

Motion-Related Resource Allocation in Dynamic Wireless Visual Sensor Network Environments

Angeliki V. Katsenou, *Student Member, IEEE*, Lisimachos P. Kondi, *Senior Member, IEEE*,
and Konstantinos E. Parsopoulos, *Member, IEEE*

Abstract—This paper investigates quality-driven cross-layer optimization for resource allocation in direct sequence code division multiple access wireless visual sensor networks. We consider a single-hop network topology, where each sensor transmits directly to a centralized control unit (CCU) that manages the available network resources. Our aim is to enable the CCU to jointly allocate the transmission power and source-channel coding rates for each node, under four different quality-driven criteria that take into consideration the varying motion characteristics of each recorded video. For this purpose, we studied two approaches with a different tradeoff of quality and complexity. The first one allocates the resources individually for each sensor, whereas the second clusters them according to the recorded level of motion. In order to address the dynamic nature of the recorded scenery and re-allocate the resources whenever it is dictated by the changes in the amount of motion in the scenery, we propose a mechanism based on the particle swarm optimization algorithm, combined with two restarting schemes that either exploit the previously determined resource allocation or conduct a rough estimation of it. Experimental simulations demonstrate the efficiency of the proposed approaches.

Index Terms—Resource allocation, cross-layer optimization, DS-CDMA, H.264/AVC, clustering, nash bargaining solution, visual sensor networks, particle swarm optimization.

I. INTRODUCTION

RECENT advances in wireless communication systems and digital electronics have enabled the development of low-cost, low-power, multi-functional sensor nodes that are small in size and communicate untethered over short distances [1]. Our interest is focused on *Wireless Visual Sensor Networks* (WVSNs) that impose several network, time, and quality constraints in video transmission. Visual sensors can harvest information from the physical environment, perform simple processing on the extracted data and transmit it to remote locations. Mobile or stationary cameras can be used for multiple purposes, such as surveillance, tracking of physical activities and analysis of distant explorations of hazardous areas, among others [2].

Traditional WVSNs are comprised of distributed nodes equipped with video cameras and a *Centralized Control Unit*

(CCU). Such a typical WVSN is considered in this paper. The nodes share the same frequency and communicate with the CCU over the network layer. The CCU applies channel and source decoding to obtain the received video from each node. A significant issue that arises in WVSNs is the resource allocation among the network nodes. Aiming at achieving an acceptable quality level of the delivered videos, each node has different resource requirements (e.g., source coding rate and delay constraints). Moreover, the transmission of each node has an effect on the other nodes due to interference. Thus, the received video quality is degraded. Taking into consideration the aforementioned factors along with each node's time-varying resource requirements, we propose a quality-driven joint strategy for the optimal allocation of the network resources, i.e., transmission power, source and channel coding rate. Two different node grouping approaches are considered and evaluated.

Another important issue that emerges in various multimedia applications, such as area surveillance, is the prioritization of the visual sensors in order to enhance the delivered video quality for selected cameras. Specifically, the network's resource allocation can be heavily dependent on the video data "importance", as it is dictated by the corresponding application. More specifically, the recordings of the cameras with high motion require a larger source coding rate than low motion videos. High motion videos are more sensitive to channel losses. Furthermore, assuming a fixed budget for source and channel coding rate, less strong channel coding will have to be used for high motion videos. On the other hand, low motion videos can be encoded at a lower source coding rate. This will leave more bits available for channel coding and will allow adequate transmission using a lower transmission power. Thus, cameras capturing scenes of high motion have different requirements than cameras capturing scenes of low motion. Therefore, we use a prioritization scheme based on the amount of motion detected in a video sequence. According to it, visual sensors that record higher motion receive a proportionally higher priority compared to the low motion sensors, and hence achieve relatively higher delivered video quality.

A last challenging issue that is addressed in this paper refers to the dynamic nature of the WVSN environment, which is implied by the changing motion level of the recorded scenes, the changing application requirements, network load variations etc. Thus, we propose an online optimization procedure that is based on the *Particle Swarm Optimization* (PSO) algorithm, combined with two different swarm initialization schemes. Our procedure is capable of retaining high end-to-end video

Manuscript received May 8, 2012; revised August 2, 2013; accepted October 2, 2013. Date of publication October 18, 2013; date of current version November 7, 2013. This work was supported by the Air Force Office of Scientific Research, Air Force Material Command, and USAF under Grant FA8655-12-1-0001. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Sanghoon Lee.

The authors are with the Department of Computer Science and Engineering, University of Ioannina, Ioannina 45110, Greece (e-mail: akatseno@cs.uoi.gr; lkoni@cs.uoi.gr; kostasp@cs.uoi.gr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2013.2286323

quality by properly re-allocating the network's resources each time a node's detected motion level changes significantly.

A. Related Work

The problem of establishing a cooperative framework for the network resource allocation among wireless nodes has been considered in previous work [3]–[9]. Furthermore, game theory has proved to be a field that can provide efficient methodologies for such problems.

The problem of allocating the bit rate simultaneously for multiple video streams across time slots using a central controller was addressed in [3] based on a competitive equilibrium approach, the Edgeworth box. The authors made an assumption that adjacent time slots have similar rate-distortion curves, therefore in each time slot and for each node the central controller reoptimizes the resource allocation for the current and the future time slots. The problem objective was to maximize the network throughput, which does not necessarily lead to the maximum possible video quality at the receiver. Moreover, assuming similar rate-distortion curves in adjacent time slots does not result in good quality for the cases where a sudden change in the captured motion level occurs. Opposed to that, our proposed update mechanism effectively addresses low to high changes in the rate-distortion curves and quickly re-allocates the network resources.

A similar bit rate allocation problem among a number of wireless video stations over an 802.11e-based network was tackled by employing two different bargaining solutions, namely the *Nash Bargaining Solution* (NBS) and the *Kalai-Smorodinsky Bargaining Solution* (KSBS) in [5], [6]. In both works the utility function was based on the end-to-end video distortion with the motivation to depict the quality impact of the resource allocation. Besides that, an algorithm to determine the bargaining powers according to the allocated bit rate per user is proposed in [5]. In [6], the previous research problem was extended by considering the wireless stations as autonomous users that are able to exchange information and use the same game-theoretic criteria to allocate the bit rate and the transmission time. In our work, we consider the NBS criterion and a similar utility function but for a different resource allocation problem. Moreover, we utilize different rules for the determination of the bargaining powers of the nodes. Nevertheless, both [5], [6] assumed static network conditions for which the resource allocation is considered valid. With a similar system set-up, the authors of [4] extended the work of [5], [6] for multihop transmission and studied three types of resource allocation solutions (centralized optimization, game-theoretic approaches and distributed greedy approaches). Moreover, they pointed out the weakness of their centralized optimization approach to adapt to the source and network dynamics due to its high complexity.

Apart from the bit rate allocation problem, NBS was also used in [7] for the cross-layer resource allocation of a *Direct Sequence Code Division Multiple Access* (DS-CDMA) WVSN with an N -class motion-based node partitioning. Equal bargaining powers were used for each visual sensor. Thus, the transmission power and network resources were

equally allocated for all sensors within each class. However, no specific clustering algorithm was proposed for the partitioning and the presented resulting quality was estimated based on the representative centroid values of each cluster. Opposed to that, in the present work we perform k -means clustering and, after the resource allocation based on the centroid values, we estimate the video quality by using each node's real parameters. In another CDMA-based WVSN, where the processing gain per node can be adapted, we performed a quality-driven resource allocation to the nodes [8]. Nevertheless, that work assumed that the network conditions are static and all nodes have the same bargaining powers in the game. The NBS was also used for the resource allocation in a multihop video transmission [9]. The network resources were allocated for both the source and the relay nodes aiming at optimizing the end-to-end video quality at the receiver. Moreover, a motion-driven optimization criterion based on the weighted aggregation of video distortions was introduced. Furthermore, the nodes were grouped into classes based on their physical position and a similar captured amount of motion for each cluster was assumed. Again, the proposed solution concerned a certain time instance of this network. Thus, the optimization process needed to be repeated in order to ensure an up-to-date resource allocation and acceptable level of network performance.

Finally, in our work we achieve motion-related proportional video quality enhancement of the nodes. Particularly, we employ motion-related bargaining powers to formulate our NBS criterion. So far, the resource allocation methods of power and joint source-channel coding rate in video communication applications have not taken into consideration this requirement and formulation. Related work in video transmission has only considered the tradeoff of video quality and channel coding rate selection with regard to the video distortion impact [10]–[13]. However, the proposed approaches considered only a single video sequence transmission. Moreover, the proposed approaches [11], [13] were applied on a binary symmetric channel and, additionally, the assumption that the distortion resulting after compression can be ignored was made in [13].

B. Contribution and Structure of the Paper

We study how the CCU can efficiently allocate the network resources of a DS-CDMA WVSN taking into consideration the individual rate-distortion characteristics of each node. The paper extends the concepts introduced in our previous work [14], [15], which only included preliminary results, and brings the following contributions:

- (a) *Investigating the quality, power and complexity tradeoffs by considering individual nodes versus clustered nodes:* We study and compare two approaches for resource allocation in WVSNs with a different tradeoff of quality, power and complexity. The first one allocates the resources individually for each sensor according to its individual rate-distortion characteristics, while the second clusters them according to the recorded level of motion and allocates the same resources among

TABLE I
LIST OF ABBREVIATIONS

Abbreviation	Definition	Abbreviation	Definition
AWGN	Additive White Gaussian Noise	PSNR	Peak Signal to Noise Ratio
BPSK	Binary Phase-Shift Keying	PSO	Particle Swarm Optimization
CCU	Centralized Control Unit	PSO-PS	PSO with prior resource allocation
CDMA	Code Division Multiple Access	PSO-RE	PSO with rough estimation of the resources
DS-CDMA	Direct Sequence Code Division Multiple Access	RCPC	Rate Compatible Punctured Convolutional codes
e.NBS	NBS with equal bargaining powers	ROPE	Recursive Optimal per-Pixel Estimate
MAD	Minimizing the Average Distortion	TC1	Test Case 1
MAI	Multiple Access Interference	TC2	Test Case 2
MMD	Minimizing the Maximum Distortion	URDC	Universal Rate Distortion Characteristics
NAL	Network Abstraction Layer	VCL	Video Coding Layer
NBS	Nash Bargaining Solution	w.NBS	NBS with motion-related bargaining powers
NP	Nash Product	WVSN	Wireless Visual Sensor Network

the cluster nodes. We present experimental results that clearly demonstrate the energy and complexity costs in order to achieve a higher end-to-end video quality.

- (b) *Motion-related prioritization of nodes using game theory*: The recordings of the high motion cameras are susceptible to channel errors, thus reach the CCU with low quality, opposed to the low motion videos which are more robust. It is therefore sensible that cameras capturing scenes of high motion have different requirements. So far, the resource allocation methods in video communication applications have not taken into consideration this requirement. Therefore, we provide a simple yet effective game theoretic motion-related prioritization of the network resources and enhance the end-to-end video quality of the high motion nodes. Furthermore, our approach demonstrates the advantage of computing locally at each node (not at the CCU) the motion characteristics of each recorded video. This also means that the required computation is partially decentralized.
- (c) *Dynamically changing environment*: Unlike previous work that considers static WVSNs, we consider a dynamic environment with variations in the amount of motion in the recorded scenes through time. For this reason, we introduce two PSO-based algorithms for the re-allocation of resources after each change in the environment. The two approaches differ in the swarm initialization, as follows:

- (i) *PSO with prior resource allocation (PSO-PS)*: The swarm is initialized closely to the resource allocation of its previous state, i.e., before the appearance of the change in the environment.
- (ii) *PSO with rough estimation of the resources (PSO-RE)*: The swarm is initialized by using a rough first estimation of the resource allocation, based on the expected transmission power of each node.

The rest of the paper is structured as follows. In Section II, the reference architecture of the WVSN considered is described. In Section III, the actual optimization problem is formulated and the proposed approaches and optimization criteria are detailed. The proposed optimization algorithms are detailed in Section IV. The experiments and results are discussed in Section V, and, finally, in Section VI conclusions

are derived. For the reader's convenience and easy reference, we summarize abbreviations and their definitions in Table I.

II. SYSTEM MODEL

We consider a single-hop topology with K nodes that transmit directly to the CCU. DS-CDMA is used for the physical layer, thus all nodes transmit over the same frequency. For a single bit transmission, L chips are transmitted by a node. Hence, each node k is associated with a spreading code sc_k (vector of length L). Namely, in order to transmit the i -th bit of a bitstream, node k actually transmits $b_k(i)sc_k$, which is a vector of L chips with $b_k(i) \in \{-1, 1\}$, depending on the value of the transmitted bit [16]. Consequently, the transmission bit rate R_k for node k can be expressed as:

$$R_k = \frac{R_{\text{chip},k}}{L}, \quad (1)$$

where $R_{\text{chip},k}$ is the chip rate of the reference node k .

We assume that the interference received from all other transmitting nodes can be modeled as *Additive White Gaussian Noise* (AWGN) [17]. The background and thermal noise are assumed to be negligible compared with the interference and, hence, they can be ignored. The k -th node operates at a transmission power level S_k , for each of the K nodes of the WVSN. Due to attenuation, a reduced power $S_{r,k}$ is received at the CCU. This is equal to:

$$S_{r,k} = E_k R_k, \quad (2)$$

where E_k is the energy-per-bit, and R_k is the total bit rate in bits/sec expressed as the ratio of the source coding rate $R_{s,k}$ in bits/sec to the channel coding rate $R_{c,k}$, which is a dimensionless number (see Section II-B), i.e.,

$$R_k = \frac{R_{s,k}}{R_{c,k}}, \quad (3)$$

for a node k , with $k = 1, 2, \dots, K$.

The energy-per-bit to *Multiple Access Interference* (MAI) ratio for each reference node k , becomes:

$$\frac{E_k}{I_0} = \frac{S_{r,k}/R_k}{\sum_{\substack{j=1 \\ j \neq k}}^K S_{r,j}/W_t}, \quad (4)$$

where $I_0/2$ is the two sided noise power spectral density due to MAI in Watts/Hertz, W_t is the total bandwidth in Hertz and $S_{r,j}$ is the received power of the interfering nodes in Watts [17].

In order to calculate the received power at the node of interest from the neighboring nodes, several radio propagation models can be employed depending on the terrain profile. In the present work, we assume that all visual sensors have clear line of sight and therefore we use the *Two-Ray Ground Reflection* model. This model considers both the direct path and a ground reflected propagation path between transmitter and receiver. The received power $S_{r,k}$ at distance d from the transmitting node k is proportional to the transmission power S_k in Watts and inversely proportional to the distance raised to the fourth power, i.e., $S_{r,k} \propto S_k/d^4$, and can be expressed as:

$$S_{r,k} = G_t G_r h_t^2 h_r^2 \frac{S_k}{d^4}, \quad (5)$$

where G_t , G_r , are the antenna gains in dB and h_t , h_r , are the antenna heights of the transmitter and the receiver in meters, respectively [18].

A. Source Coding

For the source coding of the captured video sequences, the H.264/AVC video coding standard is used. The H.264/AVC design consists of two conceptual layers: the *Network Abstraction Layer* (NAL) and the *Video Coding Layer* (VCL). The NAL, which was created to fulfill the network-friendly design objective, formats data and provides header information for conveyance by transport layers or storage media. All data are encapsulated in NAL units, each of which contains an integer number of bytes. The NAL unit structure provides a generic form for use in both packet-oriented and bitstream-based systems. The format of NAL units is identical in both environments, except that each NAL unit is preceded by a unique start code prefix for re-synchronization in bitstream-oriented transport systems. The VCL is specified to efficiently represent the content of the video data and fulfill the design objective of enhanced coding efficiency. It is similar in spirit to designs found in other standards in the sense that it consists of a hybrid of block-based temporal and spatial prediction in conjunction with scalar-quantized block transform coding [19], [20].

B. Channel Coding

Channel codes permit reliable communication of an information sequence over a noisy channel that introduces bit errors and distorts the transmitted signal. Extra information is added in the data sequence at the encoder utilizing an error-correcting code, which is used to recover the original data at the receiver. A common type of channel coding that is extensively used in many applications is convolutional coding. *Rate Compatible Punctured Convolutional* (RCPC) codes [21] are deployed in the present paper for channel coding. However, other types of channel codes could be used. The idea of puncturing, in combination with the concept of rate compatibility, offers the benefit of using the same Viterbi decoder for all convolutional codes derived from the same mother code. The Viterbi decoder

uses a maximum likelihood sequence estimation procedure. The upper bounds for bit error probability $P_{be,k}$ are given by the following inequality:

$$P_{be,k} \leq \frac{1}{P} \sum_{d=d_{free}}^{\infty} c_d P_{d,k}, \quad (6)$$

where P is the code period, d_{free} is the code free distance, c_d is the information error weight, and $P_{d,k}$ is the probability of the wrong path selection at distance d for the node-of-interest k [21]. $P_{d,k}$ depends on the channel conditions, namely the energy-per-bit to MAI ratio, and for coherent *Binary Phase-Shift Keying* (BPSK) modulation on an AWGN channel it is given by:

$$P_{d,k} = Q\left(\sqrt{2dR_{c,k} \frac{E_k}{I_0}}\right), \quad (7)$$

where $R_{c,k}$ is the channel coding rate and $Q(\cdot)$ is the function that expresses the probability that a standard normal random variable Z is equal to or larger than x , i.e. $Q(x) = P(Z \geq x) = (\sqrt{2\pi})^{-1} \int_x^{\infty} e^{-z^2/2} dz$ [21].

C. Video Distortion Estimation Model

Lossy compression and channel errors result in the distortion of the video quality at the receiver. In order to estimate the expected video distortion at the sender node k , we assume the following *Universal Rate Distortion Characteristics* (URDC), as in [22]:

$$E[D_{s+c,k}] = \alpha_k \left[\log_{10} \left(\frac{1}{P_{be,k}} \right) \right]^{-\beta_k}, \quad (8)$$

where $D_{s+c,k}$ is the video distortion at the decoder given the source and channel coding rate for node k , $P_{be,k}$ is the bit error probability, and the parameters $\alpha_k \in \mathbb{R}_+$, $\beta_k \in \mathbb{R}_+$ depend both on the motion level of the video sequence and the source coding rate.

Algorithm 1 details how we determine the values of parameters α_k and β_k for each node at the encoder. For the accurate estimation of the decoder distortion $E[D_{s+c,k}]$ at the encoder, the *Recursive Optimal per-Pixel Estimate* (ROPE) algorithm [23] is used. The ROPE algorithm recursively calculates the first and second moments of the decoder reconstruction of each pixel $E[D_{s+c,k}^p]$, while it accurately takes into account all relevant factors, i.e. quantization, packet loss, error propagation and error concealment. Due to the fact that ROPE uses the *Packet Loss Rate* (PLR) to compute the overall expected MSE distortion of a pixel, we associate it with the BER, as explained in [16].

Figure 1 shows the α_k and β_k values for three different source coding rates ($R_s = 32$ kbps for $CS = 1$, $R_s = 48$ kbps for $CS = 2$ and $R_s = 64$ kbps for $CS = 3$) for 10 well-known video sequences. These video sequences have variable amounts of motion. In order to quantify the amount of motion, we used the *Full Search Block Matching Algorithm* [24] and estimated the motion vectors of the video sequences. Then, we computed the mean magnitude of the motion vectors, which revealed a clear grouping of the videos according to the amount of motion. Particularly,

Algorithm 1 Pseudocode for Computing Parameters (α, β) at Each Node k

```

1: for three different  $P_{be}$  do
2:   for each frame  $m$  do
3:     for each pixel  $p$  do
4:       Estimate  $E[D_{s+c,k}^p]$  according to the ROPE algo-
         rithm [23].
5:     end for
6:     Estimate  $E[D_{s+c,k}^m] = \frac{1}{\# \text{ of pixels}} \sum_p E[D_{s+c,k}^p]$  for frame
          $m$ .
7:   end for
8:   Estimate the Mean  $E[D_{s+c,k}]$  for the video sequence.
9: end for
10: Use Least Squares to estimate  $(\alpha_k, \beta_k)$  from the pairs of
     $(E[D_{s+c,k}], P_{be,k})$ .

```

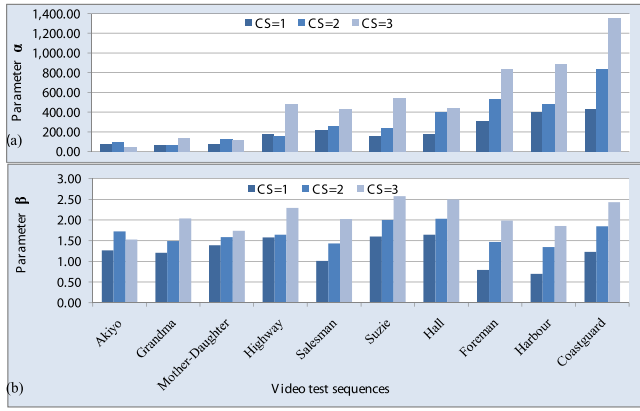


Fig. 1. Parameter α and β values for three different source coding rates.

- (i) for “Akiyo”, “Grandma” and “Mother-Daughter” the mean magnitude of the motion vectors ranges from 0.0119 to 0.3630;
- (ii) for “Suzie”, “Salesman”, “Hall” and “Highway” from 0.5570 to 0.8394;
- (iii) for “Harbour”, “Coastguard” and “Foreman” from 0.9038 to 1.7272.

Based on this, henceforth we use the notions low, medium and high for videos from these three groups, respectively.

As we can notice from Fig. 1(a), α_k values tend to be low for low motion video sequences (e.g. for “Akiyo”), and as the amount of motion increases, they get higher (e.g. for “Salesman”), too. Another important observation is that the α_k values of high motion sequences show a strong increasing tendency as the source coding rate increases. As far as parameter β_k is concerned, we observe from Fig. 1(b) that the β_k values are moving increasingly within a narrow range as the source coding rate increases for all video sequences. Considering all of the above, we conclude that the values of parameter α_k can be used for the relative quantification of the motion level since there is a clear distinction of the values’ ranges for low, medium or high amount of motion.

III. PROBLEM FORMULATION AND PROPOSED APPROACHES

Under the consideration of a constant spreading code length, L , and the constraint of identical chip rate, $R_{chip,k}$, for all network nodes, the transmission bit rate R_k is correspondingly constant within the network, as derived by Eq. (1) and is henceforth denoted as R . Furthermore, the constant bit rate R induces that the ratio $R_{s,k}/R_{c,k}$ from Eq. (3), is identical for each node. This means that the selection of the source and channel coding rate $(R_{s,k}, R_{c,k})$ is joint, and we denote this coding set selection with CS_k for each node k .

Under these assumptions, the present paper copes with the problem of optimally allocating to each node k , the source coding and channel coding rate, CS_k , as well as the transmission power level S_k , such that a function $\mathcal{F}(\cdot)$ of the overall end-to-end expected distortions of each node, $\mathcal{F}(E[D_{s+c,1}], \dots, E[D_{s+c,K}])$, is minimized. Putting it formally, if we define the vectors:

$$CS = (CS_1, CS_2, \dots, CS_K)^T \quad S = (S_1, S_2, \dots, S_K)^T,$$

then the problem can be defined as follows:

$$\min_{CS, S} \mathcal{F}(E[D_{s+c,1}], E[D_{s+c,2}], \dots, E[D_{s+c,K}]), \quad (9)$$

subject to:

$$CS \in \{CS^1, CS^2, \dots, CS^M\}, \quad M \in \mathbb{N}^*, \quad (10)$$

$$\frac{R_s^1}{R_c^1} = \frac{R_s^2}{R_c^2} = \dots = \frac{R_s^M}{R_c^M} = R, \quad (11)$$

$$S_{\min} \leq S \leq S_{\max}, \quad (12)$$

where M is the number of the available source and channel coding rates.

As stated in the above problem formulation, the source coding rate and the channel coding rate assume values from discrete sets of the same cardinality, M , while the transmission power assumes values from a continuous range. It can be easily shown that combining Eqs. (3), (4), (5), (6), (7) and (8), the expected distortion $E[D_{s+c,k}]$ for node k can be written as a function of the source coding and channel coding rates, $R_{s,k}$ and $R_{c,k}$, and the transmission powers of all nodes, S . Hence, we can write the expected distortion of node k as $E[D_{s+c,k}](R_{s,k}, R_{c,k}, S)$ or $E[D_{s+c,k}](CS_k, S)$ [16]. This distortion-related function depends on the deployed criteria, which are described in the following sections.

A. Proposed Approaches

The resource allocation problem was considered under the following two different node-grouping schemes. The motivation was to explore the end-to-end video quality versus the computational efficiency of these schemes.

- (a) *Independent WVSAN nodes*: In this approach, each node acts as individual with its own content- and time-varying video and transmission parameters. The network resources are assigned to each node separately and the respective delivered video quality is estimated. Figure 2 depicts an example topology, while Fig. 3 illustrates a block diagram of the WVSAN considered. Each one of the

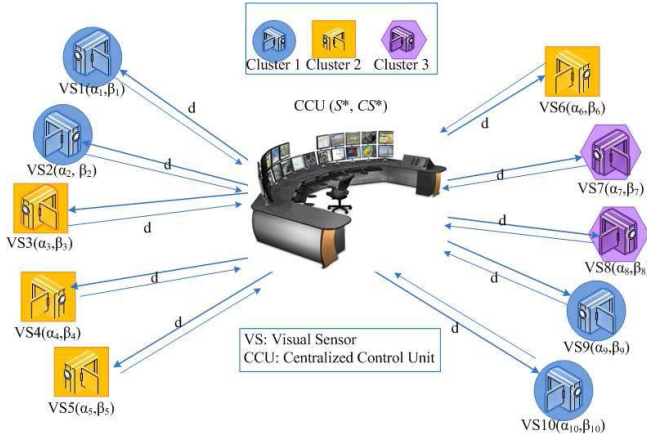


Fig. 2. Example of a WVSN topology with 10 nodes.

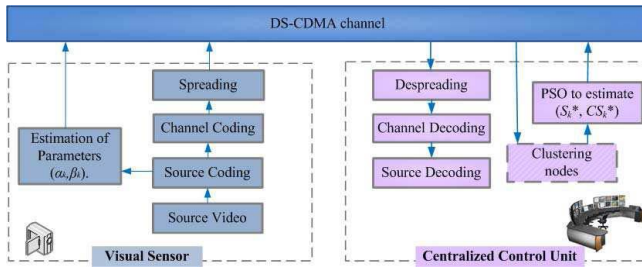


Fig. 3. Block Diagram of the WVSN system model considered.

visual sensors records a different scene and transmits its individual α_k and β_k parameters to the CCU. We assume that the sensor positions are static and that the CCU is aware of the exact WVSN topology. The CCU estimates the energy-per-bit to MAI ratio for each reference node k (using Eq. (4)). Based on the employed fairness criterion, the CCU computes the optimal resource allocation and transmits the corresponding values of the parameters, (CS_k^*, S_k^*) , as feedback to each node k .

- (b) *Clustered WVSN nodes*: This approach is motivated by the need for effectively reducing the computational complexity. We propose the use of clustering according to the individual video content-related parameters. As a result, each node becomes a member of a cluster represented by its centroid. The transmission parameters are allocated according to the centroids of the clusters. Based on those, we estimate the expected video quality of each node. The idea of clustering according to the motion level of the recorded scenes is depicted in Fig. 2 by considering the different colors and shapes of the nodes. In the same vein with the previous approach, each visual sensor transmits its individual (α_k, β_k) parameters to the CCU. The CCU performs clustering according to those parameters as shown in Fig. 3 and allocates the network resources based on the cluster centroids $(\alpha_{\text{cntrd},ct}, \beta_{\text{cntrd},ct})$ (where $ct \in \{1, 2, \dots, C\}$, where C denotes the number of clusters). Then, each node depending on the cluster that it belongs to, receives from the CCU the optimal cluster resources $(CS_{\text{cntrd},ct}^*, S_{\text{cntrd},ct}^*)$.

B. Clustering Algorithm

The general purpose of clustering is to form groups of similar objects. Several partitional clustering algorithms could be employed for the WVSN node clustering. In the present work we employ the k -means algorithm, which is a prototype-based clustering technique that simply divides the objects into non-overlapping clusters, i.e., each object belongs to only one cluster [25]. An additional advantage of k -means is its linear complexity in all relevant factors, namely the number of iterations, I_C , the number of clusters, C , and the dimensionality, n , of the search space. In the present work, we chose to cluster the WVSN nodes according to the motion level of the transmitting video sequences. Parameter α is a salient indication of the motion level. However, the formula for the expected distortion (Eq. (8)) requires both parameters α and β for each cluster. Thus, besides α , we also need a representative β for each cluster. Each formed cluster is represented in the resource allocation by its centroid's $(\alpha_{\text{cntrd},ct}, \beta_{\text{cntrd},ct})$ values. Obviously, this representation results in a significant reduction of the number of optimization parameters. Thus, the problem's complexity is improved, rendering our approach especially useful for real-time applications.

C. Optimization Criteria

According to the aforementioned problem definition, for both approaches we need to employ distortion-related functions in Eq. (9). We considered four such optimization criteria. The first one (see Section III-C.1) is defined as the minimization of the average end-to-end distortion among all nodes. The second criterion (see Section III-C.2) focuses on the minimization of the maximum distortion of the network. The third and fourth criteria (see Section III-C.3) lie in the field of Game Theory and they perform a bargaining game among the nodes. The difference among the two criteria is the distinct definition of the bargaining powers.

1) *Minimizing the Average Distortion (MAD)*: The MAD criterion minimizes the average end-to-end video distortion over all WVSN nodes by optimally determining the vectors of source coding and channel coding rates, CS^* , and transmission powers, S^* , i.e.,

$$(CS^*, S^*) = \arg \min_{CS, S} \frac{1}{K} \sum_{k=1}^K E[D_{s+c,k}], \quad (13)$$

subject to the constraints of Eqs. (10)–(12). Obviously, this criterion emphasizes the average performance of the network, allowing some individual nodes to assume higher distortion values than the rest.

2) *Minimizing the Maximum Distortion (MMD)*: The MMD criterion minimizes the maximum end-to-end distortion over all WVSN nodes by optimally determining the vectors of source coding and channel coding rates, CS^* , and transmission powers, S^* , i.e.,

$$(CS^*, S^*) = \arg \min_{CS, S} \max_k E[D_{s+c,k}], \quad (14)$$

subject to the constraints of Eqs. (10)–(12). In the system considered, reducing the distortion of a node by increasing

its transmission power causes more interference to the other nodes and thus increases their distortion. Thus, if we want to minimize the worst distortion among the nodes, we will increase the transmission power of the worst node. This will increase the distortion of the rest of the nodes. The MMD solution occurs when an “equilibrium” is reached, i.e. all nodes have the same distortion, and, increasing a node’s transmission power will increase the distortions of the other nodes, thus leading to a higher maximum distortion.

3) *Nash Bargaining Solution (NBS)*: As previously mentioned, the WVSAN node requirements to achieve higher video end-to-end quality, are conflicting. If the nodes behaved selfishly, they would try to maximize their video quality regardless of the choices made by the other nodes. This would result in all nodes transmitting using the highest available transmission power, leading to excessive interference, and thus in quality degradation (as already explained in the previous criterion). Therefore, we use the game-theoretic Nash Bargaining Solution (NBS) to organize a bargaining game. This game refers to a conflict of interest situation in which the nodes agree to negotiate in order to conclude a mutually beneficial agreement. The nodes will either reach an agreement or will fail [26]. It is sensible that a node will join the game only if it is ensured that its utility will be at least as much as it would be without cooperation (this utility is called the *disagreement point* and is denoted as dp).

For each node k from the set $\mathbf{K} = \{1, 2, \dots, K\}$, we define a *utility function* U_k , which reflects the benefit for this node. In the present work, the utility function reflects the video quality of the received video, i.e. the *Peak Signal to Noise Ratio* (PSNR):

$$U_k = \text{PSNR}_k = 10 \log_{10} \frac{255^2}{E[D_{s+c,k}]}, \quad (15)$$

where $E[D_{s+c,k}]$ is the expected video distortion for node k . \mathbf{U} denotes the feasible set of all possible utility allocations $U = (U_1, U_2, \dots, U_K)^\top$.

We define the NBS as a function $f : \langle \mathbf{U}, dp \rangle \rightarrow \mathbb{R}^K$ (with $dp \in \mathbf{U}$), that assigns to every bargaining problem $\langle \mathbf{U}, dp \rangle$ a unique member of \mathbf{U} . The NBS $f(\mathbf{U}, dp)$ satisfies the following axioms [26]:

- (I) *Feasibility*: $f(\mathbf{U}, dp) \geq dp$. This axiom ensures that each node will gain a profit at least as high as its disagreement point.
- (II) *Pareto Efficiency*: If there is $y > f(\mathbf{U}, dp)$, then $y \notin \mathbf{U}$. This axiom requires that there is no other solution within the feasible set that gives a greater utility to at least one node while giving the same utility to the other nodes [27].
- (III) *Invariance to Equivalent Utility Representations*: Given any strictly increasing affine transformation $\tau(\cdot)$, it holds that $f(\tau(\mathbf{U}), \tau(dp)) = \tau(f(\mathbf{U}, dp))$. This implies that the final outcome should not depend on the way the nodes calibrate their utility scales.
- (IV) *Independence of Irrelevant Alternatives*: If $dp \in \mathbf{V} \subseteq \mathbf{U}$, then $f(\mathbf{U}, dp) \in \mathbf{V}$ implies that $f(\mathbf{V}, dp) = f(\mathbf{U}, dp)$. This means that the NBS should be

independent of the availability or unavailability of irrelevant alternatives.

In order to find the solution $f(\mathbf{U}, dp)$, we first define the *Nash Product* (NP):

$$\text{NP} = (U_1 - dp_1)^{bp_1} (U_2 - dp_2)^{bp_2} \dots (U_K - dp_K)^{bp_K}, \quad (16)$$

where bp_k is the bargaining power of node k . The bargaining powers show which nodes are favored by the rules of the bargaining game.

Particularly, given a total transmission bit rate R_k , we determine U such that the NP is maximized. Mathematically, this can be expressed as follows:

$$f(\mathbf{U}, dp) = \arg \max_{\mathbf{U}} \text{NP}, \quad (17)$$

subject to the constraints of Eqs. (10)–(12). Since the utility function defined in Eq. (15) depends on the expected end-to-end distortion $E[D_{s+c,k}]$, it also depends on the source coding rate $R_{s,k}$, the channel coding rate $R_{c,k}$ and the vector of transmitted powers of all nodes S . This implies that Eq. (17) can be alternatively written as

$$(CS^*, S^*) = \arg \max_{CS, S} \text{NP}. \quad (18)$$

In the present work, we assume that $dp \in \mathbf{U}$ is the minimum acceptable quality (in terms of PSNR) and is determined by the system operator. The bargaining powers are assigned according to the rules of the bargaining game and show which player (node) is favored. Based on the different definition of the nodes’ bargaining powers, we formulate two distinct NBS-based criteria:

- (1) *e.NBS Criterion*: We assume that there is no necessity for favoring some of the nodes against the others. Thus, all nodes are equally treated. The bargaining powers are equal for all K nodes and we define that:

$$bp_k = \frac{1}{K}, \quad k \in \mathbf{K}. \quad (19)$$

This criterion is a special case of the NBS in which the symmetry axiom is valid. The axiom explains in simple words that, in the case where the players’ labels are swapped, each one will still get the same payoff as it would originally get [27].

- (2) *w.NBS Criterion*: According to this criterion, we propose the assignment of different bargaining powers to the nodes that are in accordance with the amount of motion recorded in the video scenes. As explained earlier in Section II-C, a salient metric for the level of motion in a video sequence is parameter α_k from the deployed rate-distortion model of Eq. (8). The higher the motion level in a video sequence, the higher the value of parameter α_k , and vice versa. Thus, for the w.NBS criterion, we define the bargaining power of each node k as the fraction:

$$bp_k = \frac{\alpha_k}{\sum_{k=1}^K \alpha_k}, \quad k \in \mathbf{K}. \quad (20)$$

This implies that the higher the motion level of a node, the higher its bargaining power, as well.

IV. EMPLOYED OPTIMIZATION ALGORITHMS

The *Particle Swarm Optimization* (PSO) algorithm served as our optimizer in all cases. PSO belongs to the category of *Swarm Intelligence* methods, which draw inspiration from the collective intelligence that emerges in physical systems of living organisms [28]. Since its development, PSO has gained immense popularity due to its simplicity and efficiency in a wide range of applications [29]. PSO has also been used for resource allocation in WWSNs [7], [14], [15] and for optimized video transmission over erasure channels [12]. Also, the use of PSO does not require convexity of the problem or the constructions of hierarchies of (mixed-integer) convex relaxations and the approximation of the problem's nonlinearities with convex under- and concave overestimators (i.e., the convex and concave envelopes of the objective function). This can be particularly useful in practical situations. These features as well as its robustness in solving dynamic problems [29] and its recognized tolerance in uncertain environments, motivated its use in our problem.

If the general n -dimensional global optimization problem $\min_{x \in X \subset \mathbb{R}^n} \mathcal{F}(x)$ is under investigation, PSO uses a set of search points, called the *swarm*, $S = \{x_1, x_2, \dots, x_N\}$, $x_i \in X$, $i \in I = \{1, 2, \dots, N\}$, to probe the search space. Each search point in S is called a *particle* $x_i = (x_{i1}, x_{i2}, \dots, x_{in})^\top \in X$, $i \in I$. Each particle is randomly initialized in X and iteratively probes it. Its motion is based on stochastically determined position shifts, called *velocity*, $v_i = (v_{i1}, v_{i2}, \dots, v_{in})^\top$, $i \in I$, while it retains in memory the *best position* $p_i = (p_{i1}, p_{i2}, \dots, p_{in})^\top$, $i \in I$, it has ever visited. The best positions serve as the guiding mechanism that attracts the particles towards the most promising regions of the search space, i.e., regions that possess lower function values.

Naturally, the rapid convergence of the particles to the best points detected so far in the search space fosters the danger of getting trapped in local minimizers. This premature convergence effect has been alleviated by introducing the concept of *neighborhood* to the particles [30], [31]. Putting it formally, a neighborhood of the i -th particle is defined as a set $NB_{i,s} = \{j_1, j_2, \dots, j_s\}$, $j_k \in I$, $k = 1, 2, \dots, s$, $i \in NB_{i,s}$, which consists of the indices of all the particles with which it can exchange information. Then, the neighborhood's best position is given by the following equation:

$$p_{gi} = \arg \min_{j \in NB_{i,s}} \{\mathcal{F}(p_j)\}. \quad (21)$$

The neighborhood's best position is used to update the i -th particle's velocity at each iteration. The parameter s (with $1 \leq s \leq N$) defines the number of particles that constitute the neighborhood.

Based on the definitions above, the iterative scheme of PSO is defined as follows [32]:

$$v_{ij}(t+1) = \chi \left[v_{ij}(t) + c_1 \mathcal{R}_1 (p_{ij}(t) - x_{ij}(t)) + c_2 \mathcal{R}_2 (p_{gi,j}(t) - x_{ij}(t)) \right], \quad (22)$$

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1), \quad (23)$$

Algorithm 2 Pseudocode of PSO-PS

Require: Previous and current parameters $\alpha_{k,t-1}, \alpha_{k,t}$ and $\beta_{k,t-1}, \beta_{k,t}$ for each node k , previous resource allocation $(CS^*, S^*)_{t-1}$, parameter α_k variation threshold δ_α .

```

1: loop
2:   if any  $|\alpha_{k,t-1} - \alpha_{k,t}| > \delta_\alpha$  then
3:     Run the clustering algorithm to form the new clusters.
     /* Only for Clustered WWSN nodes */
4:     Initialize half of the swarm around the previous
     allocation  $(CS^*, S^*)_{t-1}$ .
5:     Initialize the other half of the swarm in random
     positions within the predefined ranges.
6:     Run the PSO update function.
7:     Send to the nodes the updated resource allocation
      $(CS^*, S^*)_t$ .
8:   end if
9: end loop

```

where, $i = 1, 2, \dots, N$; $j = 1, 2, \dots, n$; the parameter χ is the *constriction coefficient*; acceleration constants c_1 and c_2 are called the *cognitive* and *social* parameter, respectively; and $\mathcal{R}_1, \mathcal{R}_2$, are random variables uniformly distributed in the range $[0, 1]$. It shall be noted that a different value of \mathcal{R}_1 and \mathcal{R}_2 is sampled for each i and j in Eq. (22). Also, the best position of each particle is updated at each iteration, as follows:

$$p_i(t+1) = \begin{cases} x_i(t+1), & \text{if } \mathcal{F}(x_i(t+1)) < \mathcal{F}(p_i(t)), \\ p_i(t), & \text{otherwise} \end{cases}, \quad i \in I. \quad (24)$$

This PSO variant was introduced in [32], and based on its stability analysis, the parameter set $\chi = 0.729, c_1 = c_2 = 2.05$, was determined as a satisfactory setting that produces a balanced convergence speed of the algorithm. Nevertheless, alternative settings have been introduced in relevant work [33].

A. PSO-Based Algorithms for Efficient Update

As previously mentioned, the problems that are studied in the present paper require the application of the optimization procedure each time a change in the recorded amount of motion occurs, due to the consequent need for resource reallocation. Since optimization from scratch can be laborious and time consuming, we proposed two techniques which effectively reduce the time consumption. Both techniques are inspired by the application of PSO in dynamic optimization environments and differ in the swarm initialization. The first technique requires the information of the prior WWSN resource allocation for the swarm initialization, while the second uses the information of the recorded amount of motion per sensor. Exploiting prior information is a common practice, especially in cases where changes of the environment are not expected to radically modify the objective function.

1) *PSO With Prior Resource Allocation (PSO-PS)*: According to this technique, the previous resource allocation is used to initialize the swarm, as reported in Algorithm 2. PSO-PS requires that PSO has run at least for the first time to allocate

Algorithm 3 Pseudocode of PSO-RE

Require: Previous and current parameters $\alpha_{k,t-1}, \alpha_{k,t}$ and $\beta_{k,t-1}, \beta_{k,t}$ for each node k , and parameter α_k variation threshold δ_α .

- 1: **loop**
- 2: **if** any $|\alpha_{k,t-1} - \alpha_{k,t}| > \delta_\alpha$ **then**
- 3: Run the clustering algorithm to form the new clusters.
 / Only for Clustered WVSNS nodes */*
- 4: Make a rough estimation of the transmission power based on Eq. (25).
- 5: Initialize half of the swarm at this rough estimation \hat{S} .
- 6: Initialize the other half of the swarm in random positions within the predefined ranges.
- 7: Run the PSO update function.
- 8: Send to the nodes the updated resource allocation $(CS^*, S^*)_t$.
- 9: **end if**
- 10: **end loop**

the WVSNS resources based on the initial motion levels. The proposed updating mechanism is triggered whenever the change of the currently recorded motion level is higher than a user-defined threshold, δ_α , from the previous one. This threshold highly depends on the number of sensors in the network. In the case of the Independent Nodes approach, Step 3 of Algorithm 2 is omitted.

2) *PSO With Swarm Initialization Based on a Rough Estimation (PSO-RE)*: A close inspection of the resource allocation results from our previous work [14], [15], that used the same optimization criteria, revealed that the transmission power of each node k can be associated with the amount of motion of the recorded scenes. Particularly, the higher the motion level of a node, the higher its transmitting power. Based on this observation, we can make a rough estimation of the transmission power of the node-of-interest k , according to its relevant motion level, as follows:

$$\hat{S}_{k,t} = \frac{\alpha_{k,t}}{\min(\alpha_t)} S_{\min}, \quad (25)$$

where α_t is the vector of the $\alpha_{k,t}$ values for each node k , $\alpha_t = (\alpha_1, \alpha_2, \dots, \alpha_K)^\top$ and $\min(\alpha_t)$ is the minimum element of vector α . In order to comply with the problem formulation, $\hat{S}_{k,t}$ should lie in the transmission power range, i.e. $S_{\min} \leq \hat{S}_{k,t} \leq S_{\max}$, where S_{\min} and S_{\max} are the lowest and highest valid transmission powers, respectively. Then, for half of the particles of the swarm, we initialize the components that correspond to the transmission powers, at the estimated value, $\hat{S} = (\hat{S}_1, \hat{S}_2, \dots, \hat{S}_K)^\top$, as described in Algorithm 3. We do not use rough estimation for the source and channel coding rates, due to the fact that these parameters lie in a rather limited range, i.e., they can assume only three possible values (see Table II). Opposed to PSO-PS, PSO-RE can be utilized even for the first resource allocation. Similarly to PSO-PS, the updating mechanism of PSO-RE is triggered whenever the change of the currently recorded motion level from the previous is higher than a user-defined threshold, δ_α .

TABLE II
NOTATION AND PARAMETER VALUES

Description	Notation	Value(s)
Number of WVSNS nodes	K	10
Total Bandwidth	W_t	2.5 MHz
Total Transmission Rate	R_k	96 kbps
Valid Coding Sets	CS	{1:(32 kbps, 1/3), 2:(48 kbps, 1/2), 3:(64 kbps, 2/3)}
Transmission Power	S_k	[0.050, 0.500] W
Video Sequence Format		QCIF
Frame Rate		15 fps
RCPC Mother Code Rate		1/4
Disagreement Point	dp	24dB (same for all nodes and test cases)
Transmitter Antenna Height	h_t	3.0 m
Receiver Antenna Height	h_r	3.0 m
Propagation Distance	d	120 m
Transmitter Antenna Gain	G_t	2.5 dB
Receiver Antenna Gain	G_r	2.5 dB
PSO Swarm Size	SS	TC1: 150
	TC2	$SS_{\text{MAD}} = SS_{\text{wNBS}} = 30, SS_{\text{MMD}} = 50, SS_{\text{eNBS}} = 40$
PSO Maximum # of Iterations (for the initial motion levels)	I_{\max}	TC1: $I_{\max}^{\text{MAD}} = 870, I_{\max}^{\text{MMD}} = 1000, I_{\max}^{\text{eNBS}} = I_{\max}^{\text{wNBS}} = 700$ TC2: $I_{\max} = 100$ for all criteria

V. EXPERIMENTAL RESULTS

For the evaluation of our proposed approaches, a number of test cases were considered. The test cases were selected such that real situations are resembled, where each sensor node may record a scene of different motion. We considered a single-hop WVSNS topology, similar to the one depicted in Fig. 2, where all nodes are equidistant from the CCU and have the same antenna gain and height for the transmitter and the receiver, respectively.

In our formulation, the source and channel coding rates for each node are selected from a prespecified discrete set of combinations (R_s, R_c). As we can see in Table II, these combinations are denoted with the labels 1, 2 and 3, respectively. Thus, PSO has to decide between 1, 2, and 3, regarding the combination of (R_s, R_c). However, PSO was originally defined as a real-valued optimization algorithm. In relevant work [7], [14], [15], integer problems are typically tackled by rounding the variables' (real) values to the nearest integers, without modifying PSO's dynamic. Thus, in our case, we let the corresponding PSO's coordinate lie within the real interval [0.6, 3.4] and, by rounding to the nearest integer, we obtain either 1, 2, or 3. The default PSO variant reported in Section IV is used. Since PSO is a stochastic algorithm, for each problem instance we conducted 30 independent experiments in order to ensure validity of the results. The swarm size SS and the maximum number of PSO iterations I_{\max} were dependent on the test case considered. I_{\max} is the maximum number of iterations among the 30 experiments. During each experiment, the best detected solution was recorded. Our experiments confirmed that the optimal power allocation is not unique, but there is a set of optimal transmission powers S^* that satisfy the proposed optimization criteria. To explain this, we have to refer to Eq. (4). If all received powers are multiplied by the same constant, then E_k/I_0 remains the same. This leads to a set of optimal solutions. Hence, we present the results with the lowest values for transmission power for all criteria and both test cases.

A. Testing of the two Proposed Approaches and Results

The first test case (denoted as TC1) consisted of a WVSNS with 10 nodes admitting different fields of view. We assumed that the fundamental difference between the recorded scenes

TABLE III

EXPERIMENTAL RESULTS FOR ALL CRITERIA FOR TC1 FOR THE INITIAL MOTION LEVELS, NAMELY RESULTING VIDEO QUALITY PSNR (dB), ALLOCATED TRANSMISSION POWER S (W), SOURCE AND CHANNEL CODING RATES (R_s (kbps), R_c)

Nodes	TC1 Independent Nodes											
	MAD			MMD			e.NBS			w.NBS		
	PSNR	S	(R_s, R_c)	PSNR	S	(R_s, R_c)	PSNR	S	(R_s, R_c)	PSNR	S	(R_s, R_c)
"Akiyo"	36.6234	0.0062	(64,2/3)	33.8121	0.0055	(32,1/3)	37.9888	0.0066	(64, 2/3)	29.4613	0.0055	(32,1/3)
"Grandma"	38.0863	0.0050	(32,1/3)	33.8121	0.0050	(32,1/3)	35.5382	0.0050	(32, 1/3)	29.9293	0.0050	(32,1/3)
"Mother-Daughter"	35.3274	0.0053	(32,1/3)	33.8121	0.0053	(32,1/3)	35.9640	0.0054	(32, 1/3)	30.5080	0.0057	(32,1/3)
"Highway"	34.6034	0.0066	(32,1/3)	33.8121	0.0071	(32,1/3)	34.8619	0.0064	(32, 1/3)	32.0027	0.0089	(32,1/3)
"Salesman"	34.0969	0.0094	(64,2/3)	33.8121	0.0108	(64,2/3)	34.0756	0.0088	(64, 2/3)	30.1598	0.0098	(32,1/3)
"Suzie"	34.8449	0.0064	(32,1/3)	33.8121	0.0068	(32,1/3)	35.2492	0.0063	(32, 1/3)	31.9931	0.0085	(32,1/3)
"Hall"	34.8972	0.0065	(32,1/3)	33.8121	0.0068	(32,1/3)	35.3416	0.0064	(32, 1/3)	32.2888	0.0088	(32,1/3)
"Foreman"	32.6868	0.0111	(64,2/3)	33.8121	0.0148	(64,2/3)	32.0829	0.0098	(64, 2/3)	35.5397	0.0272	(64,2/3)
"Harbour"	31.5727	0.0108	(48,1/2)	33.8121	0.0172	(64,2/3)	30.6105	0.0089	(48, 1/2)	31.5017	0.0191	(48,1/2)
"Coastguard"	33.5929	0.0110	(64,2/3)	33.8121	0.0133	(64,2/3)	33.3102	0.0101	(64, 2/3)	42.1464	0.0414	(64,2/3)
Total Power		0.0783			0.0924			0.0735			0.1398	

was the amount of motion in each one of them. Therefore, we used the first 150 frames of 10 well-known YUV sequences of various amounts of motion and computed the (α_k, β_k) parameters for each node. The values of (α_k, β_k) for the three available source coding rates and for all video sequences are depicted in Fig. 1. Regarding the complexity of the (α_k, β_k) computation, we need to note that the complexity of the ROPE algorithm is not high and can therefore run for real-time applications [23], [34]. The complexity of Algorithm 1 is not high as well, since it is executed for a small number of different P_{be} . Beyond this, the Least Squares Regression complexity depends on the number of training examples and features, which in our case is rather low.

In the second test case (denoted as TC2), we performed k -means clustering of the 10 nodes of TC1 using the parameters (α_k, β_k) as data set, and allocated equally the network resources within the same cluster. We have conducted experiments with different values for the number of clusters, namely $C = 2, \dots, 5$. For each one of the clusters, three centroids (pairs of $(\alpha_{\text{cntrd},ct}, \beta_{\text{cntrd},ct})$ values) were produced with regard to the three admissible source coding rates. After determining at the CCU the clusters' transmission power level and the source and channel coding rates based on the centroids, the real PSNR values were calculated for each node of the cluster based on each node's individual (α_k, β_k) values.

Tables III and IV report the allocated resources and the resulting video quality in terms of PSNR for all four criteria in both test cases. As we observe, from the video quality point of view for both test cases, MAD and e.NBS favored the low motion sequences, while w.NBS gave priority to the higher motion levels. MMD achieved the same quality level for all nodes for TC1.

In order to compare the resulting video quality when using a different number of clusters C , we computed the mean absolute PSNR difference, $|\Delta PSNR|$ (with $|\Delta PSNR_k| = |PSNR_{TC1,k} - PSNR_{TC2,C,k}|$ the absolute PSNR difference per node k) and the standard error $\sigma(|\Delta PSNR|)$. These values are illustrated in Fig. 4. We needed the mean value of the absolute differences, due to the fact that for some nodes the PSNR difference is positive, while for others negative. This is due to the distance of each node's (α_k, β_k) values from its cluster's centroid, $(\alpha_{\text{cntrd},ct}, \beta_{\text{cntrd},ct})$. Since (α_k, β_k) values

are important for the resource allocation process, the distance of those values from the centroids strongly affects the allocated resources and the video quality. This becomes more evident in the results for w.NBS, in which the α_k values are used to favor the high motion nodes. This criterion has the highest $|\Delta PSNR|$ and $\sigma(|\Delta PSNR|)$ compared to all the other criteria. On the other hand, for the other optimization criteria, using three or four clusters results in $|\Delta PSNR| < 1$ dB. Another important observation, that is also related to the (α_k, β_k) values distance from its cluster's centroid $(\alpha_{\text{cntrd},ct}, \beta_{\text{cntrd},ct})$, is that a different number of clusters results in the lowest $|\Delta PSNR|$ for this experimental setup. Particularly, from the quality point of view the optimal number of clusters for MAD and w.NBS is three, while for e.NBS is four and for MMD five. Please, also note, that the results reported in Table IV refer to these numbers of clusters per criterion.

From the transmission power allocation point of view, in both test cases and with every criterion, the allocated transmission powers are in accordance with the motion level of the transmitted video sequences. Namely, the transmission powers for the nodes with high motion are higher than the transmission powers of the low motion nodes. Moreover, the following inequality is valid for both test cases: $\sum_{k=1}^K S_{e.NBS,k}^* < \sum_{k=1}^K S_{MAD,k}^* < \sum_{k=1}^K S_{MMD,k}^* < \sum_{k=1}^K S_{w.NBS,k}^*$. It is important to point out that e.NBS results in the lowest total transmission power, while the highest total transmission power results from w.NBS and this is considered as the tradeoff for achieving higher quality for the nodes of high motion. Moreover, although more or less transmission power per node is allocated in some cases, the sum of the allocated transmission power in the two test cases, as given by the total values at the bottom of Tables III and IV, is on average around 6.33% higher than the total transmission power allocated in TC2 for all optimization criteria. This implies that, besides the higher computational cost, TC1 also induces a higher energy cost compared to TC2.

In Table V we give an example of the execution times of the two proposed approaches for each experiment performed in an Intel Core Quad CPU @2.33GHz with 2GB RAM. As far as the quality/complexity tradeoff of the different approaches is concerned, we shall consider several issues. Firstly, taking into account that the number of the particles and number of iterations both depend on the number of the WWSN nodes,

TABLE IV

EXPERIMENTAL RESULTS FOR ALL CRITERIA FOR TC2 FOR THE INITIAL MOTION LEVELS, NAMELY RESULTING VIDEO QUALITY PSNR (dB), ALLOCATED TRANSMISSION POWER S (W), SOURCE AND CHANNEL CODING RATES ($R_s(kbps)$, R_c)

Nodes	TC2 Clustered Nodes											
	MAD with $C = 3$			MMD with $C = 5$			e.NBS with $C = 4$			w.NBS with $C = 3$		
	PSNR	S	(R_s, R_c)	PSNR	S	(R_s, R_c)	PSNR	S	(R_s, R_c)	PSNR	S	(R_s, R_c)
"Akiyo"	35.1377	0.0050	(32, 1/3)	33.4973	0.0050	(32, 1/3)	38.1085	0.0063	(64, 2/3)	29.7106	0.0050	(32, 1/3)
"Grandma"	35.7730	0.0050	(32, 1/3)	34.2068	0.0050	(32, 1/3)	35.9387	0.0050	(32, 1/3)	30.5917	0.0050	(32, 1/3)
"Mother-Daughter"	35.6163	0.0050	(32, 1/3)	33.8135	0.0050	(32, 1/3)	35.8071	0.0050	(32, 1/3)	29.6520	0.0050	(32, 1/3)
"Highway"	34.5432	0.0060	(32, 1/3)	35.0769	0.0078	(32, 1/3)	34.9931	0.0061	(32, 1/3)	30.8157	0.0074	(32, 1/3)
"Salesman"	30.5164	0.0060	(32, 1/3)	30.8589	0.0078	(32, 1/3)	30.8051	0.0061	(32, 1/3)	28.1244	0.0074	(32, 1/3)
"Suzie"	35.0654	0.0060	(32, 1/3)	35.6072	0.0078	(32, 1/3)	35.5221	0.0061	(32, 1/3)	31.2813	0.0074	(32, 1/3)
"Hall"	35.0105	0.0060	(32, 1/3)	33.8165	0.0065	(32, 1/3)	35.4802	0.0061	(32, 1/3)	31.1192	0.0074	(32, 1/3)
"Foreman"	33.3120	0.0108	(64, 2/3)	34.4701	0.0151	(64, 2/3)	32.4323	0.0096	(64, 2/3)	36.6813	0.0290	(64, 2/3)
"Harbour"	32.0850	0.0108	(64, 2/3)	33.1662	0.0151	(64, 2/3)	31.2637	0.0096	(64, 2/3)	35.2306	0.0290	(64, 2/3)
"Coastguard"	34.4099	0.0108	(64, 2/3)	33.8164	0.0126	(64, 2/3)	33.3340	0.0096	(64, 2/3)	38.5301	0.0290	(64, 2/3)
Total Power	0.0712			0.0879			0.0695			0.1316		

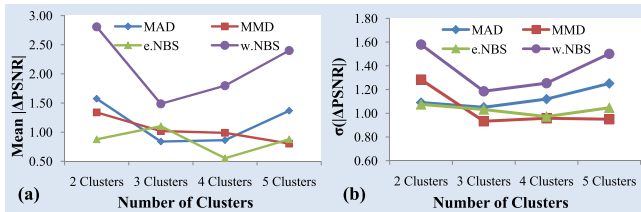


Fig. 4. Comparison of (a) the $|\Delta PSNR|$ and (b) $\sigma(|\Delta PSNR|)$ values per criterion and varying number of clusters.

the optimization algorithm complexity is directly related to the number of nodes included in the WVSAN considered. The worst case complexity for the first approach, following the definition in [35], is $O(SS \times I_{\max} \times n)$, while for the second $O(I_C \times C \times n) + O(SS \times I_{\max} \times n)$. Taking into consideration the SS and the I_{\max} from Table II for TC1 and TC2, it is evident that the worst case complexity of the second approach is lower than that of the first. This fact is verified by the execution times reported in Table V. On the other hand, in TC2 the number of WVSAN nodes K and the selected number of clusters C regulate the complexity. It is also clear that the complexity of the second approach will be equal to the first approach only if the number of clusters approximates the number of WVSAN nodes.

Taking into consideration the quality and power tradeoffs, it is suggested that the optimization criterion has to be decided each time with respect to the application requirements that may set specific restrictions on the delivered video quality. For example, if a specific application requires to favor certain nodes according to the recorded motion level, then it is evident that the w.NBS is recommended. If the system requires the best possible quality for the low motion video sequences, then we recommend the e.NBS. Finally, for the case that the system requirements demand similar quality levels, we recommend the MMD deployment.

Concerning the issue of which approach is preferable according to the delivered video quality for each node, it is clear that this exclusively depends on the application considered. Using clustering, we effectively reduce the dimensionality of the optimization problem, a fact that makes the system (computationally) less complicated, thus more time-efficient. Moreover, the WVSAN power consumption is reduced. However, it still fosters the danger of degradation of the received video quality.

TABLE V
EXECUTION TIMES PER EXPERIMENT IN ms

	$C = 2$	$C = 3$	$C = 4$	$C = 5$	TC1
MAD	0.009	0.019	0.035	0.075	2.750
MMD	0.009	0.019	0.076	0.190	3.170
e.NBS	0.009	0.020	0.037	0.090	2.420
w.NBS	0.009	0.021	0.053	0.230	2.460

B. Testing of the PSO-Based Algorithms in Dynamic WVSANs and Results

For the testing of both PSO-based algorithms, the case that the initial motion levels of the recorded videos change through time was considered. In order to assess the performance of the proposed mechanism, large motion variations per time instance were used. Namely, at each time instance one node's motion level changes from high to low or vice versa. For all criteria and test cases, the swarm size was the same as the one used in the initial resource allocation. On the other hand, the maximum number of iterations was dependent on the number of nodes with different motion level compared with the initial states. Therefore, we use the *maximum number of objective function evaluations*, i.e. the product of I_{\max} and SS . Our experimental setting consisted of different levels of the maximum number of function evaluations that were provided to the PSO approaches. Each approach was run until the maximum number of evaluations was exceeded, and its best solution was recorded. For each algorithm, our interest was focused on the corresponding level that allowed the algorithm to be successful in all the 30 independent experiments. This level can be considered as the expected number of function evaluations required by the algorithm to achieve convergence. These levels per algorithm and test case with regard to the percentage of nodes with different motion levels than the initial ones are reported in Figs. 5 and 6.

A close inspection of the graphs in Fig. 5 reveals that for TC1 the number of objective function evaluations has an increasing tendency as the percentage of nodes with different motion level from the initial levels increases. On the contrary, for TC2 the number of objective function evaluations using PSO-PS demonstrates a smoother increasing tendency, and in some cases remains the same or slightly decreases. It is

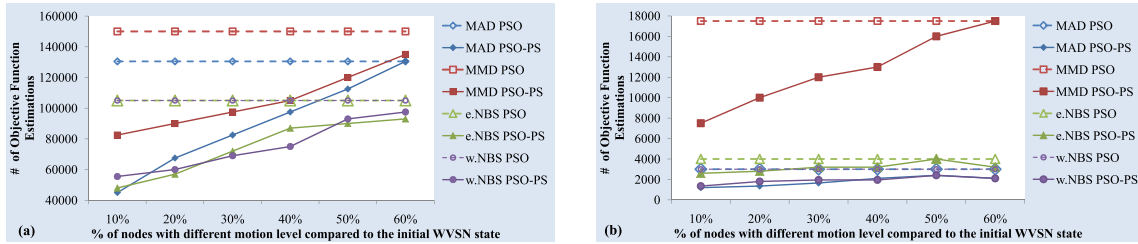


Fig. 5. PSO-PS maximum number of Objective Function Evaluations vs. % of nodes with different motion level for all criteria for (a) TC1 and (b) for TC2 (e.g. 20% means that 2 out of 10 nodes have different motion level compared to its initial motion level).

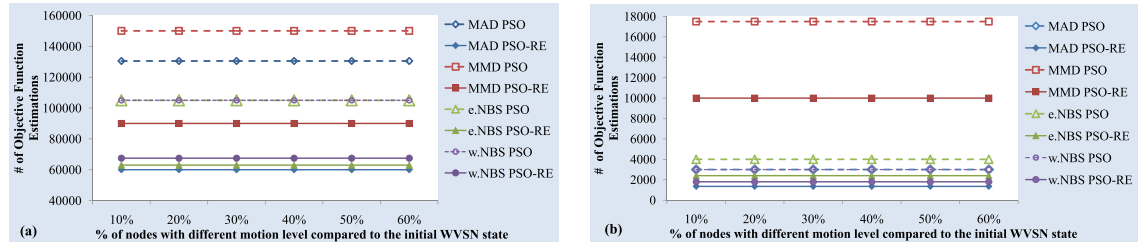


Fig. 6. PSO-RE maximum number of Objective Function Evaluations vs. % of nodes with different motion level for all criteria for (a) TC1 and (b) for TC2.

remarkable that using the approach of clustered nodes results in a much smaller number of algorithm iterations to reach convergence. For example, for the MAD criterion, in the case that 10% of the initial motion levels have changed, using clustering results in 97.33% fewer objective function evaluations. Moreover, the MMD criterion requires the highest number of maximum objective function evaluations in order to converge to the optimal solution compared to all other criteria in all cases.

As already explained, PSO-PS initializes the swarm around the prior solution, which was computed based on the initial α_k values, and does not result in a constant maximum number of objective function evaluations. Opposed to this, PSO-RE performs the rough estimation of the transmitted power based on the current time instance α_k values. Hence, it locates the swarm close to the optimal solution, and it converges after a small and constant maximum number of function evaluations every time there is a change of motion in the scenery for all the criteria and test cases, as observed in Fig. 6. Furthermore, similarly to PSO-PS, TC2 converges faster than TC1.

Regarding these two optimization algorithms, it is important to point out that both algorithms succeed in optimally allocating the WVSN resources quicker than the PSO algorithm did in this and our previous work [14], [15], as well. Additionally, as anticipated, the resulting optimal solutions from PSO-PS and PSO-RE are identical for the same motion allocation. Moreover, the PSO-RE requires far fewer iterations to converge for all optimization criteria and both test cases. Hence, it outperforms PSO and PSO-PS in terms of computation time.

VI. CONCLUSION

We have presented two approaches for the problem of quality-driven cross-layer optimization of DS-CDMA

WVSNs. We considered a single-hop topology, where each sensor transmits directly to the CCU, which handles the task of resource allocation. The two approaches have different quality/energy/complexity tradeoffs: the first considers independent visual sensors, while the second clusters them according to the recorded motion level. Overall, selecting the approach of clustering instead of having independent nodes offers the advantages of a time efficient resource allocation update and of a slightly reduced power consumption. However, a reasonable quality decline for some nodes might be the tradeoff. For the resource allocation, we considered four different optimization criteria: the MAD, the MMD, the e.NBS and the w.NBS. Using MAD results in an optimal average end-to-end quality. The MMD criterion achieves the same quality levels at the receiver. e.NBS and w.NBS apply the cooperative NBS by using equal and motion-based bargaining powers, respectively. The dynamic nature of the recorded scenery requires a constant allocation of resources. Thus, we introduced two PSO-based algorithms for the fast reallocation of the WVSN resources. The two algorithms differ in the swarm population initialization step; PSO-PS initializes half of the swarm based on the prior resource allocation, while PSO-RE performs a rough estimation of the resources based on their current relevant motion levels. Both algorithms offer enhanced time efficiency compared to PSO. However, PSO-RE ensures a constant and low maximum number of objective function evaluations, which is of great importance especially for real-time applications.

REFERENCES

- [1] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A survey on wireless multimedia sensor networks," *Comput. Netw. J.*, vol. 51, pp. 921–960, Mar. 2007.
- [2] S. Soro and W. Heinzelman, "A survey of visual sensor networks," *Adv. Multimedia*, vol. 2009, pp. 640386–1–640386–21, May 2009.
- [3] M. Tiwari, T. Groves, and P. C. Cosman, "Competitive equilibrium bitrate allocation for multiple video streams," *IEEE Trans. Image Process.*, vol. 19, no. 4, pp. 1009–1021, Apr. 2010.

- [4] H. P. Shiang and M. van der Schaar, "Information-constrained resource allocation in multicamera wireless surveillance networks," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 4, pp. 505–517, Apr. 2010.
- [5] H. Park and M. van der Schaar, "Bargaining strategies for networked multimedia resource management," *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3496–3511, Jul. 2007.
- [6] H. Park and M. van der Schaar, "Fairness strategies for wireless resource allocation among autonomous multimedia users," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 2, pp. 297–309, Feb. 2010.
- [7] K. Pandremmenou, L. P. Kondi, and K. E. Parsopoulos, "Optimal power allocation and joint source-channel coding for wireless DS-CDMA visual sensor networks using the Nash bargaining solution," in *Proc. Int. Conf. Acoust., Speech Signal Process.*, May 2011, pp. 2340–2343.
- [8] A. V. Katsenou, L. P. Kondi, K. E. Parsopoulos, and E. S. Bentley, "Quality-driven power control and resource allocation in wireless multi-rate visual sensor networks," in *Proc. 19th IEEE Int. Conf. Image Process.*, Orlando, FL, USA, Sep./Oct. 2012, pp. 1117–1120.
- [9] 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. 19th IEEE Int. Conf. Image Process.*, Orlando, FL, USA, Sep./Oct. 2012, pp. 1102–1104.
- [10] N. Thomos, S. Argyropoulos, N. V. Boulgouris, and M. G. Strintzis, "Robust transmission of H.264/AVC video using adaptive slice grouping and unequal error protection," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2006, pp. 593–596.
- [11] T.-L. Lin and P. C. Cosman, "Efficient optimal RCPC code rate allocation with packet discarding for pre-encoded compressed video," *IEEE Signal Process. Lett.*, vol. 17, no. 5, pp. 505–508, May 2010.
- [12] Y. Zhang, S. Qin, and Z. He, "Transmission distortion-optimized unequal loss protection for video transmission over packet erasure channels," in *Proc. IEEE ICME*, Jul. 2011, pp. 1–6.
- [13] S. Paluri, K. K. R. Kambhatla, S. Kumar, B. Bailey, P. Cosman, and J. Matyjas, "Predicting slice loss distortion in H.264/AVC video for low complexity data prioritization," in *Proc. 19th IEEE ICIP*, Sep./Oct. 2012, pp. 689–692.
- [14] 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 Int. Conf. Image Process.*, Brussels, Belgium, Sep. 2011, pp. 149–152.
- [15] A. V. Katsenou, L. P. Kondi, and K. E. Parsopoulos, "On the use of clustering for resource allocation in wireless visual sensor networks," *Proc. SPIE*, vol. 8305, pp. 83050S-1–83050S-13, Jan. 2012.
- [16] E. S. Bentley, L. P. Kondi, J. D. Matyjas, M. J. Medley, and B. W. Suter, "Spread spectrum visual sensor networks resource management using an end-to-end cross layer design," *IEEE Trans. Multimedia*, vol. 13, no. 1, pp. 125–131, Feb. 2011.
- [17] Y. S. Chan and J. W. Modestino, "A joint source coding-power control approach for video transmission over CDMA networks," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 10, pp. 1516–1525, Dec. 2003.
- [18] T. Rapaport, *Wireless Communications Principles and Practise* (Prentice Hall Communications Engineering and Emerging Technologies Series), 2nd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 2001.
- [19] T. Wiegand and G. Sullivan, "The H.264/AVC Video coding standard standards in a nutshell," *IEEE Signal Process. Mag.*, vol. 24, no. 2, pp. 148–153, Mar. 2007.
- [20] T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 657–673, Jul. 2003.
- [21] J. Hagenauer, "Rate-compatible punctured convolutional codes (RCPC codes) and their applications," *IEEE Trans. Commun.*, vol. 36, no. 4, pp. 389–400, Apr. 1988.
- [22] L. P. Kondi and A. K. Katsaggelos, "Joint source-channel coding for scalable video using models of rate distortion functions," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Salt Lake City, UT, USA, May 2001, pp. 1377–1380.
- [23] H. Yang and K. Rose, "Advances in recursive per-pixel end-to-end distortion estimation for robust video coding in H.264/AVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 7, pp. 845–856, Jul. 2007.
- [24] Y.-C. Lin and S.-C. Tai, "Fast full-search block-matching algorithm for motion-compensated video compression," *IEEE Trans. Commun.*, vol. 45, no. 5, pp. 527–530, May 1997.
- [25] P. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Reading, MA, USA: Addison-Wesley, 2006.
- [26] M. J. Osborn and A. Rubinstein, *Bargaining and Markets*. New York, NY, USA: Academic, 1990.
- [27] K. Binmore, *Playing for Real: A Text on Game Theory*. Oxford, U.K.: Oxford Univ. Press, 2007.
- [28] J. Kennedy and R. C. Eberhart, *Swarm Intelligence*. San Mateo, CA, USA: Morgan Kaufmann, 2001.
- [29] K. E. Parsopoulos and M. N. Vrahatis, *Particle Swarm Optimization and Intelligence: Advances and Applications*. Hershey, PA, USA: IGI Global, 2010.
- [30] J. Kennedy, "Small worlds and mega-minds: Effects of neighborhood topology on particle swarm performance," in *Proc. IEEE Congr. Evol. Comput.*, Jul. 1999, pp. 1931–1938.
- [31] P. N. Suganthan, "Particle swarm optimiser with neighbourhood operator," in *Proc. IEEE Congr. Evol. Comput.*, Washington, DC, USA, Jul. 1999, pp. 1958–1961.
- [32] M. Clerc and J. Kennedy, "The particle swarm-explosion, stability, and convergence in a multidimensional complex space," *IEEE Trans. Evol. Comput.*, vol. 6, no. 1, pp. 58–73, Feb. 2002.
- [33] I. C. Trelea, "The particle swarm optimization algorithm: Convergence analysis and parameter selection," *Inf. Process. Lett.*, vol. 85, pp. 317–325, Mar. 2003.
- [34] T. Stockhammer and S. Wenger, "Standard-compliant enhancement of JVT coded video for transmission over fixed and wireless IP," in *Proc. 4th Int. Workshop Distrib. Comput.*, Capri, Italy, Sep. 2002, pp. 1–9.
- [35] T. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed. Cambridge, MA, USA: MIT Press, 2009.

Angeliki V. Katsenou received the Diploma degree in electrical and computer engineering and the M.S. degree in signal and image processing from the University of Patras, Patras, Greece, in 2004 and 2007, respectively. She is currently pursuing the Ph.D. degree with the Department of Computer Science and Engineering, University of Ioannina, Greece. She has participated in several EC-funded ICT projects. Her research interests include video coding, wireless visual sensor networks and game-theoretic approaches in multimedia communication systems.

Lisimachos P. Kondi received the Diploma degree in electrical engineering from the Aristotle University of Thessaloniki, Greece, in 1994, and the M.S. and Ph.D. degrees in electrical and computer engineering from Northwestern University, Evanston, IL, USA, in 1996 and 1999, respectively. From 1999 to 2000, he was a Post-doctoral Research Associate at Northwestern University. He is currently an Associate Professor with the Department of Computer Science and Engineering, University of Ioannina, Greece. His research interests include the general areas of signal and image processing and communications, including image and video compression and transmission over wireless channels and the Internet, super-resolution of video sequences, and shape coding. He is an Associate Editor of the *EURASIP Journal on Advances in Signal Processing* and has served as an Associate Editor of the *IEEE SIGNAL PROCESSING LETTERS*.

Konstantinos E. Parsopoulos serves as an Assistant Professor at the Department of Computer Science and Engineering, University of Ioannina, Greece. His research is focused on computational optimization and modeling with an emphasis on metaheuristics and high-performance algorithms. His work counts one book and more than 85 papers in scientific journals, conferences, and edited volumes, while it has received more than 2500 citations.