A Cross-Layer Algorithm for Video Transmission over Wireless Systems with Hybrid MIMO Structure

Martin B. Obando, Walter C. Freitas Jr, and Francisco R. P. Cavalcanti,

Abstract—In this paper, we propose a feasible cross-layer algorithm to improve video quality in wireless communications systems. This algorithm encompasses both the application and the physical layers. In the application layer, the classification of bits is accomplished by the video encoder. In the physical layer, the spatial multiplexing and the diversity gains provided by multiple input multiple output antenna systems with hybrid structure are exploited. Besides that, channel knowledge at the transmitter is used for antenna selection to enhance the performance, which is shown to improve the objective and subjective video quality at the end user.

Index Terms—cross-layer, hybrid MIMO structure, antenna selection, unequal error protection, video transmission.

I. INTRODUCTION

THE transmission of video over wireless communications systems presents a great challenge concerning the video quality at the end user. In this medium, the transmitted signals are subject to several propagation mechanisms, such as multi-path fading and shadowing, responsible for degrading the quality of the received signals. At the same time, the rapid development of multimedia packet based services over wireless systems has encouraged the search for new transmission schemes aiming at a higher spectral efficiency [1], [2]. This can be achieved, on one hand, by developing low bit rate encoding algorithms. On the other hand, one can focus on increasing the radio link reliability by the use of antenna diversity.

Multiple-Input Multiple-Output (MIMO) transceiver structures provide gains in spectral efficiency and robustness by exploiting spatial multiplexing and spatial diversity from rich scattering in the radio link, respectively [3]. The use of MIMO for wireless video transmission has been previously considered in several works. In [4], Diversity Embedded Space-Time Codes (DESTC) are used to send multiple layers of video simultaneously. The DESTC provide Unequal Error Protection (UEP) for different video layers, thereby delivering high video quality to users with good channel conditions, and providing an acceptable quality for users with poor channel conditions.

In [5], the authors proposed a hybrid MIMO transmit structure to implement UEP for video delivery in a MIMO system. The goal of that work was to exploit the diversity gain to provide a better protection for the high priority data, while transmitting the low priority data with spatial multiplexing to achieve high data rates.

In this paper, we propose a Cross-Layer (CL) algorithm, which is an optimization approach that encompasses layers of

the Open Systems Interconnection (OSI) model. The proposed CL algorithm, encompasses the application and the physical layers. In this algorithm, the application layer takes advantage of the video coding classifying the bit-streams by priority depending on the frame type, I frame or P frame [6]. In the physical layer, the classified bit-streams are arranged over a Hybrid MIMO Structure (HMS) to achieve the UEP. A HMS is a MIMO antenna structure that arrange in parallel schemes of spatial diversity and multiplexing in one single structure. Furthermore, an Antenna Selection (AS) mechanism improves the performance selecting an appropriate subset of the available antennas in MIMO systems. In this work, the idea to use the AS mechanism on the transmitter side is that streams with smaller diversity gains will be transmitted over the antennas with best instantaneous link quality.

The contribution of this work, in comparison with [4] and [5], consists in the concept of cooperation among layers (i.e. Cross-Layer). According to this concept, the classified bit-streams are transmitted using the UEP scheme over a HMS. In addition, the AS mechanism is used to obtain robustness over the uncoded layers of the HMS. In [4] and [5] layered video coding (i.e., scalable coding) is considered, while in the proposed algorithm a non-scalable syntax is used. Another contribution of the proposed CL algorithm is related to the transmission efficiency of each Group of Pictures (GOP) (i.e., a set of frames in the form IPPP...IPPP...) in the video sequence. When the priority bit-streams end in a GOP the remaining bit-streams of less priority are transmitted over all layers of the HMS. Furthermore, in the proposed CL algorithm, we consider a more realistic physical layer design which includes Cyclic Redundancy Check (CRC), multilevel digital modulation and turbo coding, in comparison with [5].

II. CHANNEL MODEL

We consider wireless transceivers equipped with M_{tx} transmit antennas and M_{rx} receive antennas. The wireless channel is assumed to have rich scattering and flat fading. The channel is quasi-static, i.e., the channel does not change significantly in a single block, but can vary from block to block. In each block, we can represent the sampled received signal as in [7]:

$$\mathbf{Y} = \sqrt{\frac{\rho}{M_{\text{tx}}}} \cdot \mathbf{H}\mathbf{S} + \mathbf{N}, \qquad (1)$$

in which $\mathbf{Y} \in \mathbb{C}^{M_{\text{rx}} \times T}$ is the received signal matrix, T is the number of signaling intervals, $\mathbf{S} \in \mathbb{C}^{M_{\text{tx}} \times T}$ is the transmitted signal matrix, ρ is the average Signal-to-Noise Ratio (SNR) at each receive antenna, $\mathbf{H} \in \mathbb{C}^{M_{\text{rx}} \times M_{\text{tx}}}$ is the random channel

matrix, and $\mathbf{N} \in \mathbb{C}^{M_{\text{rx}} \times T}$ is the additive noise matrix. The entries of $\mathbf{H}$ and $\mathbf{N}$ are i.i.d. Zero Mean Circularly Symmetric Complex Gaussian (ZMCSCG) considered.

III. HYBRID MIMO STRUCTURE (HMS)

The use of MIMO systems is exploited mainly in two ways. The former is related to the improved link reliability provided by the diversity gain from, for example, the Orthogonal Space-Time Block Code (OSTBC) structures. In such structures, the same information symbol stream is transmitted from different transmit antennas in an appropriate manner creating redundancy. The second way refers to increasing the link spectrum efficiency by means of spatial multiplexing gain. Such gain can be obtained from, for example, Vertical Bell Laboratories Layered Space-Time Architecture (VBLAST) structures, in which different symbol streams are simultaneously transmitted from all the transmit antennas in the same time and frequency [8]. Another way to exploit the MIMO systems is to use the HMS that arises as a solution to the problem of achieve spatial multiplexing and diversity gains simultaneously.

In general, the transmission process of a HMS can be divided into layers, somewhat like in VBLAST. However, in contrast to VBLAST, in the HMS, a layer may consist of a stream of symbols at the output of an OSTBC, transmitted by a group of antennas; or an uncoded stream of symbols multiplexed in time, which is transmitted from a single antenna. Based on this concept of layers, the HMS combines pure diversity schemes (e.g., OSTBC) with pure spatial multiplexing schemes (e.g., VBLAST) [9]. The general transmission matrix of the HMS can be given by

$$
\mathbf{S_{HMS}} = \begin{bmatrix} \mathbf{S_{OSTBC_1}} \\ \vdots \\ \mathbf{S_{OSTBC_i}} \\ \mathbf{S_{VBLAST_{i+1}}} \\ \vdots \\ \mathbf{S_{VBLAST_j}} \end{bmatrix}, \tag{2}
$$

in which $\mathbf{S_{OSTBC_i}}$ represents a sub-matrix for layer i transmitting its symbols using an OSTBC structure, and $\mathbf{S_{VBLAST_j}}$ represents a sub-matrix for layer j transmitting its symbols using a VBLAST structure. It follows that $i + j \leq M_{\text{tx}}$.

With this idea, the HMS achieves a compromise between spatial multiplexing and transmit diversity gains through the combination of OSTBC and VBLAST layers in a parallel fashion.

IV. HYBRID MIMO RECEIVER

Since HMSs combines OSTBC and VBLAST layers in parallel, the spatially-multiplexed MIMO layers see each other as interference. Thus, Interference Cancellation (IC) algorithms, similar to those employed in VBLAST, are mandatory in the receiver [9].

For the OSTBC, the maximum-likelihood detection involves just simple linear operations in the receiver. Once all HMSs

employ at least one OSTBC layer, we consider a receiver that takes into account this simplicity. In fact, we adapt the IC algorithm in such a way that the orthogonal structure of the space-time code is preserved as much as possible in its output signal [9].

One particularly successful IC algorithm is called Succesive Interference Cancellation (SIC) [7]. In the SIC, the MIMO layers are detected sequentially. Initially, the received signal $\mathbf{Y}$ goes through a linear detector for the layer 1, whose output is used to produce a hard estimate of the symbols at this layer, $\hat{\mathbf{S}}_1$. Then, the contribution of the layer 1 to the received signal is estimated and cancelled, generating the signal $\mathbf{Y}_2$. The process is then repeated. In general, at the i-th layer, the signal $\mathbf{Y}_i$, hopefully free from the interference of the layers, $1, \cdots, i-1$, goes through a linear detector that tries to mitigate the interference from the layers $j > i$. A hard estimate of the symbol at this layer, $\hat{\mathbf{S}}_i$, is then produced, based on the output of this linear detector. Then, the contribution of this layer to the received signal $\mathbf{Y}_i$ is estimated and cancelled. This procedure yields for a given time instant t a modified received signal given by

$$
\mathbf{Y}_{i+1}[t] = \mathbf{Y}_i - \hat{\mathbf{S}}_i[t]\mathbf{H}_i[t], \tag{3}
$$

in which $\mathbf{H}_i$ corresponds to the first i-th columns of the channel matrix $\mathbf{H}$, corresponding to the channel gains associated to layer i. $\hat{\mathbf{S}}_i[t]\mathbf{H}_i[t]$ represents the estimated interference from the i-th layer. The result is that $\mathbf{Y}_{i+1}[t]$ is free from the interference that comes from the layers $1, \ldots, i$. This signal is then fed into the linear detector for the $(i + 1)$-th layer. This technique is also known as nulling and canceling algorithm [10]. This procedure eliminates the interference among different layers. For a given layer that transmits its symbols following an OSTBC, followed by the nulling procedure, it is necessary to employ the classical OSTBC detector using a spatial matched filter [10].

The performance of the SIC can be improved if the layers are detected in an appropriate order, resulting in Ordered Successive Interference Cancellation (OSIC) [7].

V. PROPOSED CROSS-LAYER ALGORITHM FOR REAL TIME VIDEO TRANSMISSION

In the proposed CL algorithm, we consider the application layer, in which a video encoder is used, and the physical layer, in which the HMS is exploited as a mechanism of bit-stream transmission.

The hybrid video encoder operates sequentially, frame by frame, generating I and P frames, in this case. The P frames achieve compression, by reducing temporal redundancy, and the I frames achieve compression, by reducing the spatial correlation among neighboring pixels in the same frame. The insertion of I frames is considered as a technique to avoid temporal error propagation. The I frame introduces a trade-off between latency and quality degradation because of its relatively large size, in comparison to P frames [11].

Without loss of generality, a particular HMS structure that consists of $M_{tx}=3$ organized into two layers is considered. The
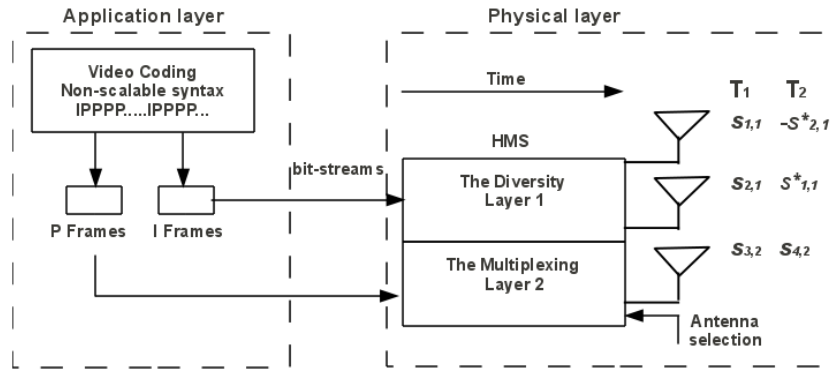
Fig. 1.   Architecture of the Cross-Layer algorithm.

transmission matrix of HMS $\mathbf{S}$ is given by

$$\mathbf{S} = \left[\begin{array}{c} \mathbf{S_{OSTBC_1}} \\ \mathbf{S_{VBLAST_2}} \end{array} \right] = \left[\begin{array}{cc} s_{1,1} & -s_{2,1}^* \\ s_{2,1} & s_{1,1}^* \\ s_{3,2} & s_{4,2} \end{array} \right], \qquad (4)$$

in which $s_{k,j}$ represents a given symbol k transmitted by layer j, and $*$ denotes the complex conjugate. The first layer (i.e., MIMO Layer 1) is composed of an Alamouti OSTBC [12], while the second layer (i.e., MIMO Layer 2) is composed of one antenna. Consequently, the considered HMS transmits $K = 4$ information symbols (i.e., $s_{1,1}; s_{2,1}; s_{3,2}$ and $s_{4,2}$) during $T = 2$ signaling intervals, so that the effective spectral efficiency of this structure is given by $\eta = (K/T) \cdot \log_2(\mu)$, in which μ is considered the cardinality of the modulation scheme. From Equation (2), the application of the HMS for a higher number of antennas is straightforward.

The AS is performed selecting the sub-channel with highest SNR to the MIMO Layer 2. This is due to the lower protection against the channel fading at this layer, as compared to the MIMO Layer 1. This can be performed using some form of limited feedback from the receiver, as pointed out in [9]. An interesting point is that it is not necessary for the transmitter to know all the Channel State Information (CSI), but just the order of the more powerful links. We denote this approach as partial CSI.

The proposed CL algorithm can be summarized in the following steps:

1) Encode video frames using I and P frames, considering the insertion of an I frame during a certain period (i.e., intra-period);
2) Classify the video encoded bit-streams as either I or P frame bit-stream type;
3) Obtain partial CSI of the channel matrix $\mathbf{H}$, so as to perform AS;
4) Allocate the I frame bit-stream into the MIMO Layer 1 and the P frame bit-streams into the MIMO Layer 2;
5) If an I frame bit-stream is transmitted from a GOP, then allocate the remaining P frame bit-streams of the GOP into both MIMO layers;
6) Do the same steps for the remaining GOPs.

Figure 1 shows the architecture of the proposed CL algorithm considering the HMS described in Equation (4).

In the application layer block, the video coding process in non-scalable syntax mode is used, where GOPs are composed of I and P frames bit-streams. At this layer, the frames of video are classified and identified during the packetization process for their transmission in the physical layer.

In the physical layer block, the bit-streams of video are allocated in the HMS layers according to their priority. In this way, the diversity layer (i.e., MIMO Layer 1) is used for I frames, which have high priority, while the P frames have less priority and are transmitted in the multiplexing layer (i.e., MIMO Layer 2). Considering the size in bits of packets of I and P frames per GOP, both layers of the HMS are used for transmit the remaining bits of the P frames in each GOP. The AS is used to obtain link reliability over the uncoded layers (i.e., MIMO Layer 2, in this case).

VI. RESULTS

In order to illustrate the performance of the proposed algorithm, the video coder H.264/AVC is employed in the application layer. The Quarter CIF (QCIF) video sequences "Akiyo" and "Grandma", with slow motion, and "Carphone" and "Foreman", with median motion, are used and classified as in [13]. The video sequences are coded at 15 fps with intra-period 19, in which each video sequences consists of 240 frames. The bit rate of video encoding is 64 kbps. At the video decoder, the error concealment technique based on a frame copy is configured. The packet size is assumed to be of 250 bytes.

In the physical layer, a CRC-24 and a 320 bit-interleaved turbo code at a coding rate of 1/3 are considered. Different matches of the bit-streams are implemented keeping a fixed spectral efficiency of 6 bps/Hz over each MIMO structure. We consider $M_{\text{tx}} = 3$ and $M_{\text{rx}} = 3$ antennas in the HMS described in Equation (4) with 8-Phase Shift Keying (8PSK) modulation. For performance comparison we also consider, the pure OSTBC H3 with 256 Quadrature Amplitude Modulation (256QAM) as proposed in [10]. On the receiver side, for detection of the signals in the HMS, an OSIC algorithm is used as in [9].

If we analyze the video encoding process per GOP, it is composed of one I frame, and the remaining ones are P frames. Then, the data rate of all P frames bit-streams in each GOP

is more than two times higher than the I frames bit-stream, on average for encoding in baseline profile to data rate of 64 kbps. For this reason, if the I frame bit-stream of the GOP has been transmitted, then the P frame bit-streams are assigned to both layers, in order to speed up the transmission (i.e., the fifth step of the proposed CL algorithm).

In the evaluation of the video transmission, the bit-streams are grouped into packets, which can be lost during the wireless transmission depending on the Packet Error Rate (PER), according to the Bit Error Rate (BER) performance of the MIMO structures, assuming independent bit errors inside each packet.

Figure 2 shows the BER performance of the HMS with and without AS. We can observe that the HMS with AS outperforms the BER of the HMS without AS in almost all SNRs. This is due to the allocation of the antennas with best instantaneous link quality to the uncode layer (i.e., MIMO Layer 2, in this case). The HMS with AS shows better BER performance in comparison with the H3 structure from a SNR of 0 to 11 dB, due to the high modulation used to reach equal spectral efficiency. The improvement of the AS in the BER performance is more evident for low SNR regime, since in high SNR the uncoded layer is capable to achieve a satisfatory performance. For SNRs from 12 to 15 dB, the H3 structure is superior, because of the higher diversity gain offered by H3 [3].
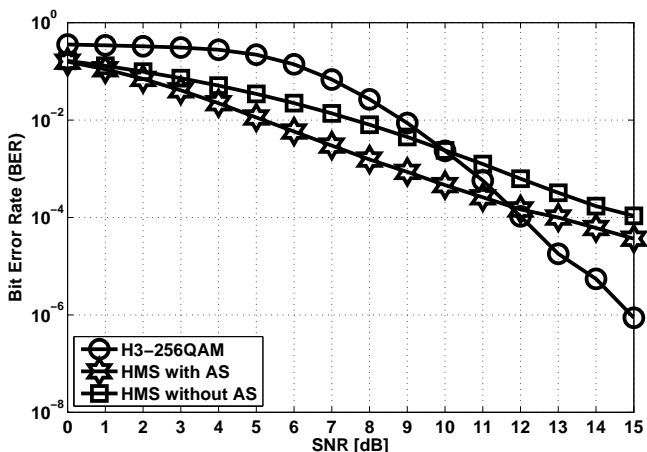

Fig. 2.   BER Performance of the MIMO structures.

In Figure 3, it is shown that the proposed CL algorithm, in comparison with the OSTBC H3, achieves a better Peak Signal-to-Noise Ratio (PSNR) performance for low SNRs (i.e., below 12 dB). For higher SNRs, the OSTBC H3 structure presents better PSNR performance because of its diversity property, but with a more complex constellation to maintain an equal efficiency spectral of 6 bps/Hz. We can also observe a superior performance of the CL algorithm, in comparison with the HMS non-CL and UEP over MIMO of [5]. This is due to the loss of bit-streams that, depending on their type, can result in a high PSNR degradation, because of the high interdependency between themselves in their video coding process (i.e., P frames) or the loss of the refresh property

TABLE I
SUBJECTIVE PERFORMANCE (SNR = 12dB)

Video Sequence	Proposed CL algorithm	H3-256QAM	HMS non-CL
Akiyo	2.79	3.58	2.37
Carphone	2.53	2.33	2.13
Foreman	2.26	2.0	1.60
Grandma	3.0	3.16	2.38

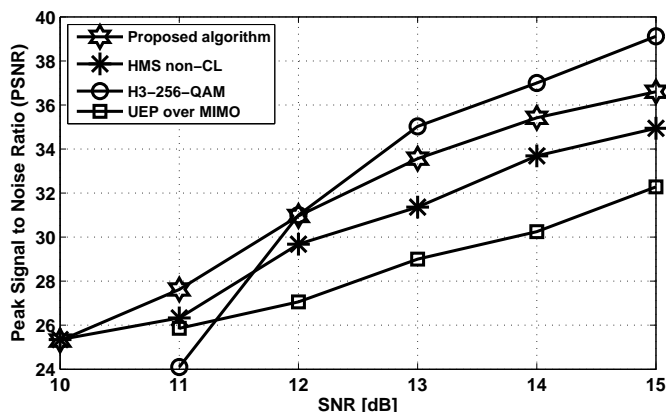reached with the I frame bit-streams.

Fig. 3.   PSNR Performance of the proposed CL algorithm.

The comparison of the proposed CL algorithm with the OSTBC H3 is interesting, since the H3 structure uses 4 signaling intervals, while the proposed algorithm uses only 2 signaling intervals, in which a constant channel is considered. This, together with the better performance of the CL algorithm in low SNRs, suggest a superior performance of the proposed method in more harsh wireless environments.

Table I presents the results of a subjective video quality evaluation that involved 20 observers. The presented videos for assessment consisted in the original video sequence, the transmitted video sequences using the proposed CL algorithm, the OSTBC H3 and the HMS non-CL. It follows an assessment method from [14]. The web page used for this subjective assessment can be found in [15].

In Table I, we can observe that the proposed CL algorithm obtains better subjective video quality when video sequences of median motion, as "Carphone" and "Foreman", are transmitted. It is due, to the lower interdependency created during the video encoding process in comparison with the slow motion video sequences as "Akiyo" and "Grandma".

VII. CONCLUSION

The proposed Cross-Layer (CL) algorithm exploits Unequal Error Protection (UEP), while the spectral efficiency of the radio link and its reliability are increased by employing the Hybrid MIMO Structure (HMS) and Antenna Selection (AS). The method results in a superior objective and subjective video quality to the end user in scenarios where the wireless environment is harsh or the video sequences are challenging for wireless transmission.

REFERENCES

[1] P. V. Pahalawatta, et al., "Review of Content-Aware Resource Allocation Schemes for Video Streaming over Wireless Network," *Wireless Communications and Mobile Computing*, vol. 7, pp. 131–142, February 2007.

[2] Z. Naor, "Multicast Video Streaming for 4G Wireless Networks," *Wireless Communications and Mobile Computing*, vol. 7, pp. 171–185, February 2007.

[3] L. Zheng, et al., "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple Antenna Channels," *IEEE Transactions on Information Theory*, vol. 49, pp. 1073–1096, May 2003.

[4] C. Bilen, et al., "Layered Video Multicast Using Diversity Embedded Space Time Codes," *IEEE Sarnoff Symposium, 2009*, pp.1–5, 2009.

[5] G. Yang, et al., "Unequal Error Protection for MIMO Systems With Hybrid Structure," *IEEE Circuits and Systems*, pp. 682–685, September 2006.

[6] H. Wang, L. Kondi, A. Luthra and S. Ci, *4G Wireless Video Communications*, A John Wiley and Sons, Ltd, 2009.

[7] A. Paulraj, R. Nabar and D. Gore, *Introduction to Space-Time Wireless Communication*, Cambridge University Press, 2003.

[8] M. Tianyu Mao, et al., "STBC-VBLAST for MIMO Wireless Communication Systems," *IEEE International Conference on Communications*, vol. 4, pp. 2266–2270, 2005.

[9] W. C. Freitas, et al., "Hybrid MIMO Transceiver Scheme with Antenna Allocation and Partial CSI at Transmitter Side," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1–5, September 2006.

[10] B. Vucetic and J. Yuan, "Space-Time Coding," *West Sussex, England ; Hoboken, NJ : Wiley, c2003. TK5102.92.V82*, 2003.

[11] S. Chakraborty, T. Frankkila, J. Peisa and P. Synnergren, *IMS Multimedia Telephony over Cellular Systems*, John Wiley & Sons, Ltd, 2007.

[12] S. M. Alamouti, "A Simple Transmit Diversity Technique for Wireless Communications," *IEEE Journal, Selected Areas in Communications*, vol. 16, pp. 1451–1458, October 1998.

[13] H. Weiwei, et al., "H.264 Motion Estimation Algorithm Based on Video Sequences Activity," *Journal of Electronics (China)*, vol. 25, pp.125–128, January 2008.

[14] ITU-T P.910 Recommendation:" Subjective video quality assessment methods for multimedia applications ", September 1999.

[15] Subjective evaluation: "http://www.gtel.ufc.br/~martin/", January 2010.

Walter Freitas Jr. was born on 1976 at Fortaleza, Brazil. He received the B.S., M.S in Electrical Engineering from Federal University of Ceará (UFC) in Fortaleza, Brazil in 2000 and 2002 respectively. In 2006 he obtained his Ph.D. from the Teleinformatic Engineering Department of the Federal University of Ceará (UFC) in Fortaleza, Brazil. Walter was a senior research of Nokia Technology Institute for two years. He is now a professor at UFC and his main area of interest concerns wireless communication systems and the application of link adaptation techniques, MIMO systems and space-time coding. Mr. Freitas is a member of IEEE Communications Society and of the Brazilian Communications Society.

Francisco R. P. Cavalcanti received his B.Sc. and M.Sc. degrees in Electrical Engineering from Federal University of Ceará, Brazil, in 1994 and 1996 respectively and a D. Sc. degree in Electrical Engineering from State University of Campinas, Brazil, in 1999. Upon graduation he joined the Federal University of Ceará, at Fortaleza, Brazil, where he is an Associate Professor and teaches Wireless Communications and Project Management at the Teleinformatics Engineering Department. In 2000 he founded and since then has directed GTEL, a research laboratory focused on the advancement of wireless communications technologies. At GTEL he manages a program of research projects in wireless communications sponsored by the Ericsson Innovation Center in Brazil. He has edited one book, published over 100 conference and journal papers and deposited 3 international patents dealing with subjects such as radio resource management, cross-layer algorithms and signal processing for wireless communications. Prof. Cavalcanti holds a Technological Development and Innovation scholarship from CNPq (Brazilian National Council for Scientific Development). He also holds a "Leadership and Management" professional certificate from the Massachusetts Institute of Technology.

Martin B. Obando received the Engineering degree in Electronic and Telecommunication from the University of Cauca, Colombia, 2000. Specialized degree in Software Development from University of San Buenaventura, Cali, Colombia, 2002. The M.Sc. degree in Electrical Engineering from the University of Brasilia, Brasilia, Brazil, 2005. He is currently pursuing the D.Sc. degree at the Department of Teleinformatic Engineering, in the Federal University of Ceará, Fortaleza, Brazil. His main research interests are in cross-layer algorithms, video transmission and video processing.