DISTORTION–AWARE JOINT SCHEDULING AND RESOURCE ALLOCATION FOR WIRELESS VIDEO TRANSMISSION

Angeliki V. Katsenou, Lisimachos P. Kondi and Evangelos Papapetrou

Department of Computer Science & Engineering, University of Ioannina, GR-45110, Ioannina, Greece

ABSTRACT

The performance of *Direct Sequence Code Division Multiple Access* (DS–CDMA) based systems highly depends on the interference caused by multiple transmissions. Interference reduction may result by using time scheduling of the transmitting nodes and by allocating accordingly the power level, and the source and channel coding rate for each node. In this work, we study the joint intra–cell scheduling and resource allocation for wireless video transmission in a hybrid DS–CDMA over *Time Division Multiple Access* (TDMA) network. We examine the case according to which a subset of the network nodes are able to simultaneously transmit (at the same time slot). The key issue is to jointly select the subset of the nodes for transmission per time slot and to allocate the available resources so that a function of the end–to–end video distortion of each node over a time frame is minimized. The experimental results show that our proposed approach results in enhanced end–to–end video quality compared to a similar DS–CDMA system.

***Index Terms*—** Video Transmission, Scheduling, Resource Allocation, TDMA, DS–CDMA.

1. INTRODUCTION

Recent advances in wireless network and video technologies have provided a plethora of different services, e.g. surveillance systems, environmental monitoring, and object tracking. In the present paper we consider the transmission of delay–sensitive video data over *Wireless Visual Sensor Networks* (WVSNs). WVSNs pose challenges to the wireless video transmission due to the error prone nature of the wireless environment. Furthermore, a major issue in multi–access WVSNs, such as *Direct Sequence Code Division Multiple Access* (DS–CDMA) WVSNs, is the interference among the transmitting nodes. Namely, each node's transmission causes interference to the other nodes' video transmissions, resulting in video quality degradation at the receiver. In most cases, due to the different rate–distortion characteristics of each recorded video, each node has different resource requirements. In order to reduce the effects of the interference caused by the simultaneous video transmission of the neighboring visual sensors, we need to establish a joint network resource allocation aiming at the enhancement of the global video quality. Another key challenge in DS–CDMA systems is to restrict the number of codes that are used for spreading. Even if the spreading codes used are orthogonal to each other, transmissions from one node cause interference to the other nodes. Hence, using scheduling helps reduce the number of required spreading codes and, consequently, helps reduce interference.

Berggren et al. [1] study a joint power control and intracell scheduling problem for supporting downlink data services in a DS–CDMA system. The authors suggest that a one–by–one transmission within a cell achieves energy savings and enhanced system capacity. Nevertheless, this cannot be applied on systems with a high number of nodes that are transmitting data under time constraints (such as video sequences). The problem addressed in [2] is to select a subset of a CDMA network nodes for transmission and jointly determine the modulation scheme, the coding scheme, the number of codes and the transmission power for those nodes with the aim of maximizing the system throughput. However, the experiments do not include video data transmission. Moreover, the maximization of the system throughput does not necessarily result in the end–to–end quality enhancement of the delivered data. Gong et al. [3] focus on the resource allocation problem. Particularly, they propose a joint bandwidth and power allocation for a multi–user decode–and–forward relay network aiming to increase the network links' capacity. Another recent work [4] considers the game–theoretic scheduling and joint channel and power allocation under the objective of the throughput maximization in *Orthogonal Frequency Division Multiple Access* (OFDMA) based cognitive radio systems. This work uses the symmetric Nash bargaining solution as an optimization criterion. All the aforementioned studies focus on the enhancement of network performance related metrics and do not take into consideration the impact on the end–to–end quality.

The impact of the resource allocation on the video quality has been considered in the recent literature. For example, in a DS–CDMA relay WVSN, the resources are allocated for both the source and the relay nodes aiming at optimizing the end–

Effort sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant number FA8655-12-1-0001. The U.S Government is authorized to reproduce and distribute reprints for Governmental purpose notwithstanding any copyright notation thereon.

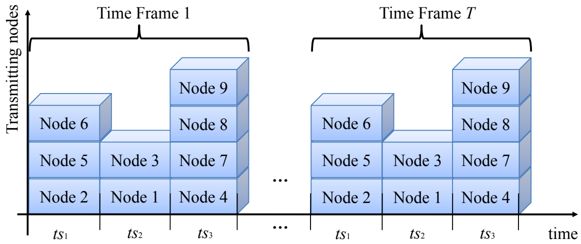


Fig. 1. Example of scheduling in the proposed HTCDMA system, where the time frame is divided in three time slots.

to–end video quality at the receiver [5]. Another resource allocation problem, namely the bit rate and transmission time allocation among a number of wireless video stations over an 802.11e–based network is considered in [6]. Game–theoretic criteria are used and the optimization objective is to increase the quality impact at the receiver.

In this paper, we focus on achieving high end–to–end video quality on a hybrid DS–CDMA, TDMA network. To this end, we need to reduce the effects of the interference caused by the simultaneous transmissions of the neighboring nodes. Therefore, we formulate the scheduling and the resource allocation as a joint end–to–end video quality–driven optimization problem. We implement an approach which considers the optimality of the solution over a specific number of time slots (time frame). According to our best knowledge, this joint problem formulation has not been considered so far in the literature. The rest of the paper is organized as follows. In Section 2, the background information for the considered system is provided. The problem formulation and the proposed approach are detailed in Section 3. The experimental results are presented in Section 4, and conclusions are drawn in Section 5.

2. CONSIDERED SYSTEM

In our system, we take advantage of the interference reduction characteristics of both DS–CDMA and TDMA. Particularly, the time scheduling property of TDMA is used to coordinate transmissions of DS–CDMA–based nodes. Thus, we utilize a hybrid DS–CDMA over TDMA system, henceforth denoted as HTCDMA. According to this HTCDMA technique, time is divided in time frames of duration t_f and each time frame is divided in N time slots of duration ts. Hence, $t_f = N\,ts$. During each time slot, more than one nodes are allowed to transmit simultaneously on a common frequency channel, as depicted in the example of Fig. 1. Each node is assigned a unique spreading code, which can be reused in consecutive time slots. Furthermore, the total transmission bit rate R_{total} in bps is equally shared among the time slots of a time frame, so that

$$R_k = \frac{R_{\text{total}}}{N} \ , \tag{1}$$

where R_k is the bit rate for each node k (with $k = 1, 2, \dots, K$) within a time slot.

A single–cell single–hop system with K wireless nodes and one base station BS is considered. In each time slot, for a single bit transmission, L chips are transmitted by a node, hence each node k is associated with a spreading code s_k. This means that, in order to transmit one bit of a bitstream, node k actually transmits $b_k\, s_k$, which is a vector of L chips with $b_k \in \{-1, 1\}$, depending on the value of the transmitted bit. The thermal and background noise are considered negligible compared to interference from the other nodes in the system. We assume that the interference received from all other nodes at the node of interest can be modeled as additive white Gaussian noise [7]. Each node k operates at a power level $S_k = E_k\, R_{\text{total}}$ in W, where E_k is the received energy–per–bit. Then, the energy–per–bit to *Multiple Access Interference* (MAI) ratio becomes

$$\frac{E_k}{I_0} = \frac{S_k / R_{\text{total}}}{\sum\limits_{j=1, j\neq k}^{K} S_j / W_{\text{t}}} \ , \tag{2}$$

where $I_0/2$ is the two sided noise power spectral density due to MAI in W/Hz, S_j is the received power of the interfering nodes in W, and W_t is the total bandwidth in Hz. The received power of a node can be directly computed from the transmission power by using a suitable radio propagation model depending on the system topology and the surrounding environment [8].

The H.264/AVC video coding standard is used for the source coding of the captured videos. For channel coding, *Rate Compatible Punctured Convolutional* (RCPC) codes are deployed [9]. For bit error probability estimation, the Viterbi's upper bounds are used, namely

$$P_{\text{be},k} \leq \frac{1}{P} \sum_{d=d_{\text{dfree}}}^{\infty} c_d P_{d,k} \ , \tag{3}$$

where P is the period of the code, d_{free} is the free distance of the code, c_d is the information error weight, and P_d denotes the probability of selecting the wrong path at distance d. For an additive white Gaussian noise channel with *Binary Phase Shift Keying* (BPSK) modulation, $P_{d,k}$ is given by $Q\left(\sqrt{2dR_{\text{c},k}E_k/I_0}\right)$, where $R_{\text{c},k}$ is the channel coding rate of node k.

Due to compression losses and channel errors, the video quality at the receiver is distorted. The channel errors are random, therefore we estimate the expected value of the video distortion $E\{D_{\text{s+c},k}\}$. For this reason we use the *Universal Rate Distortion Characteristics* (URDCs), as in [10]:

$$E\{D_{\text{s+c},k}\} = \alpha_k \left[\log_{10}\left(\frac{1}{P_{\text{be},k}}\right)\right]^{-\beta_k} \ , \tag{4}$$

where $P_{\text{be},k}$ is the bit error probability for node k, and the positive parameters α_k, β_k depend both on the motion level of the video sequence and the source coding rate $R_{\text{s},k}$. Since α_k, β_k depend on the source coding rate of node k, and $P_{\text{be},k}$ depends on the channel coding rate of node k and the received power from all interfering nodes, $E\{D_{\text{s+c},k}\}$ can be expressed as a function of the source coding rate, $R_{\text{s},k}$, the channel coding rate, $R_{\text{c},k}$, as well as the power level S of all nodes, i.e. $E\{D_{\text{s+c},k}\} = f(R_{\text{s},k}, R_{\text{c},k}, S)$. The parameters α_k, β_k are determined using least squares from a few $(E\{D_{\text{s+c},k}\}, P_{\text{be},k})$ pairs, which are obtained at the encoder using the *Recursive Optimal per-Pixel Estimate* (ROPE) as proposed in [11].

3. JOINT SCHEDULING AND RESOURCE ALLOCATION

Prior to the node scheduling and the WVSN resource allocation, each node performs a preprocessing phase, where the rate–distortion characteristics (parameters α_k and β_k) of the so far recorded scenes (a small number of video frames) for the different available source coding rates are estimated. We consider that these characteristics remain the same for the duration of a *Group of Pictures* (GoP) of a recorded video. At the end of each GoP we check if these characteristics have been significantly altered, and update them accordingly. A similar preprocessing phase has been used in [12]. The nodes communicate those rate–distortion characteristics to the BS. The BS exploits this information to optimally determine the subset of the transmitting nodes per time slot and the optimal resources (power level, source and channel coding rate) per transmitting node over a whole time frame. This process is repeated periodically every T time frames. Parameter T can be set by the BS. Moreover, it is assumed that each node is allowed to transmit only once within a time frame. Under the assumption of a constant spreading code length, L, and the constraint of identical chip rate, $R_{\text{chip},k}$, for each node k, the transmission bit rate for each node R_k is correspondingly constant during a time slot, since $R_k = R_{\text{chip},k}/L$. Furthermore, the constant bit rate R_k secures that the fraction $R_{\text{s},k}/R_{\text{c},k}$ is identical for each node.

Under these assumptions, the present paper copes with the problem of selecting a subset M_i of nodes for transmission per time slot i and the joint allocation of the source coding rate, $R_{\text{s},k}$, the channel coding rate, $R_{\text{c},k}$, as well as the power level S_k of each node k for all the time slots of a time frame, such that a function $\mathcal{F}(.)$ of the overall end–to–end expected video distortion is minimized. According to this, we need to employ a distortion–related function $\mathcal{F}(.)$ that will guide the optimization process. Therefore, in the present paper we utilize the *Nash Bargaining Solution* (NBS) and the *Minimization of the Average Distortion* (MAD).

3.1. Nash Bargaining Solution

According to NBS, each WVSN node joins the cooperative bargaining with the aim of achieving a higher utility than what it could achieve if it were to operate selfishly, without cooperation with its interfering nodes. In this bargaining, the utility of a player is interpreted in terms of end–to–end video quality. So, let the *utility function* U_k be the *Peak Signal to Noise Ratio* (PSNR) of the received video, namely:

$$U_k = 10 \log_{10} \frac{255^2}{E\{D_{s+c,k}\}} , \tag{5}$$

where $E\{D_{s+c,k}\}$ is the expected video distortion for node k, that is directly related to the source coding rate $R_{\text{s},k}$, the channel coding rate $R_{\text{c},k}$, and the power level S_k. It is reasonable to assume that each node is a rational player and would agree to cooperate only if the utility it would get is at least as high as what it would get without cooperation. This minimum acceptable amount of utility is called *disagreement point* $dp = (dp_1, dp_2, \ldots, dp_K)^\top$.

We define $\mathbf{U}$ as the feasible set of all possible utility allocations $U = (U_1, U_2, \ldots, U_K)^\top$. Each member of $\mathbf{U}$ results from a different combination of source coding rates, channel coding rates, and power levels. The NBS $\mathcal{G}(\mathbf{U}, dp)$ is a member of the feasible set that satisfies the axioms of *Feasibility*, *Pareto Efficiency*, *Invariance to Equivalent Utility Representations* and *Independence of Irrelevant Alternatives* [13].

In order to find the NBS, $\mathcal{G}(\mathbf{U}, dp)$ with $dp \in \mathbf{U}$, we have to maximize the Nash product. Particularly, we determine the subset of transmitting nodes per time slot i and the utilities vector U such that the Nash product is maximized over a time frame:

$$\max_{\substack{M_1, M_2, \ldots, M_N, \\ R_\text{s}, R_\text{c}, S}} \prod_{i=1}^{N} \prod_{k \in M_i} (U_k - dp_k)^{bp_k} , \tag{6}$$

subject to

$$U_k \geq dp_k, \tag{7}$$
$$S_{\min} \leqslant S_k \leqslant S_{\max}, \tag{8}$$
$$(R_{\text{s},k}, R_{\text{c},k}) \in \{(R_\text{s}^1, R_\text{c}^1), \ldots, (R_\text{s}^Z, R_\text{c}^Z)\},\ Z \in \mathbb{N}^* \tag{9}$$
$$\frac{R_\text{s}^1}{R_\text{c}^1} = \frac{R_\text{s}^2}{R_\text{c}^2} = \ldots = \frac{R_\text{s}^Z}{R_{\text{c},}^Z} = R_k, \tag{10}$$

where Z is the number of the available source and channel coding rates, and M_i is the set of the indices of the selected nodes for transmission per time slot i, which satisfies the following properties:

(i) $M_i \neq \varnothing, \forall\, i \in \{1, 2, \ldots, N\}$;

(ii) $M_1 \bigcap M_2 \bigcap \ldots \bigcap M_N = \varnothing$;

(iii) $M_1 \bigcup M_2 \bigcup \ldots \bigcup M_N = \{1, 2, \ldots, K\}$.

The bargaining powers $bp = (bp_1, bp_2, \ldots, bp_K)^\top$, that express which node is more favored by the bargaining rules [13], are all considered equal to $1/K$. Besides this, we assume that dp is the lowest acceptable PSNR value and is determined by the application requirements.

3.2. Minimization of the Average Distortion

The MAD criterion minimizes the average end–to–end video distortion over a time frame by optimally determining the subset of transmitting nodes per time slot i, as well as the source coding rate, $R^*_{\mathrm{s},k}$, channel coding rate, $R^*_{\mathrm{c},k}$, and power level, S^*_k for each node k, i.e.,

$$\min_{\substack{M_1,M_2,\ldots,M_N,\\ R_{\mathrm{s}},R_{\mathrm{c}},S}} \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{|M_i|}\sum_{k\in M_i} E\{D_{\mathrm{s+c},k}\}\right) \tag{11}$$

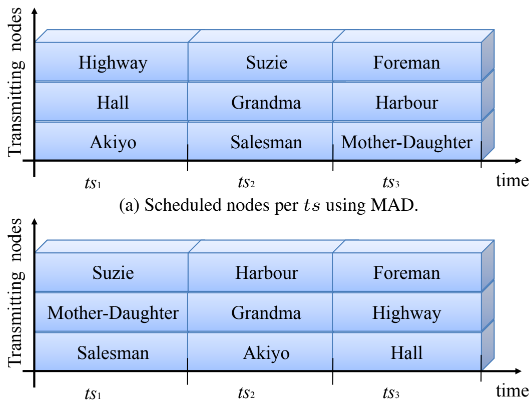
subject to the constraints of Eqs. (8)–(10) and where $|.|$ denotes the cardinality of a set.

4. EVALUATION AND RESULTS

For the evaluation of our proposed approaches, a number of experiments were conducted. We considered a single–hop WVSN topology, where each node is equidistant and has clear line of sight with its BS. Each node may record a scene of different motion. Specifically, the nine YUV QCIF sequences listed in Table 1 (with 15 fps frame rate) were used. The available bandwidth is $W_t = 1$ MHz and the total transmission rate for the HTCDMA system is $R_{\text{total}} = 288$ kbps. Moreover, we assumed that a time frame is divided in $N = 3$ time slots with $ts = 10$ ms duration each, and that three nodes are allowed to simultaneously transmit during each time slot. We present the results of the evaluation that was based on comparing the proposed HTCDMA system to a DS–CDMA with the same bandwidth, but with lower bit rate, i.e. 96 kbps, in order for each node to experience the same bit rate in both systems (see Eq. 1). For the NBS, the disagreement point for both systems and for all transmitting nodes was set to 24 dB.

In the proposed scheme, the transmit powers assume continuous values within the range $[0.0500, 0.5000]$ W. On the other hand, the selected subset of transmitting nodes per time slot and the source and channel coding rates of each node are selected from a discrete set, i.e. *Coding Set* $CS \in \{1 : (32\text{ kbps}, 1/3),\ 2 : (48\text{ kbps}, 1/2),\ 3 : (64\text{ kbps}, 2/3)\}$. Hence, the formulated optimization problems are mixed–integer problems. We employ a stochastic optimization technique, called *Particle Swarm Optimization* (PSO) [14], in order to efficiently solve the formulated problem. PSO has been used before in similar resource allocation problems over WVSNs as in [5, 12]. Considering the stochastic nature of the PSO algorithm, 30 independent experiments were executed for each problem instance to ensure the validity of the results [14].

The different node subset selection for the two criteria in the proposed HTCDMA system is illustrated in Fig. 2. For both criteria, the scheduled nodes per ts have different rate–distortion characteristics. For example, the "Mother-Daughter" video, which is of low motion, is transmitted in the same time slot with "Suzie" and "Salesman", which are



(a) Scheduled nodes per ts using MAD.

(b) Scheduled nodes per ts using NBS.

Fig. 2. Scheduled nodes for transmission in a time frame.

Table 1. Results for HTCDMA and DS–CDMA.

Proposed HTCDMA							
	MAD			NBS			
Nodes	PSNR	CS	S	PSNR	CS	S
1.Akiyo	33.4457	1	0.0500	32.2730	1	0.0523
2.Salesman	32.0021	2	0.0781	31.5022	2	0.0707
3.Grandma	32.9401	1	0.0500	32.5180	1	0.0500
4.Mother-Daughter	32.2943	1	0.0500	33.1749	1	0.0500
5.Harbour	31.2454	3	0.0983	31.0477	3	0.0954
6.Hall	33.6395	1	0.0597	33.7675	1	0.0500
7.Highway	33.3946	1	0.0604	33.9298	1	0.0632
8.Suzie	32.8205	1	0.0625	32.6780	1	0.0580
9.Foreman	32.4454	1	0.0533	33.5102	1	0.0632

DS–CDMA						
	MAD			NBS		
Nodes	PSNR	CS	S	PSNR	CS	S
1.Akiyo	30.4438	1	0.0528	31.0597	1	0.0523
2.Salesman	27.8555	1	0.0670	27.1846	1	0.0589
3.Grandma	30.7121	1	0.0500	31.4448	1	0.0500
4.Mother-Daughter	30.6904	1	0.0539	31.4871	1	0.0538
5.Harbour	26.5457	2	0.0980	26.9821	2	0.0868
6.Hall	29.9832	1	0.0651	30.4413	1	0.0630
7.Highway	29.7705	1	0.0653	30.0973	1	0.0627
8.Suzie	29.9789	1	0.0640	30.4265	1	0.0620
9.Foreman	26.9850	1	0.0736	27.3091	1	0.0609

considered videos of medium motion. This is a result of the employed distortion–aware functions and is important for the video quality enhancement during each ts.

Table 1 reports the obtained results for the network resource allocation using the proposed method for the considered HTCDMA and the DS–CDMA system. A close inspection of the results reveals that strong channel coding is selected for the majority of the nodes in all cases. Only for some medium or high motion nodes we have the selection of weaker channel coding rate. Besides this, the total power that is required for the video transmission is slightly lower for the NBS than the MAD criterion for both systems (1.76% for HTCDMA and 7.16% for DS–CDMA).

Regarding the resulting quality gain of the proposed HTCDMA over DS–CDMA, we consider the PSNR differ-

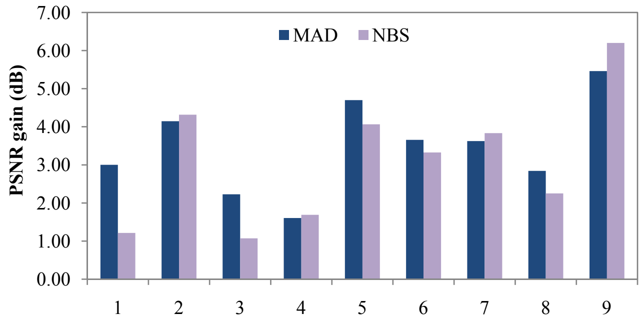


Fig. 3. PSNR gain (in dB) per node of HTCDMA over DS–CDMA for both criteria.

ence $(PSNR_k^{\text{HTCDMA}} - PSNR_k^{\text{DS-CDMA}})$, for each node k and for the two different criteria. This is depicted in Fig. 3. For both optimization criteria, the average PSNR gain is greater than 3 dB (particularly 3.47 dB for MAD and 3.11 dB for NBS), which is a considerable quality enhancement. Furthermore, it is remarkable that the quality gain is higher for the videos with higher amount of motion. This is due to the fact that HTCDMA reduces the interference among the transmitting nodes. Moreover, the scheduling is performed with regard to the resulting video distortion, thus the optimal combination of nodes is selected for transmission.

5. CONCLUSIONS

In this study, we considered the problem of the quality–driven joint node scheduling and resource allocation in a hybrid DS–CDMA over a TDMA system. In the proposed system, we use the Nash Bargaining Solution and the Minimization of the Average Distortion to decide on which nodes transmit per time slot as well as what power level, source and channel coding rates should be used in order to enhance the delivered video quality. For the formulated mixed–integer optimization problem, the Particle Swarm Optimization is employed. The evaluation of the proposed approach has shown that our approach offers the benefit of enhanced end–to–end video quality at the receiver compared to a similar DS–CDMA system that allows simultaneous transmission of all nodes.

6. REFERENCES

[1] F. Berggren, S.-L. Kim, R. Jantti, and J. Zander, "Joint power control and intracell scheduling of DS-CDMA nonreal time data," *IEEE Journ. on Selected Areas in Communications*, vol. 19, no. 10, pp. 1860–1870, October 2001.

[2] V. G. Subramanian, R. A. Berry, and R. Agrawal, "Joint scheduling and resource allocation in CDMA systems," *IEEE Trans. on Information Theory*, vol. 56, no. 5, pp. 2416–2432, May 2010.

[3] X. Gong, S. A. Vorobyov, and C. Tellambura, "Joint bandwidth and power allocation in wireless multi-user decode-and-forward relay networks," in *Proc. IEEE International Conference on Acoustics Speech and Signal Processing*, March 2010, pp. 2498–2501.

[4] Q. Ni and C. C. Zarakovitis, "Nash Bargaining Game Theoretic Scheduling for Joint Channel and Power Allocation in Cognitive Radio Systems," *IEEE Journ.on Selected Areas in Communications*, vol. 30, no. 1, pp. 70–81, January 2012.

[5] E. G. Datsika, A. V. Katsenou, L. P. Kondi, E. Papapetrou, and K. E. Parsopoulos, "Priority–based cross–layer optimization for multihop DS–CDMA visual sensor networks," in *Proc. IEEE Intern. Conference on Image Processing*, Orlando, Florida, 2012, pp. 1102–1104.

[6] H. Park and M. van der Schaar, "Fairness strategies for wireless resource allocation among autonomous multimedia users," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 20, no. 2, pp. 297–309, Feb. 2010.

[7] Y. S. Chan and J. W. Modestino, "A joint source coding–power control approach for video transmission over CDMA networks," *IEEE Journ. on Selected Areas in Communications*, vol. 21, no. 10, pp. 1516–1525, 2003.

[8] T. Rapaport, *Wireless Communications Principles and Practise*, Prentice Hall Communications Engineering and Emerging Technologies Series. Prentice Hall, 2nd edition, 2001.

[9] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. on Communications*, vol. 36, no. 4, pp. 389–400, April 1988.

[10] L. P. Kondi and A. K. Katsaggelos, "Joint source-channel coding for scalable video using models of rate distortion functions," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Salt Lake City, 2001, pp. 1377–1380.

[11] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE Journ. on Selected Areas on Communications*, vol. 18, no. 6, pp. 966–976, June 2000.

[12] A. V. Katsenou, L. P. Kondi, and K. E. Parsopoulos, "Resource Management for Wireless Visual Sensor Networks Based on Individual Video Characteristics," in *Proc. IEEE International Conference on Image Processing*, Brussels, September 2011, pp. 149–152.

[13] K. Binmore, *Playing for Real: A Text on Game Theory*, Oxford University Press, 2007.

[14] K. E. Parsopoulos and M. N. Vrahatis, *Particle Swarm Optimization and Intelligence: Advances and Applications*, Information Science Publishing (IGI Global), 2010.