levels of the transmission chain. In particular the so-called "physical layer controller" normally needs full CSI, in order e.g. to optimize multicarrier modulation bit loading parameters, whereas the optimization at application layer may be performed according to a reduced version of CSI (e.g information averaged over longer time interval).

DRI provides further elements related to the channel decoding process. It is based on the concept of soft decision, where the final result about the value of a bit is not simply either 0 or 1, but also the likelihood value of it.

The SRI is further information produced by the source coder that is exploited at the destination side and possibly also by the other entities concerned in the JSCC/D chain along the data path at the radio transmitter nodes, in order to optimize the QoS resulting from the decoding process of the video stream.

The SAI results from the analysis of the decoding process of the video stream. It is generated by the destination terminal and exploited at the radio receiver to set the working parameters of the channel decoder and demodulator module. In a further design step, it can be exploited even at the transmitting terminal in order to improve the performance of the channel coding and modulation and the resulting QoS.

NSI reports about the availability of network resources across the data path. Such information can be represented by the QoS performance parameters. For example delay, delay-variation (jitter) and packet loss gives an idea of the network load conditions. Although the bottleneck is on the wireless channel, some impairments can be introduced in the other part of the network too. Therefore, NSI can be effectively exploited at the source coder to better tune the amount of the generated rate and coding parameters in general, as well as at each radio transmitter node.

The maximization of the received video quality is the target of the system optimization process. Information on quality evaluation of preceding frames is thus of critical importance in order to tune the system parameters. The quality evaluation should be performed "on-the-fly", without reference to the transmitted frame. For simplicity and in order to refer to a widely known metric, we will refer anyway in the following to PSNR, assuming the knowledge of the correct PSNR value for system optimization.

In addition, the JSCC/D system requires information exchange in set-up phase, for initial capacity negotiations and authentication process. In particular, in the set-up phase the joint controllers collect the information on the available system options (e.g. available channel encoders and available channel coding rates, available modulators,...) and constraints.

Figure 2 depicts in more detail the transmitter side of the system under consideration, implemented with some restrictions in terms of options available in the PHOENIX basic transmission chain demonstrator.

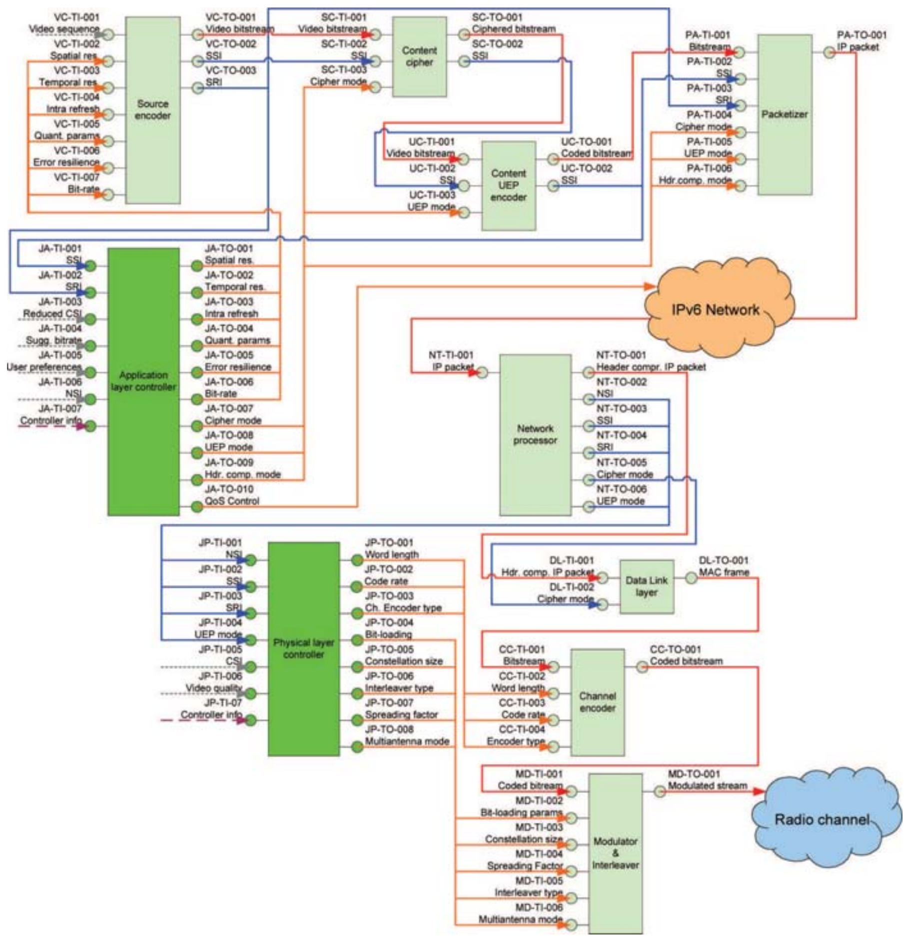

Fig. 2. The PHOENIX signalling scheme - Transmitter side.

B. Network Transparency.

In order to allow the exchange of control and signaling information between concerned JSCC/D entities, it is necessary to tackle the issue of the Network Transparency. This is somehow an abstract idea of making the underlying network infrastructure almost invisible, from which the term transparency, to all the entities involved in the jointly optimization of the source and channel (de)coder, as well as of the (de)modulator. Almost transparent is related to the fact that the telecommunication infrastructure by its own inevitably affects in some extent the overall system, e.g. by introducing delay, loss and various types of errors, but without actually interacting with the control-plane of the concerned deployed devices and providing sufficient delivery guarantees to the video streams in order to accomplish a defined quality of service (QoS) for the end-user. The goal is twofold: to realize communication exchanges between differently located entities into the network (including the end-terminals) not to interact anyhow with non-JSCC/D aware devices. The primary goal is referred to the capability of transferring signalling and control information between both different network nodes and link layers as needed, in a transparent manner, in spite of the strict rules of the ISO OSI model, which impose a modular and independent design of each link layer of a network node with well defined interfaces and the delivering through a telecommunication infrastructure that carries data only of a specific format (IP datagram, in the case). The second objective aims to ensure as much as possible backward compatibility, not only with the existing standards as also addressed by the first goal, but even with the nowadays telecommunication infrastructure that constitutes the basis for the next generation networks, allowing for a smooth migration to IPv6-enabled devices eventually supporting JSCC/D functionalities. Some mechanisms that could implement the concept of the Network Transparency are: IPv6 data packets and extension headers, ICMPv6 messages, direct socket-to-socket communication, external databases and service profiles stored in shared memory spaces. Other possible methods relies either on the introduction of adaptation layers at the transmitter and receiver side to allow for the exchanges implicated by the joint source and channel (de)coding system [7], or the exploitation of already existing and deployed ad-hoc signalling protocols.

C. JSCC/D Controllers.

The system controllers are represented by two distinct units, namely the “physical layer (PHY) controller” and the “application layer (APP) controller”. The latter collects information from the network (NSI: packet loss rate, delay and delay jitter) and from the source (e.g. SSI) and has availability of reduced channel state information and of the quality metric of the previously decoded frame. According to these information, it produces controls for the source encoder block (e.g. quantization parameters, frame rate, error resilience tools to activate) and for the network. In our first model, we considered the controller as a finite state machine characterized by a number of states, each corresponding to a combination of the said parameters. The transitions among the states are determined by the input parameters. In particular, a low quality value associated to a negative trend will cause a transition to a state characterized by a higher robustness. Given the source bit-rate associated to the chosen state, the code rate available for signal protection is evaluated considering the total $R_t = R_S/R_C$ constraint. That information is provided to the physical layer controller.

The PHY controller's task is to provide controls to the physical layer blocks, i.e. the channel encoder, modulator and interleaver. In the simplified case considered in section III, where modulation is fixed, the controller decides, similarly as in [6], the channel coding rate for each different sensitivity source layer, with the goal of minimizing the total distortion D_{S+C} with the constraint of the average channel coding rate R_C. In a more complex scenario, the controller also sets the parameters for bit-loading in multicarrier modulation, interleaver characteristics and performs a trade off with receiver complexity.

III. Example application

We describe here an example application of the proposed approach, in order to show the achievable gain of the controllers-managed system described. The first PHOENIX demonstration platform allows to evaluate the performance of H.264 and MPEG-4 video transmission over the system described. We consider in the following MPEG-4 video transmission over a wireless channel affected by time-correlated Rayleigh fading and Log-normal shadowing, with unequal error protection through rate compatible convolutional codes [2] (performed by the PHY controller described above) and joint source and channel rate allocation with a total bit-rate constraint (performed by the APP controller described above). Ciphering is performed at content level; UDP/IP packetization, UDP-lite, IPv6 network and MAC header addition are taken into account. Simple BPSK modulation is assumed. Two receiving antennas with ideal selection combining are considered. The error resilience technique in [5] is also applied. The main parameters considered are reported in table I. 64 s of total simulation have been considered, i.e. 64 cycles of one second each. The controllers perform thus system adaptation every 1s time-step.

Even with the simple and sub-optimal adaptation algorithms implemented, the gain achievable in terms of perceived video quality when JSCC-D Controllers are included in the transmission chain is absolutely remarkable. In the following, two different simulations are compared:

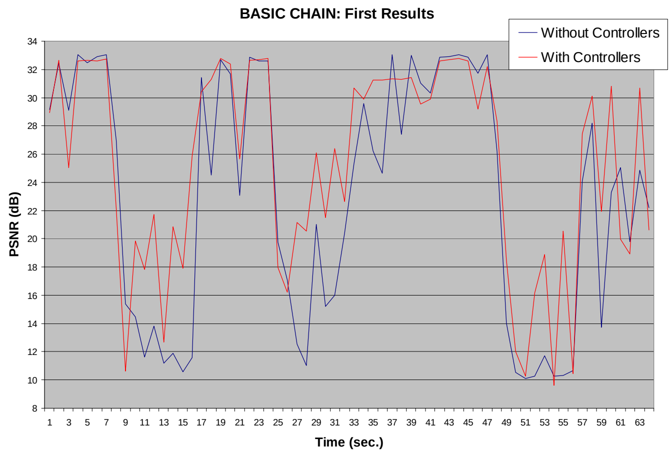

Fig. 3. PSNR vs. time for simulation 1 (no adaptation) and simulation 2 (adaptation)

Fig. 4. Frame 996. Adapted (left); non-adapted (right).

in simulation 1 the adaptation algorithms are not applied and we considered fixed parameters: a video frame-rate $F_r = 30fps$, Intra-refresh every 15 frames and quantization parameters $q_I = 8$, $q_P = 12$, giving a source bit rate of about 400 kbps. Equal error protection is performed on the bitstream output by the network. Considering the redundancy provided by network headers, channel coding rates for the two layers are $R_{C1} \simeq 0.66, R_{C2} \simeq 0.66$.

On the contrary, in simulation 2 the described algorithms for the APP and PHY JSCC-D controllers are activated. Figure 3 reports the PSNR vs. time curve.

In good channel conditions the two solutions are in general equivalent whereas when the channel meets fading, simulation 2 shows an improvement with respect to simulation 1. On the whole, Simulation 2 reaches a gain of 2.31 dB in terms of mean PSNR with respect to Simulation 1. Even more sensible improvements are achievable in terms of perceived visual quality, as it is clear from the example frames shown in Fig. 4, representing the case of good channel conditions, where the two schemes provide similar results, and in Fig. 5, representing the case of bad channel conditions, where the improvement achievable with the adapted scheme is more evident.

IV. Conclusions

A global approach for realistic network-aware joint source and channel system optimization has been described in the paper. The information to be exchanged among the

TABLE I

Simulation parameters considered.

Test Video Sequence	
Video Sequence:	foreman
Video Format:	CIF (352x288)
Frame Rate:	30 fps
Duration:	64 seconds
Source Coding	
Encoded sequence frame rate:	30, 15, 7.5 fps
Group of Video Objects dimension:	30, 15, 8 frames
Intra refresh period:	30, 15, 8 frames
Initial value for I-quantizer:	8
Initial value for P-quantizer:	12
Approximate size of MPEG-4 packets:	1000 bits
Packetisation	
Number of priority layers:	2
Network	
Number of nodes in the IPv6 network:	10
Mean node delay :	3 ms
Mean node packet loss:	100 ppm
Bottleneck rate:	10000 kbps
Buffer size at the bottleneck :	100000 bytes
Radio Link	
Number of CRC bits:	6
Channel encoder:	convolutional
Mother code rate:	1/5
Constraint length:	6
Code generators (in octal):	(75,71,73,65,57)
Puncturing period:	8
Code rates considered:	1/5, 2/9, 1/4, 2/7, 1/3, 2/5, 1/2, ...
Interleaving (on the single packet):	random
Modulation:	BPSK
Number of RX antennas:	2 (ideal Selection Combining)
Maximum coded bitrate:	1 Mbps
Radio channel:	non-selective block fading channel
Median Es/No:	4 dB
Slow fading:	uncorrelated Log-Normal distributed $\sigma = 4$ dB coherence time (=slow-fading block duration): 8000 ms
Fast fading:	time-correlated Rayleigh distributed doppler frequency: 2 Hz channel gain sample time (=fast-fading block duration): 0.1 ms

Fig. 5. Frame 1818. Adapted (left); non-adapted (right).

system block are described and the concept of “controllers” introduced. A basic demonstrator is then described. Simulation results confirm the validity of the described approach. In particular the basic example of quality controlled scheme proves the robustness of the scheme when channel conditions become particularly harsh.

V. Acknowledgement

This work has been partially supported by the European Commission in the framework of the “PHOENIX” IST Project. The whole PHOENIX IST project consortium is also gratefully acknowledged.

References

[1] C. E. Shannon, “A mathematical theory of communication,” *Bell System Technical Journal*, vol. 27, pp. 379-423, 623-656, July-Oct. 1948.

[2] J. Hagenauer and T. Stockhammer, “Channel coding and transmission aspects for wireless multimedia,” *Proceedings of the IEEE*, vol. 87, no. 10, Oct. 1999.

[3] M. G. Martini, M. Chiani, “Proportional Unequal Error Protection for MPEG–4 video transmission”, *Proc. IEEE ICC 2001*, Helsinki, June 2001.

[4] M. G. Martini, M. Chiani, “Robust Transmission of MPEG–4 Video: Start Codes Substitution and Length Field Insertion Assisted Unequal Error Protection”, *Picture Coding Symposium - PCS 2001*, Seoul, April 2001.

[5] M.G.Martini, M.Chiani, “Joint Source-Channel Error Detection with Standard Compatibility for Wireless Video Transmission”, *Proc. IEEE WCNC 2002*, Orlando, Florida, March 2002.

[6] M.G. Martini, M. Chiani, “Rate-Distortion models for Unequal Error Protection for Wireless Video Transmission,” *IEEE VTC 2004 Conference*, Milan, Italy, May 2004.

[7] C.Lamy-Bergot and P. Vila, “Multiplex header compression for transparent cross-layer design,” *IEEE Int. Conf. on Networking (ICN’04)*, Pointe-á-Pitre, French Polynesia, March 2004.