GAME-THEORY-BASED CROSS-LAYER OPTIMIZATION FOR WIRELESS DS-CDMA VISUAL SENSOR NETWORKS

Lisimachos P. Kondi

University of Ioannina, Dept. of Computer Science [1]
Ioannina, Greece

Elizabeth S. Bentley

Air Force Research Laboratory [2]
Rome, NY, USA

ABSTRACT

We propose a game-theory-based cross-layer optimization scheme for wireless Direct Sequence Code Division Multiple Access (DS-CDMA) visual sensor networks. The scheme uses the Nash Bargaining Solution (NBS), which assumes that the nodes negotiate, with the help of a centralized control unit, on how to allocate resources. The NBS takes into account the video quality each node could achieve without making an agreement. The cross-layer optimization scheme determines the source coding rate, channel coding rate, and transmission power for each node. We compare the proposed game-theory-based scheme with competing schemes that minimize the average or maximum distortion among the nodes. Experimental results are presented and conclusions are drawn.

***Index Terms*—** Visual sensor networks, cross-layer optimization, game theory, Nash bargaining solution, DS-CDMA.

1. INTRODUCTION

This work is concerned with the cross-layer resource allocation for wireless visual sensor networks. These networks are comprised of typically low-weight distributed sensor nodes, equipped with video cameras, that can communicate with a centralized control unit at the network layer. The centralized control unit performs channel and source decoding to obtain the received video from each node. The control unit transmits information to the nodes in order to request changes in transmission parameters, such as source coding rate, channel coding rate, and transmission power. Applications of visual sensor networks include surveillance, automatic tracking and signaling of intruders within a physical area, command and control of unmanned vehicles, and environmental monitoring.

A major problem in a wireless visual sensor network is how to allocate the resources among the nodes. In a Direct Sequence Code Division Multiple Access (DS-CDMA) system, the transmission of a node causes interference to the transmissions of the other nodes. Thus, increasing the transmission power of one node will improve the received video quality of its transmission but will also degrade the received video qualities of the other nodes. It becomes clear that a joint optimization of the parameters of all nodes is required. In our previous work [1, 2], we proposed cross-layer optimization schemes that minimize either the average video distortion of all nodes, or the maximum distortion among the nodes. The former optimization criterion minimizes the average distortion but does not consider fairness issues among the nodes. The latter criterion optimizes the worst video quality among all the nodes in order to be fair.

[1] This research was supported by a Marie Curie International Reintegration Grant within the 7th European Community Framework Programme.

[2] Approved for Public Release; Distribution Unlimited: 88ABW-2010-0088, 25 JAN 2010

In this paper, we propose an alternative resource allocation between the nodes, which is based on the Nash Bargaining Solution (NBS) from game theory [3]. The NBS allocates resources as a result of a negotiation between the nodes with the help of the centralized control unit. The solution takes into account the video quality each node could achieve if it were to operate in a selfish manner, without negotiating. The video quality of each node should be at least as high as the quality it could achieve without negotiation. The NBS has been used before in communications problems such as the multiuser channel allocation for Orthogonal Frequency Division Multiple Access (OFDMA) networks [4]. It has also been used in video streaming for the allocation of the total bit rate among several video users [5]. However, in [5], no specific network setup is assumed. Also, the achievable video qualities in case of no negotiation (disagreement point) are arbitrarily selected. To the best of our knowledge, the present work is the first that applies the NBS to a video-quality based optimization of a wireless visual sensor network, where the video qualities at the disagreement point are determined based on what each node could achieve selfishly, and not arbitrarily selected.

The rest of the paper is organized as follows. In section 2, the basic architecture of the considered wireless visual sensor network is presented. In section 3, the proposed game-theory-based cross-layer optimization scheme is discussed. In section 4, experimental results are presented. Finally, in section 5, conclusions are drawn.

2. VISUAL SENSOR NETWORKS

We next describe the basic architecture of the considered wireless visual sensor network. DS-CDMA is used at the physical layer, while H.264 is used for source coding and Rate Compatible Punctured Convolutional (RCPC) codes are used for channel coding.

2.1. DS-CDMA

This work considers a wireless visual sensor network that utilizes DS-CDMA. In DS-CDMA, all users (nodes) transmit on the same frequency. In order to transmit a single bit, a node actually transmits L "chips". Thus, each node k is associated with a spreading code (signature sequence) $\mathbf{s}_k$, which is a vector of length L. Thus, in order to transmit the ith bit of a bit stream, node k actually transmits $b_k(i)\mathbf{s}_k$, which is a vector of L chips and $b_k(i)$ is either 1 or -1, depending on the value of the bit that is being transmitted. The node of interest receives interference from the other nodes. It is reasonable to assume that the interference can be approximated by Additive White Gaussian Noise (AWGN) [6]. Since user k has an associated power level in watts, $S_k = E_k R_k$, the energy per bit to Multiple Access

Interference (MAI) ratio becomes

$$\frac{E_k}{N_0} = \frac{\frac{S_k}{R_k}}{\sum_{j \neq k}^{K} \frac{S_j}{W_t}}; k = 1, 2, 3, ..., K \tag{1}$$

where E_k is the energy-per-bit, $N_0/2$ is the two-sided noise power spectral density due to MAI in watts/hertz, S_k is the power of the node-of-interest in watts, R_k is the transmitted bit rate in bits per second, S_j is the power of interfering node j in watts, and W_t is the total bandwidth in hertz [6]. R_k is taken to be the total bit rate used for source and channel coding. Assuming K users, R_k can be expressed as

$$R_k = \frac{R_{s,k}}{R_{c,k}}; k = 1, 2, 3, ..., K \tag{2}$$

where $R_{s,k}$ is the source coding rate for node k and $R_{c,k}$ is the channel coding rate for node k. Since $R_{s,k}$ has units of bits per second and $R_{c,k}$ is a dimensionless number, R_k will be measured in bits per second.

2.2. Source Coding

The video captured by the nodes is compressed using the H.264/AVC video coding standard. H.264/AVC has two conceptual layers, the video coding layer (VCL) and the network abstraction layer (NAL). The VCL forms the main part of the H.264/AVC and performs the required tasks for video compression to efficiently represent the content of the video data. The NAL achieves the network-friendly objective of H.264/AVC. It defines the interface between the VCL and the broad variety of systems and transport media. All data is encapsulated in NAL units, which contain an integer number of bytes. The NAL unit structure can be used in packet-based and bitstream-based systems. The difference in formatting lies in a unique start code prefix for resynchronization preceding the NAL unit in bitstream-based systems [7].

2.3. Channel Coding

In this work, we use Rate Compatible Punctured Convolutional (RCPC) codes for channel coding [8]. In our calculations, we use Viterbi's upper bounds on the bit error probability, P_b, given by

$$P_b \leq \frac{1}{P} \sum_{d=d_{free}}^{\infty} c_d P_d \tag{3}$$

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 is the probability that the wrong path at distance *d* is selected [8]. An AWGN channel with binary phase-shift keying (BPSK) modulation has a P_d given by

$$P_d = Q\left(\sqrt{\frac{2dR_cE_b}{N_0}}\right) \tag{4}$$

where $Q(.)$ is the Q-function for a Gaussian random variable, R_c is the channel coding rate and E_b/N_0 is the energy-per-bit normalized to the single-sided noise spectral density measured in watts/hertz.

3. OPTIMAL RESOURCE ALLOCATION

A centralized control unit at the network layer determines how network resources should be allocated amongst the nodes. It can request changes in transmission parameters, such as the source coding rates, channel coding rates, and transmission power levels. The constraint is that the chip rate be the same for all nodes. Assuming that all the nodes use the same spreading code length L, a constraint on the chip rate is equivalent to a constraint on the bit rate. In our previous work [1, 2], we used as optimization criterion either the minimization of the average distortion among the nodes, or the minimization of the maximum distortion among the nodes. In this paper, we propose the use of the Nash Bargaining Solution (NBS) from game theory [3]. The application of the NBS to our resource allocation problem is discussed next.

3.1. The Nash Bargaining Solution

The goal of our resource allocation problem is to determine the source coding rate, channel coding rate, and transmission power of each node, subject to a target bit rate constraint. The received video of node k will have expected distortion $E\{D_{s+c,k}\}$. As in [5], we define the *utility* x_k for node k as

$$x_k = \frac{c}{E\{D_{s+c,k}\}}, \tag{5}$$

where c is a positive constant.

Let $\mathbf{X}$ be the *feasible set* that consists of all possible vectors $(x_1, x_2, \ldots, x_K)$. Thus, each element of $\mathbf{X}$ comes from a different combination of source coding rates, channel coding rates and powers for the K nodes. The feasible set $\mathbf{X}$ is assumed to be convex, closed and bounded above.

Given the nature of the DS-CDMA channel, increasing the transmission power of one node will improve its video quality but will also degrade the video quality of the other nodes due to increased interference. Thus, the nodes should negotiate (with the help of the centralized control unit) in order to decide on a mutually acceptable member of set $\mathbf{X}$ (operating point). The result of the negotiation should give all nodes at least as high a utility as what they would get if they were to decide on their parameters independently, otherwise the agreement would do harm to them rather than good. Thus, if the negotiation fails, each node will decide on its parameters independently, with the goal of maximizing its utility, regardless of what the other nodes decide to do. Thus, the disagreement point $\mathbf{d} = (d_1, d_2, \ldots, d_K)$ is the vector of utilities that each video node can get without making a deal.

In our case, if no deal is made and the node of interest wants to maximize its video quality regardless of what the other nodes are doing, it will need to transmit with maximum power, since this will maximize its E_k/N_0 from Eq. (1), regardless of the other users' transmission powers. Clearly, this applies to all video nodes. Thus, $\mathbf{d}$ will consist of the utilities that each of the nodes gets when all nodes transmit with maximum power. Thus, to find $\mathbf{d}$, we will need to find the optimal allocation between source coding and channel coding for each node, given that all nodes transmit with maximum power. This corresponds to a *Nash equilibrium*, because the strategy of each of the nodes is a best reply to the strategies of the other nodes. (If a node wants to maximize its video quality regardless of what the other nodes are doing, it will have to transmit with maximum power).

Let us now define the *bargaining set*. The bargaining set consists of all Pareto-efficient payoff profiles (elements of $\mathbf{X}$) that assign all nodes at least as much as they can get at the disagreement point (without making a deal).

An agreement is Pareto-efficient when there is no other feasible agreement that all nodes prefer. Thus, a utility allocation $\mathbf{x} = (x_1, x_2, \ldots, x_K)$ is not Pareto-efficient if there is another allocation where each node gets a larger utility.

The Nash Bargaining Solution $F(\mathbf{X}, \mathbf{d})$ is a member of the bargaining set that satisfies the following axioms [3].

1. $F(\mathbf{X}, \mathbf{d}) \geq \mathbf{d}$.
2. $\mathbf{y} > F(\mathbf{X}, \mathbf{d}) \Rightarrow \mathbf{y} \notin \mathbf{X}$.
3. Given any strictly increasing affine transformation $\tau(.)$, $F(\tau(\mathbf{X}), \tau(\mathbf{d})) = \tau(F(\mathbf{X}, \mathbf{d}))$.
4. If $\mathbf{d} \in \mathbf{Y} \subseteq \mathbf{X}$, then $F(\mathbf{X}, \mathbf{d}) \in \mathbf{Y} \Rightarrow F(\mathbf{Y}, \mathbf{d}) = F(\mathbf{X}, \mathbf{d})$.

The first two axioms stipulate that the solution should lie in the bargaining set. The third axiom means that the solution should not depend on how the nodes calibrate their utility scales. Thus, if we scale the utility x as $Ax + B$, where $A > 0$, it should not make a difference in the solution. Also, since scaling doesn't matter, the selection of the constant c in the definition of utility is not important. The fourth axiom is the *Independence of Irrelevant Alternatives*. If $\mathbf{x} \in \mathbf{Y} \subseteq \mathbf{X}$ and $\mathbf{x}$ is the solution to the problem when the feasible set is $\mathbf{X}$, then $\mathbf{x}$ should also be the solution when the feasible set is $\mathbf{Y}$. Since points in $\mathbf{X}$ that do not belong in $\mathbf{Y}$ were not chosen as the solution to the problem where the feasible set is $\mathbf{X}$, their unavailability in the problem where the feasible set is $\mathbf{Y}$ should be irrelevant.

To find the Nash Bargaining Solution, we need to find the source coding rate, channel coding rate and power for each node (which determine the vector of utilities $\mathbf{x} = (x_1, x_2, \ldots, x_K)$) so that the Nash product is maximized [3]:

$$F(\mathbf{X}, \mathbf{d}) = \arg\max_{\mathbf{x}} (x_1 - d_1)^{\alpha_1} (x_2 - d_2)^{\alpha_2} \cdots (x_K - d_K)^{\alpha_K}, \tag{6}$$

subject to the requirement that $\mathbf{x} \geq \mathbf{d}$, where α_i is the *bargaining power* of node i. The bargaining power of a node depends on whether it is advantaged or disadvantaged by its role in the bargaining game [3]. In our setup, there is no reason to assume that some nodes are more advantaged than others, thus we assume that $\alpha_i = 1$ for all i. Then, the Nash product becomes

$$F(\mathbf{X}, \mathbf{d}) = \arg\max_{\mathbf{x}} (x_1 - d_1)(x_2 - d_2) \cdots (x_K - d_K). \tag{7}$$

Thus, in order to apply the Nash Bargaining Solution to our cross-layer optimization problem, we will first need to determine vector $\mathbf{d} = (d_1, d_2, \ldots, d_K)$, which contains the maximum utility for each video node when all nodes transmit at maximum power

Once $\mathbf{d}$ is found, we will just need to determine the source coding rates, channel coding rates, and transmission powers, that will maximize the Nash product (Eq. (7)).

We partition the sensor nodes into two classes: Nodes that image low-motion scenes (low-motion nodes) and nodes that image high-motion scenes (high-motion nodes). Let us now assume that there are K_{low} low-motion nodes and K_{high} high-motion nodes, where $K_{low} + K_{high} = K$. Then, the NBS will be a vector (x_{low}, x_{high}), such that

$$F(\mathbf{X}, \mathbf{d}) = \arg\max_{\mathbf{x}} (x_{low} - d_{low})^{K_{low}} (x_{high} - d_{high})^{K_{high}}, \tag{8}$$

where $\mathbf{d} = (d_{low}, d_{high})$ is the disagreement point. Here, we assume two classes of nodes. However, the scheme can be generalized to any number of classes.

3.2. Optimization Solution

The expected distortion $E\{D_{s+c,k}\}$ (and thus the utility x_k) of node k depends on the source coding rate $R_{s,k}$, the channel coding rate $R_{c,k}$ and transmission power S_k selected for user k, as well for the transmission powers $S_i, i \neq k$, selected for all the other users (interferers). Thus, the goal of the optimization is to determine the $R_{s,k}$, $R_{c,k}$ and S_k, $k = 1, \ldots, K$ so that the Nash product in Eq. (7) is maximized. The problem is a discrete optimization problem, that is, $R_{s,k}$, $R_{c,k}$, and S_k can only take values from discrete sets $\mathbf{R_s}$, $\mathbf{R_c}$, and $\mathbf{S}$, respectively, i.e., $R_{s,k} \in \mathbf{R_s}$, $R_{c,k} \in \mathbf{R_c}$, $S_k \in \mathbf{S}$.

Since it would be prohibitively complex to experimentally obtain the expected distortion for each user for all possible combinations of source coding rate, channel coding rate, and power level, we have used the Universal Rate Distortion Characteristics. These characteristics model the expected distortion as a function of the bit error rate after channel decoding. As in [9], we assume the following model for the URDC for each user k

$$E\{D_{s+c,k}\}(R_{s,k}, P_b) = a \left[\log_{10}\left(\frac{1}{P_b}\right)\right]^{-b} \tag{9}$$

where $a > 0$ and $b > 0$ are determined using mean square optimization from a few $(E\{D_{s+c,k}\}, P_b)$ pairs that are obtained experimentally. a and b depend on the video sequence and source coding rate $R_{s,k}$.

Therefore, the expected distortion $E\{D_{s+c,k}\}$ of user k is determined using the following procedure. First, for a given set of transmission powers, the E_k/N_0 is determined from Eq. (1). Then, for a given selection of $R_{c,k}$, the bit error rate P_b after channel decoding is estimated using Eqs. (3) and (4), with $E_b = E_k$. Finally, for a given selection of $R_{s,k}$, the expected distortion is estimated using Eq. (9).

4. EXPERIMENTAL RESULTS

A number of experiments were conducted, some of which are presented here. We assumed that the visual sensor network nodes belong to one of two classes, depending on the amount of motion in the scene they are viewing: low-motion class and high-motion class. The "Foreman" video sequence was used to represent the scene viewed by a high-motion node, while the "Akiyo" video sequence was used to represent the scene viewed by a low-motion node. Thus, it is necessary to have two sets of URDC curves, one for each level of motion. The characteristics were obtained for both video sequences at a frame rate of 15 frames/s. The data points used to obtain the parameters a and b in Eq. (9) were obtained by corrupting the video stream with packet errors based on a bit error rate P_b, decoding the corrupted video bit stream with the H.264/AVC codec, calculating the distortion, repeating this experiment 300 times and then taking the average distortion. We chose $c = 255^2$ so that the definition of utility is analogous to the definition of PSNR. However, as mentioned previously, the choice of c does not affect the results.

We assumed Binary Phase Shift Keying (BPSK) modulation and RCPC codes with mother rate $1/4$ from [8]. We set the link layer packet size to 400. The total bandwidth, W_t, was set to 5 MHz. For the results presented here, we assumed a target bit rate of $R_k = 96kbps$ for all k. The set of admissible source coding rates and corresponding channel coding rates were: $\mathbf{C} \in \{1 : (32kbps, 1/3), 2 : (48kbps, 1/2), 3 : (64kbps, 2/3)\}$. The power levels were chosen from $\mathbf{S} \in \{5, 10, 15\}$ watts.

We performed the cross-layer optimization using the Nash Bargaining Solution and compared its results with the method of minimizing the average expected distortion of the nodes (MAD) and the method of minimizing the maximum distortion among the nodes (MMD) [1, 2]. We report expected PSNR values instead of expected distortion values, since they are equivalent. Tables 1, 2, 3,

	S_1 (W)	S_2 (W)	C_1	C_2	$PSNR_1$	$PSNR_2$
MAD	15	10	1	1	25.24	28.33
MMD	10	5	1	1	25.79	25.64
NBS	5	5	1	1	24.19	30.81

Table 1. Optimal resource allocation for the three criteria (MAD, MMD, NBS) for 25 high-motion and 25 low-motion users.

	S_1 (W)	S_2 (W)	C_1	C_2	$PSNR_1$	$PSNR_2$
MAD	15	10	1	1	26.13	30.32
MMD	10	5	1	1	26.48	27.82
NBS	5	5	1	1	25.45	32.83

Table 2. Optimal resource allocation for the three criteria (MAD, MMD, NBS) for 25 high-motion and 15 low-motion users.

	S_1 (W)	S_2 (W)	C_1	C_2	$PSNR_1$	$PSNR_2$
MAD	15	10	1	1	27.11	32.31
MMD	10	5	1	1	27.24	29.83
NBS	5	5	1	3	26.85	35.69

Table 3. Optimal resource allocation for the three criteria (MAD, MMD, NBS) for 25 high-motion and five low-motion users.

	S_1 (W)	S_2 (W)	C_1	C_2	$PSNR_1$	$PSNR_2$
MAD	15	10	1	1	26.60	31.30
MMD	10	5	1	1	27.24	29.83
NBS	5	5	1	1	25.45	32.83

Table 4. Optimal resource allocation for the three criteria (MAD, MMD, NBS) for 15 high-motion and 25 low-motion users.

4 and 5 show the results of the three cross-layer optimization methods for the cases of 25 high-motion and 25 low-motion nodes, 25 high-motion and 15 low-motion nodes, 25 high-motion and five low-motion nodes, 15 high-motion and 25 low-motion nodes, and five high-motion and 25 low-motion nodes, respectively. S_i is the transmission power, C_i is the source-channel coding rate combination and $PSNR_i$ is the expected PSNR in dB, where $i = 1$ refers to the high-motion nodes and $i = 2$ refers to the low-motion nodes. It can be seen that the NBS results in a higher PSNR for the low-motion nodes, compared to the other methods. This is because d_{low}, the maximum utility of the low-motion nodes when all nodes transmit at maximum power (when no deal is made), is already higher than the utility the low-motion nodes would get under MAD and MMD. Of course, with the NBS, the PSNR of the high-motion nodes decreases. By looking at the presented results, we can see that the PSNR increase of the low-motion nodes is larger than the PSNR decrease of the high-motion nodes, except when the number of low-motion nodes is much larger than the number of high-motion nodes (as in Table 5). Thus, we can say that the NBS is the preferred optimization criterion, unless the low-motion nodes heavily outnumber the high-motion nodes. The NBS solution has the additional benefit that the transmission powers of all the nodes are lower than with the other criteria.

5. CONCLUSIONS

A major problem in wireless visual sensor networks is how to allocate the available system resources fairly among the nodes. We have presented a cross-layer resource allocation scheme that is based on the Nash Bargaining Solution (NBS). The NBS allocates resources as a result of an agreement between the nodes, taking into account the video quality each node could receive without making an agreement. We have experimentally compared our criterion with two competing criteria, which minimize the average or the maximum distortion among the nodes. Our experimental results show that the NBS is the preferred optimization criterion, unless the low-motion nodes heavily outnumber the high-motion nodes.

6. REFERENCES

[1] E. S. Pynadath and L. P. Kondi, "Cross-layer optimization with power control in DS-CDMA visual sensor networks," in *Proc. of International Conference on Image Processing*, Atlanta, GA, 2006.

[2] E. S. Bentley, L. P. Kondi, J. D. Matyjas, and M. J. Medley, "Wireless visual sensor network resource allocation using cross-layer optimization," in *Proc. Visual Communications and Image Processing Conference*, San Jose, CA, 2009.

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

[4] Z. Han, J. Zhu, and K. J. R. Liu, "Fair multiuser channel allocation for OFDMA networks using Nash bargaining solutions and coalitions," *IEEE Transactions on Communications*, vol. 53, no. 8, pp. 1366–1376, Aug. 2005.

[5] H. Park and M. van der Schaar, "Bargaining strategies for networked multimedia resource management," *IEEE Transactions on Signal Processing*, vol. 55, no. 7, Jul. 2007.

[6] Y. S. Chan and J. W. Modestino, "A Joint Source Coding-Power Control Approach for Video Transmission Over CDMA Networks," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 10, pp. 1516–1525, December 2003.

[7] T. Wiegand and G. Sullivan, "The H.264/AVC Video Coding Standard Standards in a nutshell," *IEEE Signal Processing Magazine*, vol. 24, no. 2, pp. 148–153, 2007.

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

[9] L. P. Kondi and A. K. Katsaggelos, "Joint Source-Channel Coding for Scalable Video Using Models of Rate-Distortion Functions," in *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Salt Lake City, Utah, 2001, pp. 1377–1380.

	S_1 (W)	S_2 (W)	C_1	C_2	$PSNR_1$	$PSNR_2$
MAD	10	5	3	1	31.94	33.90
MMD	15	5	3	1	34.58	32.83
NBS	5	5	1	3	26.85	35.69

Table 5. Optimal resource allocation for the three criteria (MAD, MMD, NBS) for five high-motion and 25 low-motion users.