

Energy-efficient joint video encoding and transmission framework for WWSN

Othmane Alaoui-Fdili^{1,2} · François-Xavier Coudoux³ ·
Youssef Fakhri^{2,4} · Patrick Corlay³ · Driss Aboutajdine²

Received: 20 October 2015 / Revised: 6 December 2016 / Accepted: 2 June 2017 /
Published online: 7 July 2017
© Springer Science+Business Media, LLC 2017

Abstract In this paper, we propose an energy-efficient joint video encoding and transmission framework for network lifetime extension, under an end-to-end video quality constraint in the Wireless Video Sensor Networks (WWSN). This framework integrates an energy-efficient and adaptive intra-only video encoding scheme based on the H.264/AVC standard, that outputs two service differentiated macroblocks categories, namely the Region Of Interest and the Background. Empirical models describing the physical behavior of the measured energies and distortions, during the video encoding and transmission phases, are derived. These models enable the video source node to dynamically adapt its video encoder's configuration in order to meet the desired quality, while extending the network lifetime. An energy-efficient and reliable multipath multi-priority routing protocol is proposed to route the encoded streams to the sink, while considering the remaining energy, the congestion level as well as the packet loss rates of the intermediate nodes. Moreover, this protocol interacts with the application layer in order to bypass congestion situations and continuously feed it with current statistics. Through extensive numerical simulations, we demonstrate that the proposed framework does not only extend the video sensor lifetime by 54%, but it also performs significant end-to-end video quality enhancement of 35% in terms of Mean Squared Error (MSE) measurement.

Driss Aboutajdine died before publication of this work was completed.

✉ Othmane Alaoui-Fdili
othmane.alaoui.fdili@gmail.com

¹ Ecole Supérieure de Technologie de Safi, Université Cadi Ayyad, Safi, Morocco

² LRIT-CNRST 29, Faculté des sciences, Université Mohammed V de Rabat, Rabat, Morocco

³ IEMN UMR 8520, Department OAE, University of Valenciennes and Hainaut-Cambresis, Valenciennes, France

⁴ LARIT Networks and Telecommunications Team, Ibn Tofail University, Kenitra, Morocco

Keywords Energy-efficiency in WWSN · Cross-layer design · Energy consumption prediction · Video distortion prediction · Energy-efficient reliable routing

1 Introduction

With the recent developments in CMOS technology, tiny low-cost imaging sensors have been developed. As a result, the sensor nodes are able to capture visual information, which enhance all the previous sensor-based applications, especially those that are monitoring-oriented. Hence, a new derivative of the Wireless Sensor Networks (WSN), which is the Wireless Video Sensor Network (WWSN) [3, 28], has been introduced. The sensor nodes are constrained in terms of memory, processing, and the most important resource: the energy. In a WWSN, the generated visual information is huge, compared to the case of WSN, and needs to be processed prior to transmission. This operation is commonly called the compression or the encoding. The energy consumed during this processing has to be taken into account [3]. According to the literature [21], the video compression task can consume about 2/3 of the total power for video communications over wireless channels. For these reasons, energy-efficient compression schemes adapted to the WWSNs are needed.

Another important issue in the WWSNs that is strongly recommended [28], for efficient data communication, is the differentiated service paradigm. For example, based on the relevance of the video data, the routing protocol may provide differentiated treatments to the packets, by achieving high reliability only to the most significant packets. In general, achieving more reliability in a WWSN means consuming more energy. Therefore, thanks to this paradigm, significant amount of energy may be conserved. In fact, this feature should be first introduced at the application layer and is afterwards respected at the lower layers. This interaction between layers is formally called the Cross-Layering. It is considered as the most efficient solutions' design to deal with the stringent constraints of the WWSN [8], and to meet the application's requirements with a low complexity.

In this paper, a novel energy-efficient and cross-layer designed framework for video delivery over WWSN is proposed. This approach involves three layers: application, routing and MAC. We adopt an energy-efficient and adaptive intra-only video compression scheme dedicated to the WWSNs, in the application layer. Actually, this video compression scheme, which was first proposed in our previous work [4], is based on the H.264/AVC standard using intra-only coding mode. In this scheme, when an event occurs, the video encoder switches to a so called rush mode, adopts a higher Frame Rate (FR) and produces two service differentiated Macroblocks (MB) categories, namely the Region of Interest (ROI) and the Background (BKGD). Afterwards, a simple bit rate adaptation technique, called the Frequency Selectivity (FS), is applied to the BKGD in order to decrease the energy consumed during the transmission process. This operation is controlled by the Frequency Selectivity Parameter (FSP) that represents the number of the retained coefficients after the transform. A block diagram of the proposed video compression scheme in [4] is given in Fig. 1. This solution has proven that considerable energy is saved during the transmission process. However, the consumed energy during the encoding process was not considered.

The novelty of the paper compared to this previous work is that we propose here empirical models in line with the physical behavior of the energies and the distortions measured during the video encoding and transmission phases. The proposed models are expressed as functions of three controlling parameters of the video encoder [4] which are the FR, the Quantization Parameter (QP) and the FSP. To the best of our knowledge, this is the first work that proposes models considering this particular combination of controlling

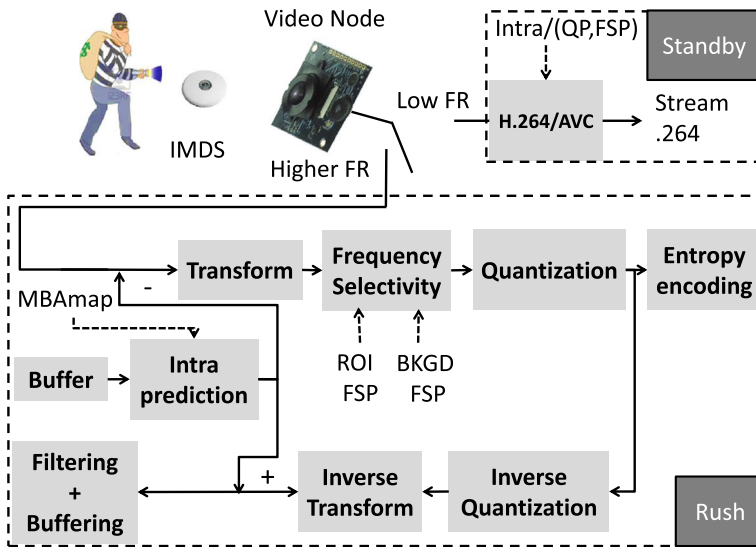


Fig. 1 Block diagram of the proposed video compression scheme [4]

parameters. Using the proposed models, the video source node may predict these energies and distortions and then adapt its video encoder’s configuration in order to decrease the global energy consumption, while meeting the desired video quality. In addition, as second major original contribution, we propose a novel Energy-efficient and Reliable Multipath Multi-priority routing protocol (ERMM) to handle the encoded video streams. In fact, ERMM considers the neighboring nodes whose the Remaining Energy (RE) is above the average, then estimates their reliability using a proposed metric, in order to deliver High Priority (HP) and Low Priority (LP) packets (i.e. ROI and BKGD) according to the desired reliability. Moreover, this protocol interacts with the application layer (i.e. video encoder) in order to bypass congestion situations and continuously feed it with current statistics. In addition, this protocol relies on the cooperation of the MAC layer as well, in order to guarantee the differentiated service, by giving the high priority packets higher chances to access the shared medium and to keep track of the current loss rates.

The main contributions of this paper can be summarized as follows: (i) empirical models for predicting the energy consumption and the distortion during the video encoding and the transmission phases are derived and evaluated, (ii) a new energy-efficient and reliable multipath multi-priority routing protocol that serves each packet according to the region which its payload data belongs (i.e. ROI or BKGD) is proposed, (iii) a novel framework that integrates and defines the interactions between all of the presented components, for network lifetime extension under a desired end-to-end video quality distortion, is described and (iv) an extensive evaluation of the proposed framework in terms of energy-efficiency, reliability, delay, and video quality is exposed.

The remainder of this paper is organized as follows. Section 2 presents a review of the cross-layer approaches and frameworks dedicated to the WVSNs. In Section 3, empirical models for predicting the energy consumption and the distortion caused by the encoding and the transmission phases are derived then validated. These models are used by the video encoder representing the first component of our framework. In Section 4, the second component of the framework, namely the proposed ERMM routing protocol is detailed. Section 5

explains the interactions between the different proposed components of the framework. In Section 6, the proposed framework is evaluated in terms of energy consumption, reliability, delay and video quality. Finally, the paper is concluded in Section 7.

2 Related works

We investigate in this section the most relevant and recent works proposing cross-layer approaches for WVSNS. In [23], Politis et al. propose a cross-layer approach involving the application and the routing layers. In the application layer, they adopt the classical implementation of the H.264/AVC standard for the video compression at the source node. In order to compensate the energy drained during this phase, the well known LEACH [15] protocol for clustering is first used. In addition, they introduce two packet scheduling algorithms, namely Baseline and Power aware, that selectively drop packets and track the video distortion generated by this drop by a proposed recursive distortion prediction model. However, the authors in [23] did not consider the energy consumed during the compression phase, especially when inter-prediction is performed.

Aghdasi et al. [1] present a cross-layer approach influencing the application, the transport and the network layers. The authors introduce the prioritization paradigm by generating two service differentiated frames using a proposed encoder called Modified-MPEG. The transport protocol uses inter-layer command messages to be aware of the network status. The packets' retransmissions are prohibited to perform real-time video transmission. A single path routing protocol based on hop-by-hop negotiation with a selected parent is adopted. However, there is no clear distinction between the energy consumed for the transmission and the encoding. Moreover, prohibiting retransmissions as well as multipath routing have a negative impact on the reliability performances, especially in the WVSNS.

A framework for image transmission with priority routing is presented by Felemban et al. [11]. In the application layer, they perform data prioritization using a bit planes based separation. For routing, the authors choose as the next hop, the neighbour whose the number of hops from the sink is the least. In MAC layer, a priority based enqueueing/dequeueing procedure is introduced. However, choosing the next hop following only one dimension metric with no load balancing mechanisms, results in concentrating the load on specific nodes, leading to their quick death. Additionally, no energy consumption evaluation of the framework is presented.

In [26], the authors propose a cross-layer framework that interconnects the application, network and MAC layers. In the application layer, the Wyner-Ziv lossy distributed source coding is used at the sensor node with variable group of pictures size. The models for predicting the distortion caused by the encoding and the packet loss are also provided. Then, the generated frames are routed using a source directed multipath routing. For service differentiation, the IEEE 802.11e MAC [16] layer is adopted and packets containing I-Frames data are given the HP, while those containing W-Frames data are processed with LP. This work is quite similar to our proposal in large, but the specific objective is different. Our work aims to extend the network lifetime under a video quality constraint, while the authors in [26] aim to maximize the capacity of the deployed network. Nevertheless, they do not provide a model for video encoding energy consumption.

The issue of network lifetime extension under a video quality constraint is addressed in [12] and [30]. In [12], the authors present a distributed algorithm for maximizing the network lifetime by jointly optimizing the source rates, the encoding powers, and the routing scheme. In fact, they first consider the energies consumed as well as the distortions of both

transmission and encoding phases in the formulation of the optimization problem. Then, Lagrangian duality is used to solve this problem. The authors in [30] propose a framework for source and channel rate adaptation, where the energy consumption of source and network encoding, error control and data communication are considered. In fact, models for predicting these quantities are proposed. Thereafter, a weighted convex optimization problem is formulated to be decomposed into a two-level optimization procedure. The Lagrangian duality is also used for solving this problem.

3 Video encoder's prediction models

This section focuses on the mathematical tools used by the framework's first component (i.e. the video encoder). Actually, these tools are empirical models enabling the video encoder to dynamically select the appropriate configuration, with the concern of decreasing the energy consumption while meeting a desired end-to-end video quality. The unique mathematical model that describes well the energy consumption and the video distortion is used in [12] and [30], and was first proposed in [14]. Nevertheless, the authors in [13] highlight that their analytical approach described in [14] cannot be easily extended to other video encoders such as H.264/AVC standard. Hence, they propose in [13] an operational approach for off-line energy consumption analysis and modelling, which can be applied to generic video encoders. It consists on the consideration of the parameters that are responsible for consuming more or less energy during the video signal compression. In our case, these parameters are the FR, the QP and the FSP.

In the present work, we adopt an intra-only encoder based on the H.264/AVC standard. Motivations behind this choice are twofolds. First, it is well-known that motion estimation and compensation are highly complex procedures. In fact, according to the experiments presented in [2], conducted on Stargate/Telos video sensor nodes, the computational energy depletion constitutes the major portion (more than 90%) of the total energy consumption as compared to the bit transmission energy. In addition, the inter coding with motion estimation consumes more than 10 times the energy drained by the intra coding. Hence, choosing intra-only coding mode significantly limits energy consumption of resource-limited video nodes. In order to decrease the high bit rate generated by this mode, we compensate it by using the frequency selectivity and adopting lower frame rate. Second, intra-only coding avoids error propagation across successive frames, hence preserving video quality. Therefore, there is no temporal losses propagations. Moreover, the Flexible Macroblock Ordering (FMO) tool [19], which is provided by the H.264/AVC standard, is used to separately encode two regions in each frame. Hence, the ROI is protected from spatial losses propagations occurred in the BKGD and vice-versa.

To the best of our knowledge, no prior work has proposed models for energy and distortion prediction for the encoding and transmission phases, considering this particular combination of controlling parameters, for an intra-only H.264/AVC based video encoder. Therefore, in order to derive appropriate models, extensive tests were carried out for energy and distortion measurements using several video sequences, on a PC with an Intel 2.93 GHz Core 2 Duo processor. The ITU-T recommends for the selection of appropriate test sequences to consider the Spatial Information (SI) and the Temporal Information (TI) of the scenes [18]. In fact, these parameters reflect the compression difficulty as well as the level of impairment that is suffered when the scene is transmitted. Furthermore, it is important to choose sequences that span a large portion of the spatial-temporal information plane [18]. Specifically, we use eight video sequences, Hall, Foreman, Soccer, Ice in QCIF

resolution (176×144), and Bus, City, Stefan and News in CIF resolution (352×288). Figure 2 illustrates the distribution of the test video sequences over the spatial-temporal information plane. As can be seen, the considered set of sequences covers a large area, showing its SI-TI diversity. In fact, according to Fig. 2, this set includes video sequences representing contents with a moderate TI and SI ranging from low to high (i.e. City, Foreman, Ice, News, Stefan and Bus), while others represent contents with a moderate SI and TI ranging from low to high (i.e. Soccer, Foreman, Ice and Hall). One can expect that only the SI index is relevant since intra-only coding is performed. However, the TI index also matters since error concealment at the decoder is performed by slice copy from the previous decoded frame, as further discussed in Section 3.2.

Each sequence is encoded at 14 different QPs and 16 different FSPs at 30 fps using our modified JM18.4 implementation of the H.264/AVC standard, as illustrated by the Fig. 1, in its baseline profile. In the rest of the paper, the curves represent the models and the points correspond to the measurements. Due to the lack of space and for the sake of visibility, only four sequences and eight QPs are shown.

3.1 Energy consumption modelling

The video source node spends its lifetime encoding and transmitting the video streams. Therefore, the energy drained by a node E_{Total} can be approximated by:

$$E_{Total} = E_{Enc} + E_{Trans} \tag{1}$$

where E_{Enc} and E_{Trans} refer to the consumed energy for encoding and transmitting the video stream respectively.

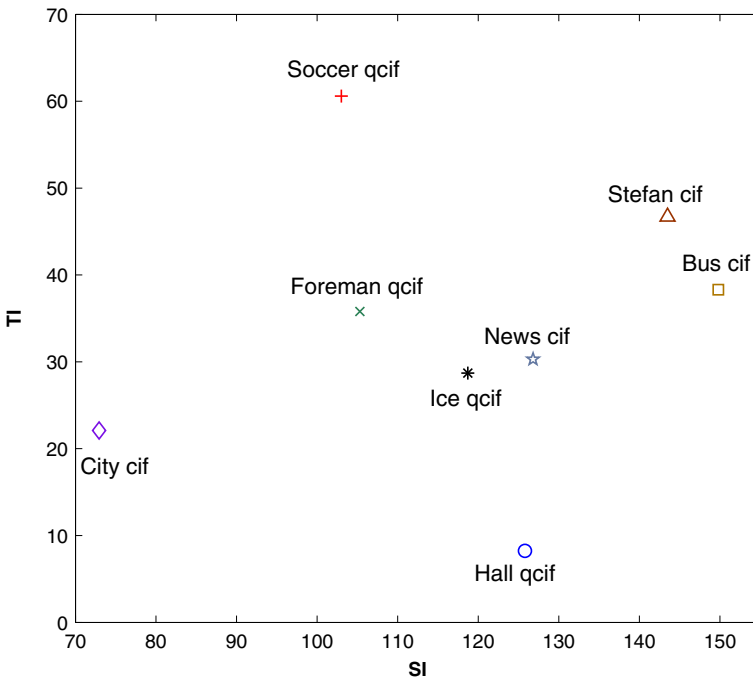


Fig. 2 The coverage of the considered video sequences over the spatial-temporal information plane

In our previous work [5], we only address the impact of FR and QP variations, on the encoding energy consumption of an intra-only H.264/AVC encoder. We have shown that the energy decreases by about the half when the Frame skip (Fskip) is increased by one (i.e. decreasing the FR by the half), which is an obvious behaviour. In this Section 3.1, we extend this approach by investigating and modelling the joint effect of QP and FSP variations on the consumed energy, for both video encoding and transmission phases.

3.1.1 Encoding energy modelling

The consumed energy for processing a specific task can be approached, as presented in [24], by a number of clock cycles dependent function, given as follows:

$$E_{Proc}(N) = N_{cyc} \cdot C_{total} \cdot V_{dd}^2 + V_{dd} \cdot (I_0 \cdot e^{\frac{V_{dd}}{n \cdot V_T}}) \left(\frac{N_{cyc}}{freq} \right) \quad (2)$$

where N_{cyc} is the number of clock cycles, C_{total} is the average capacitance switched per cycle, V_{dd} is the supply voltage, I_0 is the leakage current, $freq$ is the clock speed, V_T is the thermal voltage and n a processor dependant constant. On the other hand:

$$N_{cyc} = T_{Proc} \cdot freq \quad (3)$$

where T_{Proc} is the processing time of the task, which is in our case the encoding time returned by the encoder.

For reliable and accurate measurements, we control the running of the encoder to be in a real-time fashion and make it the only executed process in one microprocessor. Once the measurements are converted to energies, using the model of (2), they are normalized by the maximal consumed energy. The different coefficients of (2) are available on Intel's website [17].

Points in Fig. 3 show the behaviour of the average normalized measured energy consumption during the encoding process for different QPs and FSPs over 200 realisations. For a given FSP, the normalized measured energy decreases when the QP increases. The same behaviour is observed when the FSP decreases for a given QP. Those behaviours can be explained by the fact that when increasing the QP and decreasing the FSP, more zero coefficients are generated, leading to a simple entropy coding scheme. Actually, the points in Fig. 3 illustrate the behaviour of (QP,FSP)-dependent coefficients that reduce the maximal energy reached for (QP_{min}, FSP_{max}) . For each QP, the maximal value of the coefficients is reached at $FSP = FSP_{max}$, then decreases toward the minimal value reached at $FSP = FSP_{min}$. These observations suggest that the (QP,FSP)-dependent coefficients, noted $C_{EE}(\cdot)$, can be modelled by:

$$C_{EE}(QP, FSP) = \alpha_{EE}(QP, FSP_{max}) \cdot \beta_{EE}(QP, FSP) \quad (4)$$

Points in Fig. 4 report the measured $\alpha_{EE}(QP, FSP_{max})$ that expresses the effect of the quantization on the normalized consumed energy during the encoding process. Actually, this was investigated in detail in our previous work [5], where all of the coefficients were kept, which corresponds to the (QP, FSP_{max}) configuration. We have seen that $\alpha_{EE}(\cdot)$ can be predicted by the following model:

$$\alpha_{EE}(QP, FSP_{max}) = 2 - \exp^{a \cdot QP} \quad (5)$$

where a is a content-dependent coefficient obtained by minimizing the Root Mean Squared Error (RMSE) between measured and predicted data. As shown in Fig. 4, the proposed model in (5) can predict the reduction coefficients $\alpha_{EE}(\cdot)$ accurately.

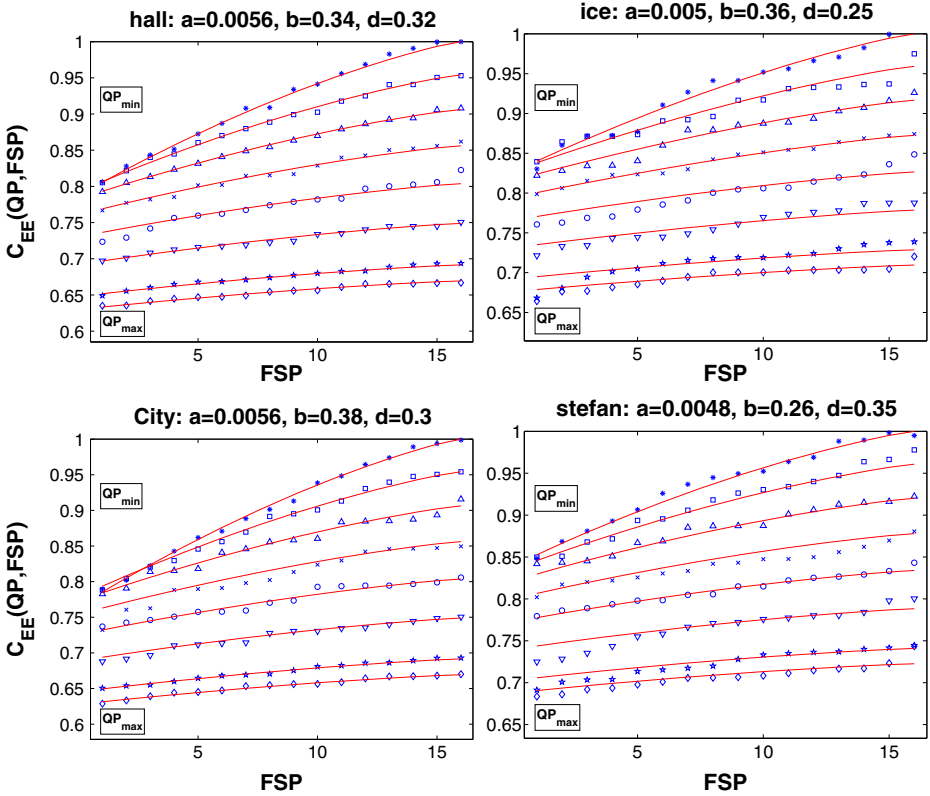


Fig. 3 Measured (points) and predicted (curves) normalized encoding energy $C_{EE}(QP, FSP)$. Curves are predicted values by the proposed model of (4)

Points in Fig. 5 reports the measured $\beta_{EE}(QP, FSP)$, representing the introduced energy consumption reduction when decreasing the FSP for a given QP. According to the measured data in Fig. 5, the appropriate functional form for $\beta_{EE}(\cdot)$ has to meet some conditions.

First, when the FS is applied with the maximal FSP, no energy consumption reduction should be observed. Hence, the unique energy consumption reduction should be the one achieved by the quantization (5). Second, the more coefficients are retained in the FS phase, the less energy-efficiency there is. Finally, when the quantization becomes more severe, the effect of the FS becomes less important. Based on the above mentioned arguments, we propose the following functional form for $\beta_{EE}(\cdot)$:

$$\beta_{EE}(QP, FSP) = \frac{1 - b \cdot e^{(c \cdot QP_{norm})} \cdot e^{(d \cdot (FSP^*_{norm})^f)}}{1 - b \cdot e^{(c \cdot QP_{norm})}} \tag{6}$$

where $QP_{norm} = \frac{QP}{QP_{max}}$, FSP^* is the dual of FSP ($FSP^* = FSP_{max} - FSP$) and $FSP^*_{norm} = \frac{FSP^*}{FSP^*_{max}}$. According to several tests, we conclude that c can be fixed to -1 and f fixed to 1.2. The parameters b and d are content dependent.

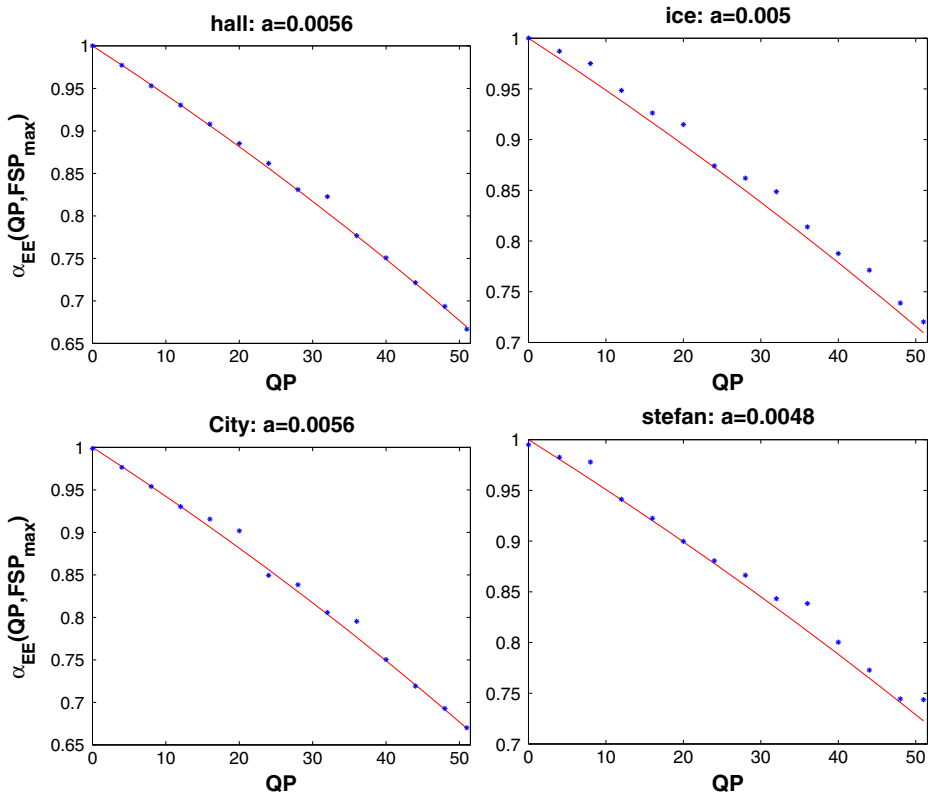


Fig. 4 Measured (points) and predicted (curves) $\alpha_{EE}(QP, FSP_{max})$. Curves are predicted values by the proposed model of (5)

As can be seen in Fig. 5, the proposed model in (6) describes well the behaviour of $\beta_{EE}(QP, FSP)$. The overall model for estimating the energy consumed during the video encoding process, considering the QP and the FSP, can be formulated by:

$$E_E(QP, FSP) = E_E(QP_{min}, FSP_{max}) \cdot \alpha_{EE}(QP, FSP_{max}) \cdot \beta_{EE}(QP, FSP) \quad (7)$$

Note that $E_E(QP_{min}, FSP_{max})$ depends mainly of the video content, the QP_{min} , the FSP_{max} , the FR, as well as the microprocessor’s features. Curves in Fig. 3 show that the model, proposed in (4), describes well the consumed encoding energy during the encoding process with an average prediction error of 6% in terms of RMSE.

3.1.2 Transmission energy modelling

The energy consumed per bit during the transmission process is given by the well-known model proposed in [15]:

$$E_{Tx} = E_{elec} + \epsilon_{fs} \times d^2 \quad (8)$$

where ϵ_{fs} is the energy consumed by the amplifier to transmit at short distance, E_{elec} is the energy dissipated in the electronic circuit to transmit and receive the signal and d is the distance between the transmitter and the receiver. Based on this model, the energy consumed

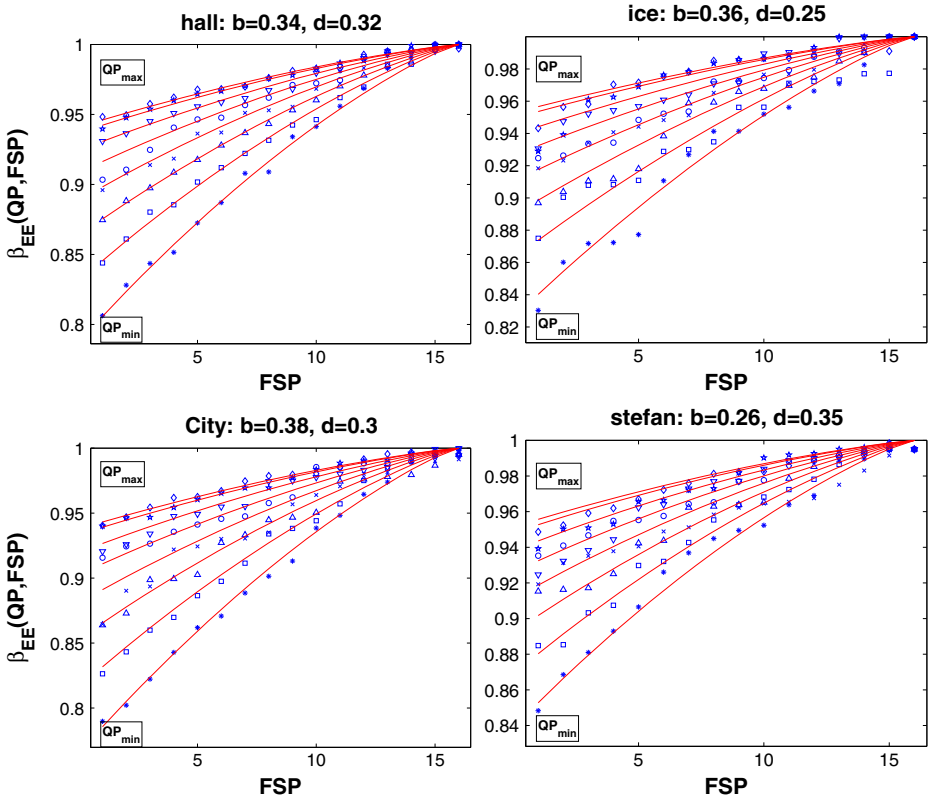


Fig. 5 Measured (points) and predicted (curves) $\beta_{EE}(QP, FSP)$. Curves are predicted values by the proposed model of (6)

for transmitting, at a given distance d at a fixed ϵ_{fs} and E_{elec} , depends directly of the data rate generated by the encoder.

Points in Fig. 6 represent the collected normalized energy consumed during the transmission phase of the conducted tests. The first observation is that the joint effect of the quantization and the FS on the energy drained for encoding is approximately the same as the one of the energy drained for the transmission. The second one is that the achievable energy-efficiency by increasing the QP and decreasing the FSP is more important in the transmission than in the encoding, which is a logical result. The reason is that during the encoding process, even if the zero coefficients makes the CAVLC simpler, there is some energy that is dissipated for their processing (comparing, counting, . . .). Moreover, for transmitting the encoded signal, considerable energy-efficiency is observed due to the efficient encoding.

Similar to Section 3.1, the points in Fig. 6 can be seen as the behaviour of (QP,FSP)-dependent coefficients that reduce the maximal energy reached also for (QP_{min}, FSP_{max}) . For each QP, the maximal value of the coefficients is reached at $FSP = FSP_{max}$, then decreases toward the minimal value reached at $FSP = FSP_{min}$. These observations suggest that these (QP,FSP)-dependent coefficients, noted $C_{TE}(\cdot)$, can be modelled by:

$$C_{TE}(QP, FSP) = \alpha_{TE}(QP, FSP_{max}) \cdot \beta_{TE}(QP, FSP) \tag{9}$$

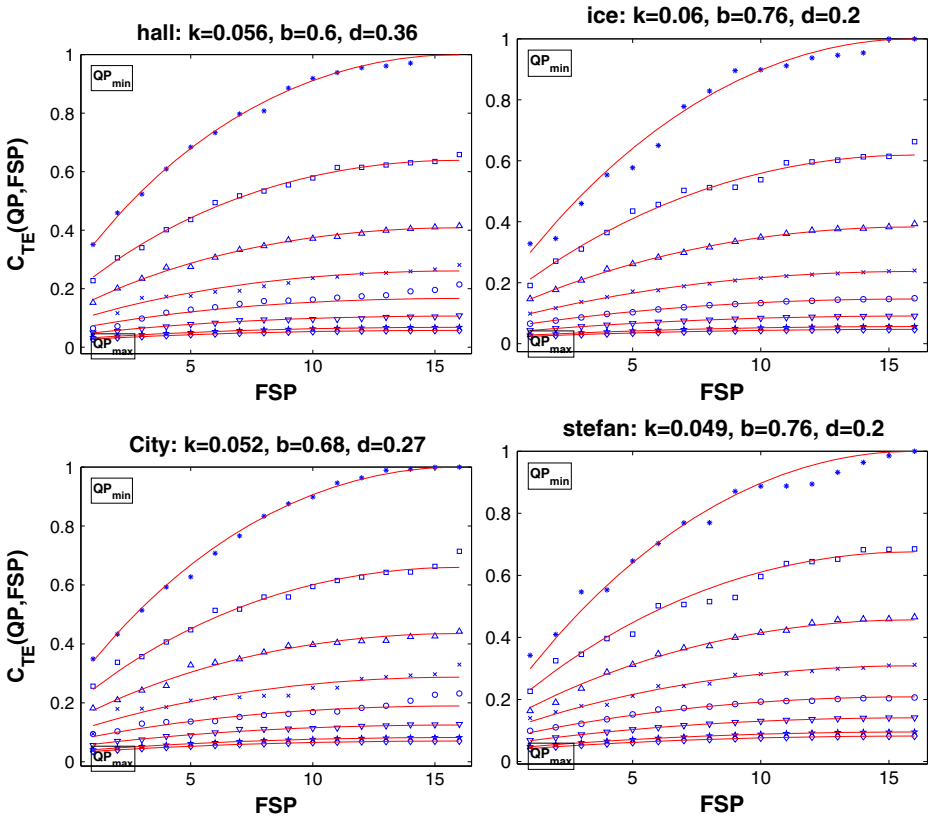


Fig. 6 Measured (points) and predicted (curves) normalized transmission energy $C_{TE}(QP, FSP)$. Curves are predicted values by the proposed model of (9)

Points in Fig. 7 report the measured $\alpha_{TE}(QP, FSP_{max})$, representing the effect of the quantization on the normalized consumed energy during the transmission process. As can be seen, $\alpha_{TE}(\cdot)$, similar to $\alpha_{EE}(\cdot)$, reaches its maximum at $QP=QP_{min}$ and its minimum at $QP=QP_{max}$. However, $\alpha_{TE}(\cdot)$ decreases faster than $\alpha_{EE}(\cdot)$ and reaches minimal values approaching 0. Therefore, we propose to model $\alpha_{TE}(\cdot)$ as an exponentially decreasing function of QP, as follows:

$$\alpha_{TE}(QP, FSP_{max}) = e^{-k \cdot QP} \tag{10}$$

where k is a content-dependent coefficient obtained by minimizing the RMSE between measured and predicted data. As shown in Fig. 7, the proposed model given by (10) can predict the reduction coefficients $\alpha_{TE}(QP, FSP_{max})$ accurately.

Points in Fig. 8 report the measured $\beta_{TE}(QP, FSP)$ that introduce energy consumption reduction when decreasing the FSP for a given QP. According to the measured data on this figure, the appropriate functional form for $\beta_{TE}(\cdot)$ has to meet the same conditions stated before, regarding the $\beta_{EE}(\cdot)$. For these reasons, we have:

$$\beta_{TE}(QP, FSP) = \frac{1 - b \cdot e^{(c \cdot QP_{norm})} \cdot e^{(d \cdot (FSP_{norm}^*)^f)}}{1 - b \cdot e^{(c \cdot QP_{norm})}} \tag{11}$$

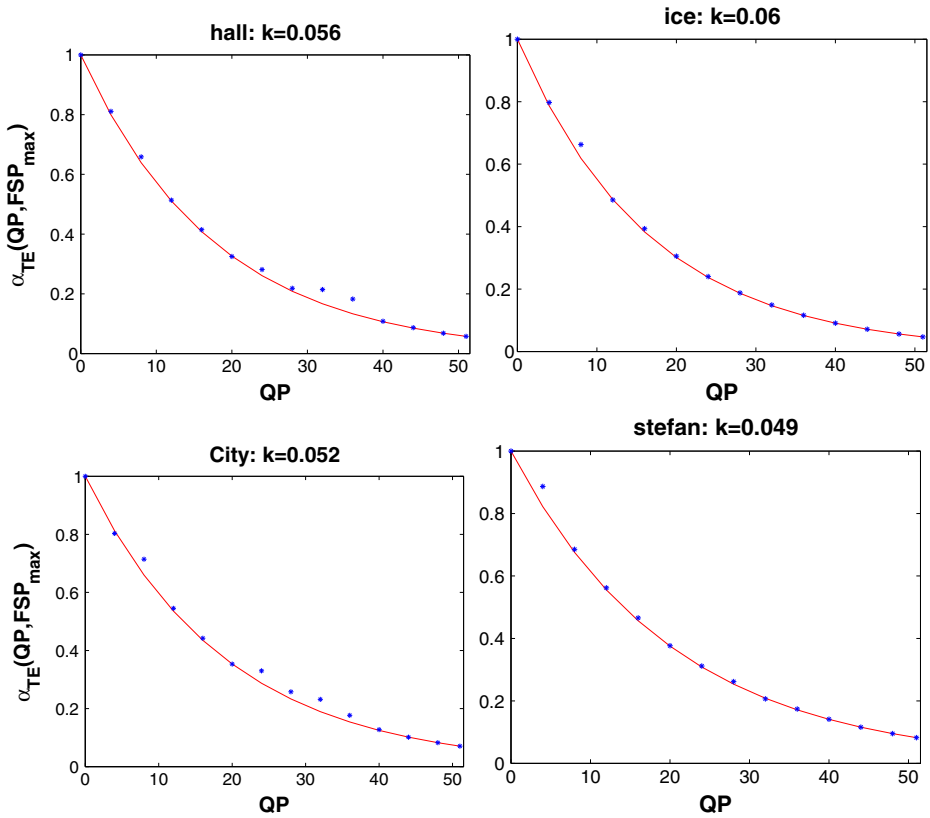


Fig. 7 Measured (points) and predicted (curves) $\alpha_{TE}(QP, FSP_{max})$. Curves are predicted values by the proposed model of (10)

According to several tests, we conclude that c can be fixed to -0.1 and f fixed to 2 . The parameters b and d are content dependent. From Fig. 8, we observe that the proposed model in (11) describes well the behaviour of $\beta_{TE}(\cdot)$.

The overall model for estimating the energy consumed during the video transmission process, considering the QP and the FSP, can be formulated by:

$$E_T(QP, FSP) = E_T(QP_{min}, FSP_{max}) \cdot \alpha_{TE}(QP, FSP_{max}) \cdot \beta_{TE}(QP, FSP) \quad (12)$$

As shown in Fig. 6, the proposed model in (9) describes well the coefficients $C_{TE}(QP, FSP)$, and hence the consumed energy during the transmission process with an average prediction error of 4% in terms of RMSE. The $E_T(QP_{min}, FSP_{max})$ depends of the radio module features, QP_{min} , FSP_{max} , and FR.

3.2 Video distortion modelling

In the following, we derive models for video distortion prediction mainly caused by both the encoding and the transmission processes. The end-to-end video distortion D_{Total} can be approximated by:

$$D_{Total} = D_E + D_T \quad (13)$$

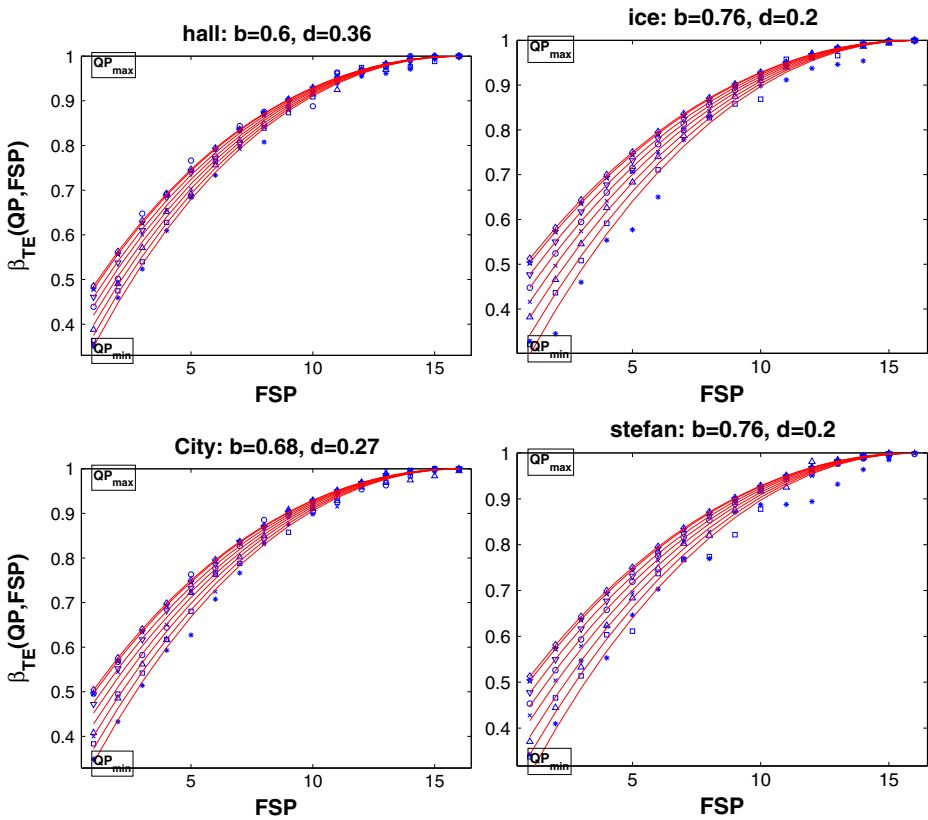


Fig. 8 Measured (points) and predicted (curves) $\beta_{TE}(QP, FSP)$. Curves are predicted values by the proposed model of (11)

where D_E and D_T refer to the distortion caused by encoding and transmitting the video stream respectively. Note that since intra-only coding is used, error propagation does not occur across frames. The distortion is expressed in terms of Mean Squared Error (MSE), and is given by:

$$MSE(X, Y) = \frac{\sum_{i=1, j=1}^{m, n} (X_{i, j} - Y_{i, j})^2}{m \cdot n} \tag{14}$$

where X is the frame of original video sequence, Y is the frame of compressed or received video sequence, m and n are frame width and height respectively.

3.2.1 Encoding distortion modelling

The distortion introduced by the encoding process can be expressed as a function of the QP and the FSP. Since we adopt a pixel level measure, the FR does not affect the video quality.

Points in Fig. 9 report the normalized measured encoding distortion. Unlike the encoding energy, increasing the QP and decreasing the FSP lead to an increase of the video distortion, which is an obvious behaviour. In fact, when the quantization and the FS are severe, more relevant information is lost. This affects directly the measured quality that drops

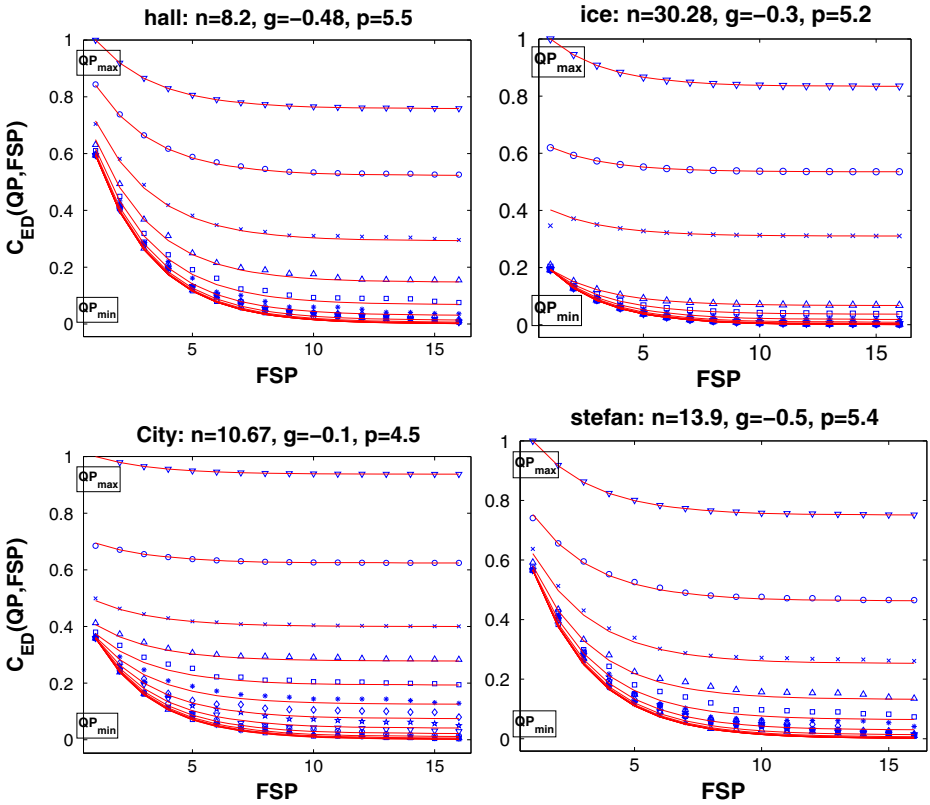


Fig. 9 Measured (points) and predicted (curves) $C_{ED}(QP, FSP)$. Curves are predicted values by the proposed model of (15)

dramatically. Note that the normalization is done by the maximal encoding distortion value, reached for (QP_{max}, FSP_{min}) .

Points in Fig. 9 can be seen as the behaviour of (QP, FSP) -dependent coefficients that reduce the maximal distortion reached for (QP_{max}, FSP_{min}) . For each QP, the maximal value of the coefficients is reached at $FSP = FSP_{min}$, then decreases toward the minimal value reached at $FSP = FSP_{max}$. These observations suggest that the coefficients $C_{ED}(\cdot)$ can be modelled by:

$$C_{ED}(QP, FSP) = \alpha_{ED}(QP, FSP_{min}) \cdot \beta_{ED}(QP, FSP) \tag{15}$$

Points in Fig. 10 report the measured $\alpha_{ED}(QP, FSP_{min})$, containing the effect of the quantization on the normalized distortion caused by the encoding process. Unlike $\alpha_{EE}(\cdot)$, $\alpha_{ED}(\cdot)$ reaches its maximum at $QP = QP_{max}$ and its minimum at $QP = QP_{min}$. In addition, $\alpha_{ED}(\cdot)$ observes a quasi-stationary phase until a given QP is reached, where $\alpha_{ED}(\cdot)$ sharply increases. Therefore, we propose to model $\alpha_{ED}(\cdot)$ by:

$$\alpha_{ED}(QP, FSP_{min}) = (1 - a_0) \cdot (QP_{norm})^n + a_0 \tag{16}$$

where a_0 is the normalized measured encoding distortion for $(QP, FSP) = (0, 1)$ and n is a content-dependent coefficient obtained by minimizing the RMSE between measured and

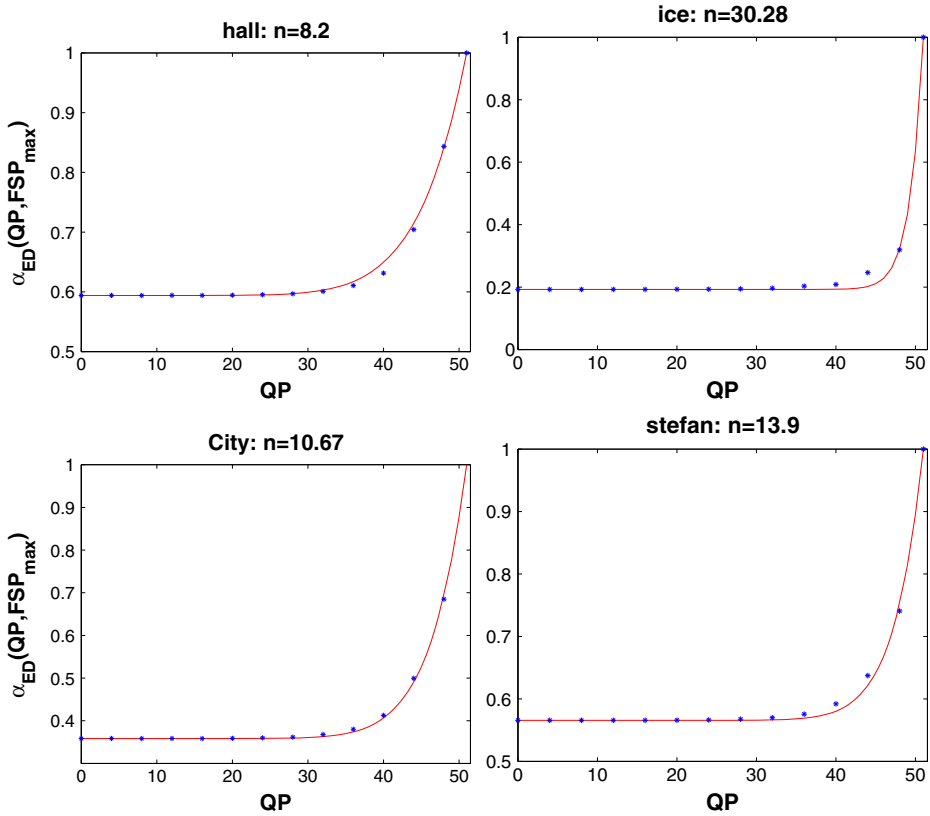


Fig. 10 Measured (points) and predicted (curves) $\alpha_{ED}(QP, FSP_{min})$. Curves are predicted values by the proposed model of (16)

predicted data. As shown in Fig. 10, the proposed model in (16) can predict the reduction coefficients $\alpha_{ED}(QP, FSP_{min})$ accurately.

Points in Fig. 11 report the measured $\beta_{ED}(QP, FSP)$ that introduce encoding distortion reduction when increasing the FSP for a given QP. According to the measured data on this figure, the appropriate functional form for $\beta_{ED}(\cdot)$ has to meet the following conditions. First, when the FS is applied with the minimal FSP, no encoding distortion reduction should be observed. Hence, the unique encoding distortion reduction should be the one due to the quantization, given by (16). Second, the more coefficients are retained in the FS phase, the less encoding distortion there is. Finally, when the quantization becomes more severe the effect of the FS becomes less important. Based on these observations, we propose the following functional form of $\beta_{ED}(\cdot)$:

$$\beta_{ED}(QP, FSP) = \frac{1 - g \cdot e^{(p \cdot (QP_{norm}^*)^2)} \cdot e^{-h \cdot FSP}}{1 - g \cdot e^{(p \cdot (QP_{norm}^*)^2)} \cdot e^{-h}} \tag{17}$$

where p, g and h are model parameters. According to several tests, h can be set to 0.41.

The overall model for estimating the video distortion due to the encoding process, considering the QP and the FSP, can be formulated by:

$$D_E(QP, FSP) = D_E(QP_{max}, FSP_{min}) \cdot \alpha_{ED}(QP, FSP_{min}) \cdot \beta_{ED}(QP, FSP) \tag{18}$$

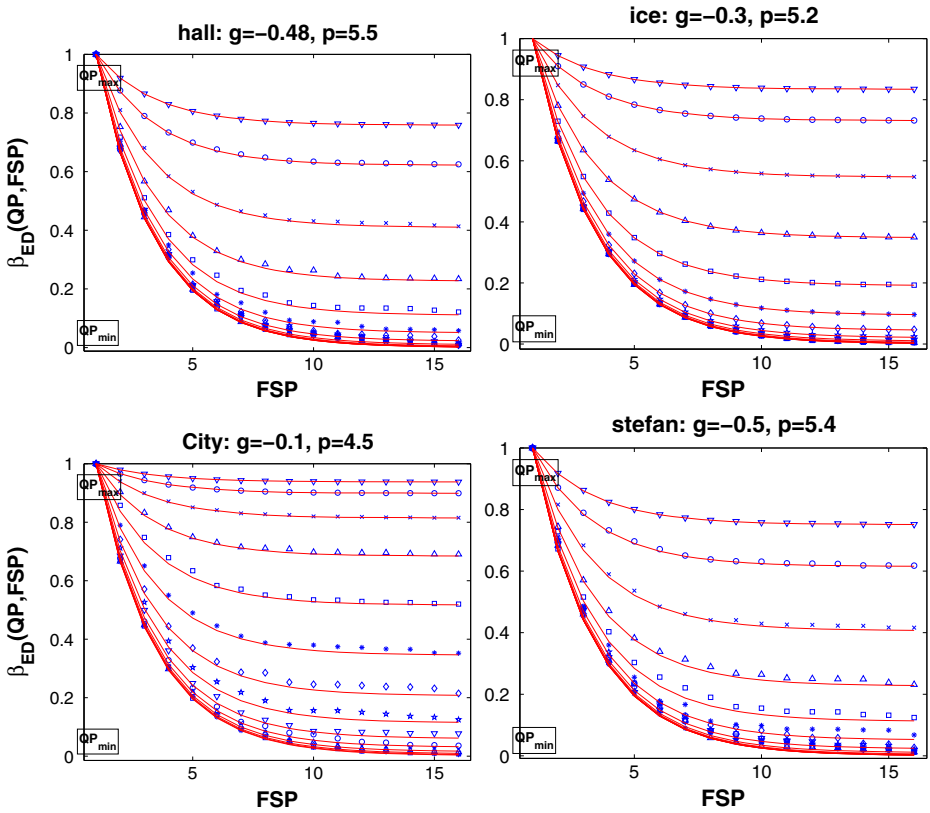


Fig. 11 Measured (points) and predicted (curves) $\beta_{ED}(QP, FSP)$. Curves are predicted values by the proposed model of (17)

As can be observed in Fig. 9, the proposed model in (15) can generate the coefficients $C_{ED}(QP, FSP)$ and thus estimate the encoding distortion with an average prediction error of 4% in terms of RMSE.

3.2.2 Transmission distortion modelling

Another factor affecting the video quality is the packet loss during the transmissions. The introduced distortion, D_T , is mainly dependent on the adopted encoding methodology (e.g. inter or intra) as well as the Packet Loss Rate (PLR). The PLRs are continuously provided by the network layer.

The introduced distortion due to the loss of a packet k can be estimated by the following widespread model [29]:

$$E[D_k] = (1 - \rho_k)E[D_{(R,k)}] + \rho_k E[D_{(L,k)}] \tag{19}$$

where $E[D_{(R,k)}]$ and $E[D_{(L,k)}]$ are the expected distortions when the packet k is either correctly received or lost, respectively. ρ_k is the loss probability of the packet k .

In the earlier video coding standards, intra pixels are coded directly using transform coding. In the H.264/AVC standard, intra-prediction is applied by first predicting the current

MB using previously coded neighbouring pixels in the current frame, then the prediction error is coded using transform coding. Therefore, when a packet belonging to a given slice is lost, the next MBs cannot be retrieved. For this purpose, in this paper, we perform partially or entire past frame slice copy loss concealment at the decoder, depending on the number of lost MBs.

Accordingly, when any packet is lost, the expected distortion $E[D_{(L,k)}]$ is caused by the frame loss concealment $E[D_{(Con,k)}]$. This last can be easily approximated at the encoder by the average difference between two successive frames, which is directly related to the inherent TI of the video sequence. When the packet is received, the distortion $E[D_{(R,k)}]$ is caused by the channel errors. In the present work, we assume that the packets are either lost or correctly received, hence $E[D_{(R,k)}] = 0$. Thereafter, (19) becomes:

$$E[D_k] = \rho_k E[D_{(Con,k)}] \quad (20)$$

Finally, the frame level distortion can be approximated by:

$$E[D^n] = \lambda \frac{\sum_{k=1}^N E[D_k]}{N} + (1 - \lambda)E[D^{n-1}] \quad (21)$$

where D^n refers to the distortion at frame n , and N is the number of the frame's packets. We introduce a temporal loss effect propagation factor λ that represents the introduced distortion caused by the concealment from previous concealed frames.

For model validation, since none of the few H.264/AVC real trace datasets available in the literature use the intra-only prediction mode, we adopt a two-state Markov chain model to generate bursty packet loss patterns, characterizing the WWSN [12]. This model is widely used to simulate the packet loss patterns in wireless sensor networks [7, 9, 12]. The average PLR is given by [12]:

$$PLR = \frac{P_{GB}}{P_{GB} + P_{BG}} \quad (22)$$

where P_{GB} and P_{BG} correspond to the transition probability from the Good to the Bad state, and from the Bad to the Good state, respectively [12].

Using this model, we generate 200 packet loss patterns for each PLR. The results are reported in Fig. 12. The proposed model approaches well the measured average distortions with an average prediction error of 9% in terms of RMSE.

4 ERMM routing protocol

This section focuses on the second component of the proposed framework, which acts at the network and MAC layers. Actually, we propose an energy-efficient routing protocol integrating the differentiation service paradigm, to efficiently handle the encoded video stream. The proposed routing protocol performs a hop-by-hop routing decision strategy that offers more scalability and flexibility to network topology and traffic changes. This protocol is actually inspired by the well known MMSPEED [10] protocol, that is considered as a reference routing protocol for the Wireless Multimedia Sensor Networks in general [3] and is used in many recent research works in both WSN and WWSN communities [22, 25]. Therefore, we aim in this section to revisit the MMSPEED protocol in order to improve its reliability and energy-efficiency, and introduce interactions with the application layer to suit the context of the WWSN. The proposed ERMM protocol operates under the same assumptions as the MMSPEED protocol [10].

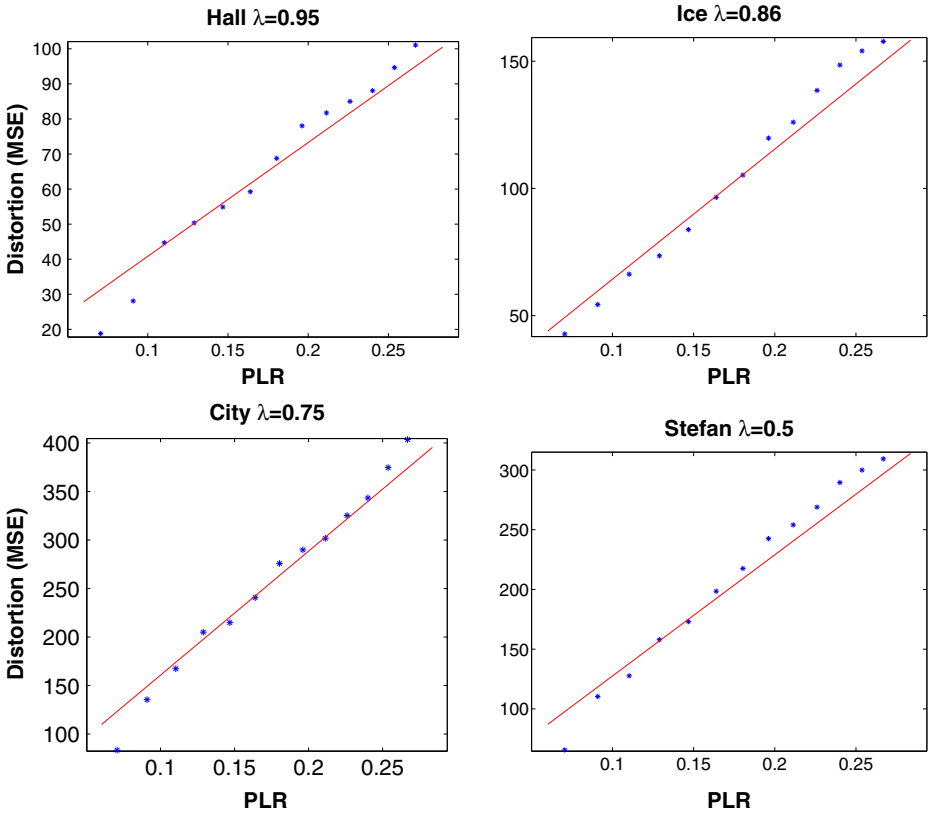


Fig. 12 Measured (points) and predicted (curves) D_T . Curves are predicted values by the proposed model of (21)

4.1 Reliability measure enhancement

In the following, we first define some terms that will further be used, then propose an enhanced reliability measure. Table 1 lists the relevant notations and variables for this algorithm. The first one is the Next-Hop Candidates (NHC) set, regrouping the nodes within the communication range and closer to the sink S than the current node i , and it is formulated as follows:

$$NHC_i(S) = \{j \in NS_i / dist(i, S) - dist(j, S) > 0\} \tag{23}$$

where $dist(i, S)$ and $dist(j, S)$ are Euclidean distances between i and S , j and S respectively, and NS_i is the neighbourhood set that contains the one-hop reachable neighbours. The reliability of a candidate node j is defined in [10] by:

$$RP_{ij}^S = (1 - PLR_{ij})(1 - PLR_{ij})^{\lceil \frac{dist(j,S)}{dist(i,j)} \rceil} \tag{24}$$

where PLR_{ij} is the average rate of the lost packets sent to j . Equation (24) expresses the probability that the packet reaches the node j , then the sink S after $\frac{dist(j,S)}{dist(i,j)}$ approximative hops separating j and S . Usually, the radio conditions vary along the path. However, in

Table 1 Notations and variables

NHC	Next-Hop candidates
NS_i	Neighbourhood set containing the one-hop reachable neighbours
RP_{ij}^S	Estimated reaching probability of a candidate node j
$ERM(j)$	Enhanced reliability measure of the node j
PLR_{ij}	Average rate of the lost packets sent to j
PLR_{ij}^l	PLR_{ij} regarding a given packet priority l
ABS^l	Available buffer size in the priority queue l
IBS^l	Initial buffer size of the priority queue l
TRP	Total reaching probability
NE	Neighbouring establishment period
RE	Remaining energy
$AvgRE$	Average remaining energy among the NHC
$NHCL_{Avg}$	Subsets of NHC containing nodes with $RE \leq AvgRE$
$NHCH_{Avg}$	Subset of NHC containing nodes with $RE > AvgRE$
ET	Energy threshold coefficient
PT	Percentage threshold of the desired reliability
LP	Low priority packet
HP	High priority packet

order to operate in a decentralized architecture, where the node has information about only its neighborhood, the PLR_{ij} is assumed to be constant over the needed hops to reach the sink S .

We propose here an enhancement of this reliability measure of (24) by considering the PLR_{ij}^l , referring to the achieved average loss rate regarding a given packet priority l , in addition to the Available Buffer Size (ABS) ABS^l of the node j . In fact, the ABS^l gives an information about the chances of the packet to be buffered in the next-hop node, independently of the success or not of the transmission. Therefore, in ERMM, the node’s reliability is expressed as follows:

$$RP_{ij}^S = (1 - ERM^l(j))(1 - ERM^l(j))^{\lfloor \frac{dist(j,S)}{dist(i,j)} \rfloor} \tag{25}$$

where $ERM(j)$ is the Enhanced Reliability Measure (ERM) of the node j for the packet priority l given by:

$$ERM^l(j) = PLR_{ij}^l * \frac{IBS^l - ABS^l}{IBS^l} \tag{26}$$

where IBS^l is the Initial Buffer Size (IBS) of the priority queue l .

The Total Reaching Probability (TRP) as presented in [10], allows to estimate the chances that a packet has to reach the sink, when it is relayed over multiple paths from a given source i and it is formulated as follows:

$$TRP_{new} = 1 - (1 - TRP_{old})(1 - RP_{ij}^S) \tag{27}$$

4.2 Neighbouring tables establishment and maintenance

The proposed routing protocol is stateless and operates on the basis of local information to dynamically construct paths toward the destination. For neighbouring tables establishment, the nodes broadcast packets during a given period NE (i.e. Neighbouring Establishment)

after network deployment. These packets contain their unique identifier (ID) and the three-dimensional coordinates referring to their respective location. Upon the reception of these messages, each node stores the ID and the coordinates of its neighbours. For neighboring tables maintenance, the nodes periodically broadcast control packets containing their respective ID, three-dimensional coordinates, RE and ABS in each level of the priority queue.

Upon the reception of each control packet, the node updates the entry of the concerned node and checks the validity of all of its entries. Thanks to the adopted IEEE 802.11e MAC layer [16], the nodes have also local statistics about the achieved PLR for each packet priority level PLR^l .

4.3 The routing algorithm

Figure 13 presents the flowchart of the proposed routing algorithm. The first step is the establishment of the NHC set, knowing the sink’s coordinates. Selecting candidates in such a manner allows the convergence of the packet toward the destination. Afterwards, the Average RE ($AvgRE$) among the NHC set is computed as follows:

$$AvgRE = \frac{\sum_{i=1}^N RE(i)}{N} \tag{28}$$

where N is the size of NHC set and the RE is obtained from the neighbouring table. Accordingly, two subsets are created on the basis of the so estimated $AvgRE$, namely the NHC_{HAvg} and the NHC_{LAvg} , which contain the nodes with RE higher and lower than the $AvgRE$ respectively. Next, for each node j of the NHC_{HAvg} subset, a score considering its reliability RP_{ij}^S and its distance from the current node i is calculated as follows:

$$score(j) = \tau RP_{ij}^S + (1 - \tau)dist(i, j)^{-1} \tag{29}$$

where RP_{ij}^S is obtained by (25) and τ is a control coefficient obtained through simulations. Considering at first stage the nodes with residual energy higher than the average

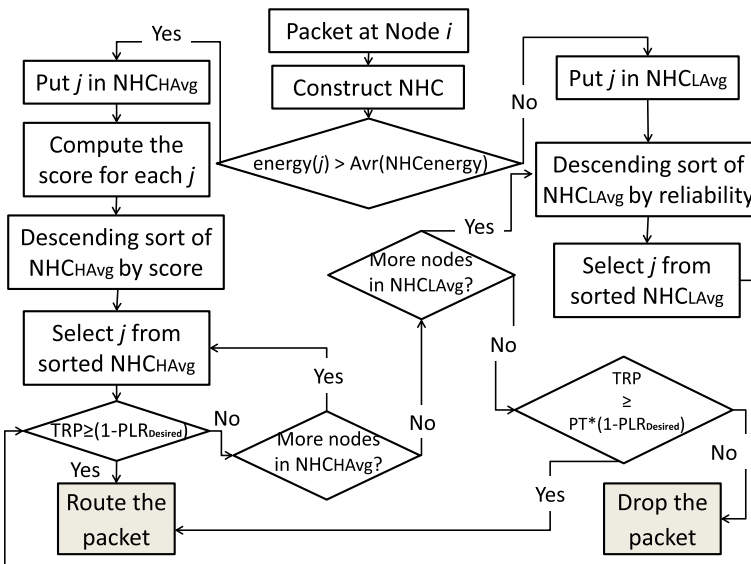


Fig. 13 ERMM routing algorithm’s flowchart

(i.e. the NHC_{HAvg} subset) aims to select the best neighbouring nodes in terms of energy, hence performing higher energy efficiency and load balancing. Taking the $dist(i, j)$ into account allows to consider neighbours close to the node i . Hence, shorter delays can be performed since the congestion level, represented by the ABS , is also taken into consideration. The nodes are then sorted according to these scores. The nodes are then selected according to their score and injected one by one in (27) until the TRP reaches the desired reliability that depends on the packet's priority. The packet is then sent to these nodes.

Whenever the nodes belonging to NHC_{HAvg} are not sufficient to overcome the specified PLR, the NHC_{LAvg} subset is used. In this situation, it is of greater importance to seek reliability than the timeliness. Therefore, the nodes belonging to NHC_{LAvg} subset are sorted according to their RP_{ij}^S , then injected one by one in (27) to enhance and reach the desired reliability. If all the nodes in NHC_{LAvg} are used and the TRP is still lower than the desired reliability (i.e. $1 - PLR_{Desired}$) by a given Percentage Threshold PT , the packet is discarded. Here, the node detects a congestion situation and notifies the source node in order to decrease the packet rate until the TRP becomes reachable for the following packets. In addition, the video source node decreases its video encoder's FR in accordance. When PT of the desired reliability is achieved, the packet is sent to the considered nodes.

4.4 Underlying MAC layer

Now that the next-hops are selected, the packet is relayed to the MAC layer to be further transmitted. If the MAC layer is already receiving or sending a packet, the current packet is enqueued according to its priority in the appropriate queue. The MAC layer continuously reports the ABS as well as the average PLR of each priority to the network layer.

In order to efficiently report the event to the sink during the rush mode, the event's data has to be prioritized across the network. This is achieved using the Enhanced Distributed Coordination Function (EDCF) provided by IEEE 802.11e standard. In the following, we assign the ROI in the rush mode to the HP, while the BKGD and the standby mode are assigned to the LP access categories. We denote the $PLR_{Current}^{HP}$ and $PLR_{Current}^{LP}$ the last stored PLR for the HP and LP packets respectively.

5 Framework overview

In this section, we introduce the proposed framework which concern is to reduce the overall energy consumption while meeting a user's desired end-to-end video quality. Actually, the proposed framework is designed with respect to the WWSN constraints (i.e. energy efficiency, video quality, reliability, bandwidth) and hence could be adopted for video surveillance applications using the WWSN, in energy-constrained environments such as the wildlife and the Smart Cities. We deem for a practical realization of the framework, a camera such as the LILIN IPD2220ES [20] could be used, since it adopts the H.264/AVC standard with ROI/BKGD separation and offers controllable multiple Frame Rates and bit rates (QP). Hence, given a video quality constraint and the observed node's remaining energy, the video encoder of our framework begins by estimating the actual transmission distortion, then it seeks the appropriate configuration in terms of the triplet (FR, QP, FSP), using the models of Section 3. Afterwards, the encoded video stream is routed with service differentiation and energy-efficiency towards the destination, using the proposed routing protocol in Section 4. The relevant notations and variables for the framework are listed in Table 1.

Specifically, during the standby mode, each frame is encoded with intra-only prediction following a given initial FR denoted by FR_{Stand} . Given the $PLR_{Current}^{LP}$, the node could approach the transmission video distortion using the model in (21). Consequently, the allowed encoding distortion is deduced. Afterwards, thanks to the encoding distortion model of (18), a solution set of QPs and their corresponding FSPs is constructed. Now, each couple is evaluated in terms of energy consumption, by using the models of (7) and (12). Afterwards, the node chooses the couple (QP,FSP) minimizing the energy consumption. On the basis of the estimated total energy consumption of the selected configuration, denoted QP^+ and FSP^+ , the node decides whether it could compress and transmit the stream or not. If the RE in the node is greater than the required energy by a given threshold ET , the node may encode the video stream and relay it to the network layer. In the opposite case, the node decreases the adopted FR_{Stand} and re-evaluate the required encoding and transmission energies in an iterative manner. When the lower permitted FR_{min} is reached, the node increases the selected QP^+ , then re-evaluates the encoding and transmission energies until the maximal permitted QP_{max} is attained. If the node is still incapable of performing the encoding, it turns-off the video module and remains as a relaying node.

In the rush mode, a higher initial FR_{Rush} is adopted. In Fig. 14, the flowchart explains the adopted procedure when two priorities coexist. This procedure is similar to the one proposed for the standby mode, with a minor adaptation. In fact, in the rush mode, the ROI and the BKGD are encoded using the same QP. By using the models of (18) and (21), in addition to the respective PLRs (i.e. $PLR_{Current}^{HP}$ and $PLR_{Current}^{LP}$), the node separately deduces the couples (QP,FSP) satisfying the distortion constraint of each region. Afterwards, the node selects the shared QPs and their respective FSPs. Finally, by using the models of (7) and (12), the node evaluates the energy consumption of each couple and selects those minimizing the total energy consumption. The selected configuration is denoted QP^+ , FSP_{ROI}^+ and FSP_{BKGD}^+ .

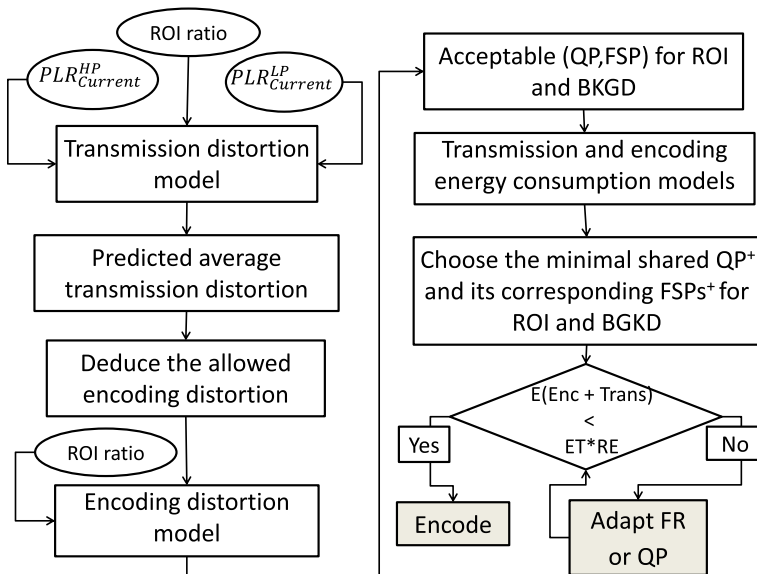


Fig. 14 Rush Mode flowchart

Once the video stream is encoded, the data is relayed to the network layer. The ERMM routing is then preformed as explained in Section 4.

6 Performance evaluations

In order to validate the proposed cross-layer framework, we conduct several tests using our modified JM18.4 implementation of the H.264/AVC standard. The video sequence Coast-guard is used, in QCIF spatial resolution at 30 fps, with (SI, TI)=(126,34). This QCIF format is composed of 99 MBs of 16x16 pixels. Since we choose to simulate the transmission of a fixed number of 3 MBs per packet, each frame is split into 33 packets.

The proposed cross-layer designed framework is composed of two major elements: (i) a video encoder performing dynamic encoding parameter selection via the prediction models of Section 3 and (ii) the ERMM routing protocol of Section 4. Therefore, we divide the current section into two subsections. In Section 6.1, we provide a view at the scale of one video source node by evaluating the impact of the dynamic selection of the video encoding parameters in terms of energy efficiency and video quality. Section 6.2 provides a view at the scale of a WWSN deployed to perform video surveillance of a given field, with the concern of delivering video streams to the sink, under a given quality constraint while extending the network lifetime as much as possible. Therefore, the overall framework is evaluated in terms of delay, reliability, energy-efficiency and video quality.

In this paper, we propose off-line models for the prediction of energy consumption and distortion, for which some parameters are content dependent. We propose as a first approach to select the models' parameters values on the basis of the (SI, TI) index. Consequently, the corresponding models parameters values of the nearest training video sequence in terms of SI and TI are selected, which corresponds in this case to News video sequence.

In Table 2, we report the different parameters used for the simulations.

6.1 Video source node evaluation

In this Subsection, we evaluate the behaviour of the video source node implementing our proposed encoding scheme. Therefore, we compare this last against a basic one. The basic encoding scheme consists in using the H.264/AVC intra-only encoder with FMO option to generate the ROI and the BKGD. The QP is set to 40 and the FR for standby and rush

Table 2 Simulation parameters

General settings			
Radio	40 m	ROI ratio	0.4
Bandwidth	250 kbps	Initial energy	100 J
Terrain	200x200 m ²	Packets per frame	33
Queue size	100 Pckts	NE	50 s
PR_{SM}	3 pps	PR_{RM}	6 pps
$PLR_{Desired}^{HP}$	0.25	$PLR_{Desired}^{LP}$	0.4
ET	0.04	PT	0.85
τ	0.8	(QP_{min}, QP_{max})	(14,40)
FR_{min}	1	(FSP_{min}, FSP_{max})	(3,12)

mode are set to 3.75 and 7.5 respectively. The initial FR_{stand} and FR_{Rush} for the proposed framework are also set to 3.75 and 7.5 respectively. The targeted maximal distortion, in terms of MSE, is fixed at 500. In this evaluation, we alternate three standby modes and three rush modes, as illustrated in Figs. 15, 16 and 17. In these figures, a round consists in collecting the current PLR regarding the packet priority $PLR_{Current}$, encoding the video signal, then transmitting it. For fair comparison, for each round the encoding schemes are subject to the same randomly generated $PLR_{Current}$.

Figure 15 shows the energy consumed by the video source node for encoding and transmitting the video signal. The proposed encoding scheme outperforms the basic one and extends the source node lifetime. In fact, this behaviour can be explained by the fact that the proposed encoding scheme adapts its configuration according to the available energy. Hence, when needed, the node decreases the encoder’s FR and progressively increases the QP in order to decrease the energy consumption and meet the desired quality. In addition, as can be seen in the end of our scheme’s Standby mode3, when the RE reaches the specified threshold ET, the node turns off its video module and remains as a relay node. In fact, this is done in order to maintain the network connectivity as long as possible. The observed average source node lifetime extension over these rounds is about 54%. This extension is of great interest specially to avoid as much as possible critical situations, such as node death during reporting an event (i.e. Fig. 15 in Rush mode2).

Figure 16 shows the average end-to-end video distortion after concealment for both schemes. While the basic scheme uses constants QP and FR values, the proposed scheme adapts its configuration according to the current PLR, desired end-to-end video distortion and the node’s RE. We observe that, for the proposed scheme, considering node’s RE during the encoding process allows the node to remain available for more video encodings and transmissions. Moreover, when the RE becomes lower, the node tends to increase the distortion, while meeting the targeted video quality, to save energy as illustrated in Fig. 16 during the Standby mode3.

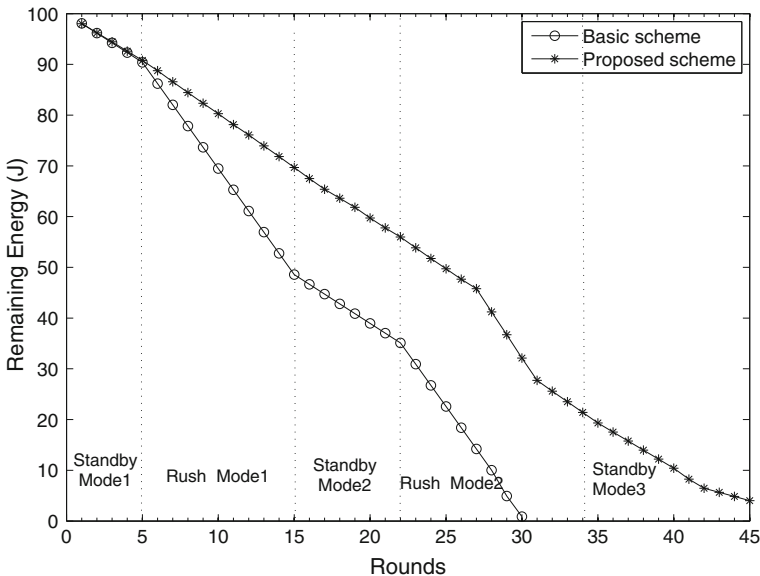


Fig. 15 Total energy consumption

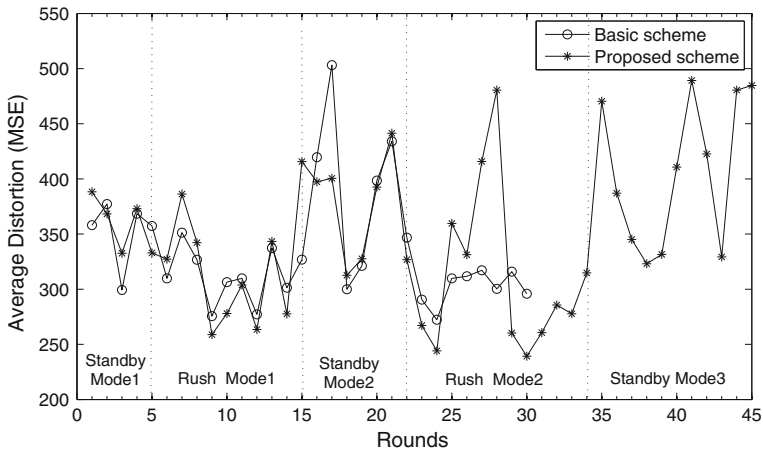


Fig. 16 Average end-to-end distortion after loss concealment

In Fig. 17, we report the adopted configuration by the proposed scheme at runtime. This encoding scheme selects, during the standby modes, higher QP^+ to save energy while providing acceptable video quality. During the rush modes, the FSP_{BKGD}^+ are lower than the FSP_{ROI}^+ to compensate the energy drained for the selected QP^+ in order to meet the desired quality.

6.2 Framework evaluation

In the following, we evaluate the proposed framework in terms of energy-efficiency, reliability, delay and end-to-end video quality. The used protocols and the proposed models are implemented in the JSIM [27] simulator. We use the energy model of a wireless AT&T Wavelan PCMCIA card implemented in the simulator.

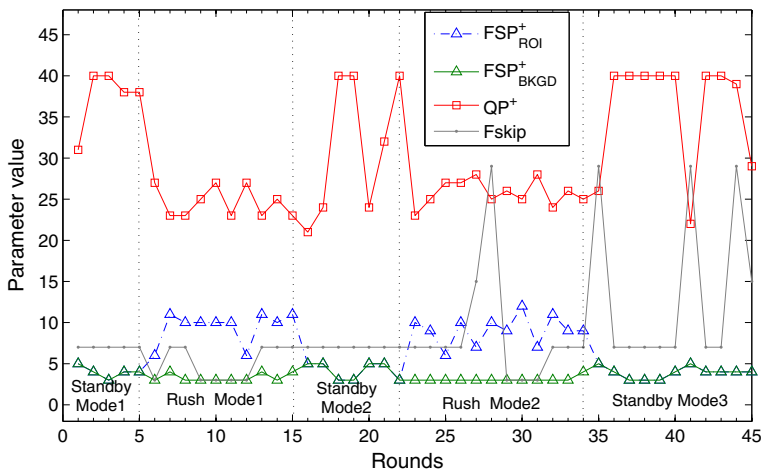


Fig. 17 Runtime parameter values

In Table 2, we report the different parameters used for the simulations. We consider 100 nodes randomly deployed in an area of interest where a given number of them are video nodes. The evaluation consists in a sequence of two standby modes separated by one rush mode, under a distortion constraint of 300. We study the impact of the number of video source nodes on the framework performance over 100 realizations. In each realization, the number of the video source nodes is fixed while their positions and the instants when they start sending are randomly chosen. We make sure that the sending periods overlap in order to evaluate the reaction of the framework in presence of a congestion situation. In addition, the proposed framework is compared against a basic one composed of the above described basic encoding scheme and the MMSPEED [10] routing protocol.

In Fig. 18, we present the average Packet Delivery Ratio (PDR) over the 100 realizations. Both frameworks are able to provide service differentiation to the HP and LP packets, thanks to the adopted 802.11e MAC layer. However, by considering the congestion level information estimated by the neighbour’s average loss rate in addition to its current ABS, the proposed ERMM protocol integrated in our framework provides higher PDR for HP and LP packets. Moreover, when needed, our framework decreases the packet generation rate (i.e. PR_{SM} or PR_{RM}) to relieve the network. Nevertheless, our approach provides the desired PDR up to 12 video source nodes for the LP stream, against 8 video source nodes for the basic framework. For the HP stream, the desired PDR of 75% is achieved up to 20 video source nodes, while it is provided only up to 10 video source nodes for the basic framework. In average, the PDR provided by our framework exceeds the one obtained by the basic framework by about 12% and 20% for the HP stream and LP stream respectively.

Figure 19 exposes the average remaining energy of the network over the 100 realizations. This energy decreases while the number of video source nodes increases, which is an obvious behaviour. As shown in Section 6.1, our framework saves a significant amount of energy at the video source node level. When combining our encoding scheme to the ERMM protocol, additional energy is saved during the routing process. Hence, the overall network lifetime is extended by about 15%.

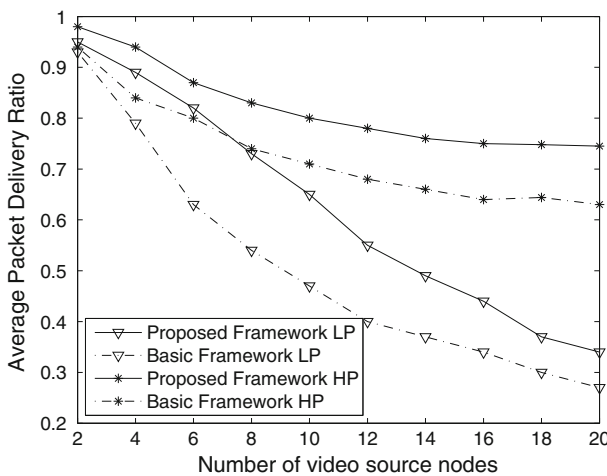


Fig. 18 Average packet delivery ratio while varying the number of video source nodes

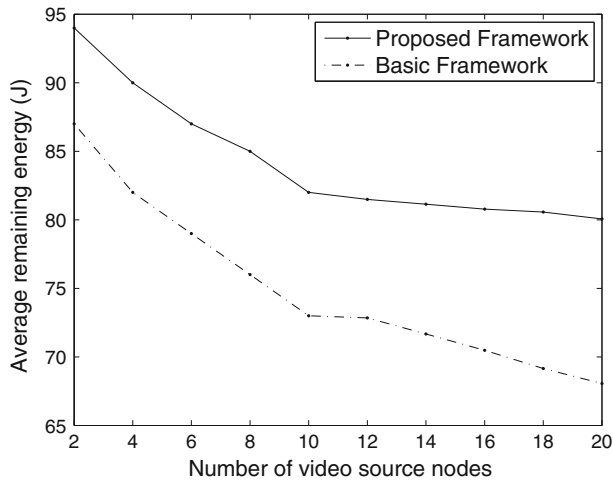


Fig. 19 Average remaining energy while varying the number of video source nodes

In Fig. 20, we report the obtained results in terms of delay over the 100 realizations. We observe that the basic framework offers a delay reduction of 3% for the HP stream and 15% for the LP stream. This behaviour is mainly caused by including the delay metric in the MMSPEED protocol, while the ERMM protocol aims to provide energy-efficiency and reliability. Nevertheless, the maximal delay of our framework does not exceed 1.2 s, which is bearable in a video-surveillance application.

The average end-to-end video distortion, over the 100 realizations, after concealment is given by Fig. 21. The proposed framework provides video quality with respect to the maximal distortion constraint, while the basic one offers this quality up to 10 video source nodes. In average, the video quality enhancement is about 35%, expressed by the decrease

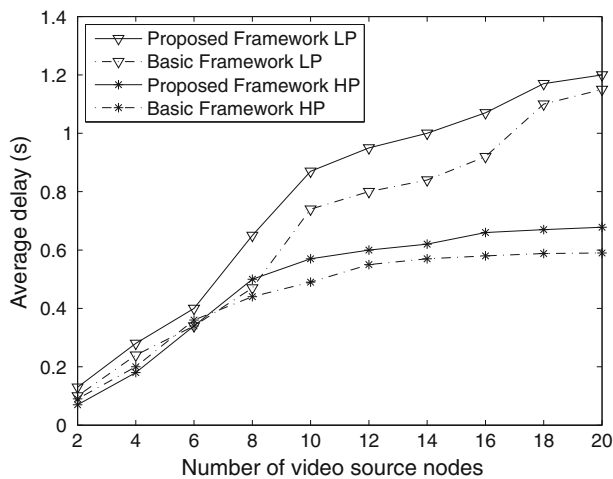


Fig. 20 Average delay while varying the number of video source nodes

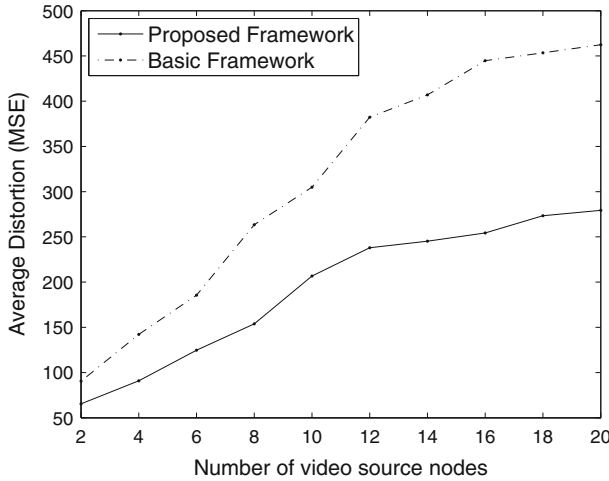


Fig. 21 Average end-to-end distortion after concealment while varying the number of video source nodes

of the measured MSE for the proposed framework. This decrease is achieved thanks to the enhanced reliability of the proposed ERM protocol, in addition to its interactions with the video encoder, that dynamically adapts its configuration to remain below the defined distortion threshold while extending the network lifetime.

Finally, for guidance, we report in Table 3 a comparative study between our proposed framework and some of the existing ones in the literature. In fact, performances comparison is not always obvious due to the fact that the simulation environment (i.e node’s initial energy budget, network density, number of video source nodes, available bandwidth, considered video distortions, considered consumed energies) varies from an approach to another.

Therefore, in Table 3, we present the average performance in terms of the considered energies, delay, PDR, number of video source nodes, the consumed energy for encoding and transmitting a video frame and the received video quality. As can be seen on Table 3, the proposed framework in average handles the tradeoff between network energy-efficiency and received video quality.

Table 3 Average performances comparison with existing frameworks. EE: Encoding Energy, TE: Transmission Energy, J/F: Joules per Frame

Ref.	Energy	Nbr of src	Delay (s)	PDR	Energy (J/F)	Quality (dB)
Our	EE+TE	2–20	0.2–0.92	0.96–0.55	13–52 (10^{-3})	30–23.7
[30]	EE+TE	10–30	N/A	N/A	N/A	28–19
[6]	EE+TE	1–12	0.1–0.25	N/A	1.5–3.5	28–18
[11]	TE	1–3	1–6	0.95–0.4	N/A	32
[26]	TE	2–10	0.1–1.7	N/A	0.15–1.1 (10^{-3})	33–22
[23]	TE	1	N/A	N/A	46 (10^{-3})	25
[1]	TE	1	N/A	N/A	33 (10^{-3})	40

7 Conclusion

This paper proposed a novel framework for network lifetime extension under an end-to-end video distortion constraint, for video delivery over a WWSN. The proposed framework relies on the cooperation of an energy-efficient encoding scheme, based on the H.264/AVC standard, and an Energy-efficient and Reliable Multipath Multi-priority routing protocol. First, we derived empirical models for energy consumption and distortion prediction during the encoding and the transmission phases. Second, we integrated them to our proposed video encoding scheme to dynamically adapt its configuration on the basis of the underlying network layer's feedbacks and the information about the remaining energy. Third, we proposed a routing protocol that considers the remaining energy, the packet loss rate and the available buffer size in nodes' multi-priority queues, in order to perform differentiated service, reliable and energy-efficient routing. Simulation results, in terms of energy, reliability, video quality and delay proved the efficiency of the proposed framework. In fact, our framework achieves 35% of end-to-end video quality enhancement, expressed in terms of MSE decrease, and 54% of video source node lifetime extension, resulting in increasing by 15% the overall network lifetime.

Acknowledgements This research was supported by the excellence fellowships program of the National Center for Scientific and Technical Research (CNRST) of Morocco (G01/003) and the Franco-Moroccan cooperation program in STICs for the research project "RECIF".

The authors would also like to thank the editorial office Christian Malan, the editor-in-chief, the guest editors as well as the anonymous reviewers for their efforts and valuable comments and suggestions that have led to improvements in this paper.

References

1. Aghdasi H, Abbaspour M, Moghadam M, Samei Y (2008) An energy-efficient and high-quality video transmission architecture in wireless video-based sensor networks. *Sensors* 4529–4559. doi:[10.3390/s8074529](https://doi.org/10.3390/s8074529)
2. Ahmad JJ, Khan HA, Khayam SA (2009) Energy efficient video compression for wireless sensor networks. In: 43rd annual conference on information sciences and systems (CISS). IEEE, pp 629–634. doi:[10.1109/CISS.2009.5054795](https://doi.org/10.1109/CISS.2009.5054795)
3. Akyildiz IF, Melodia T, Chowdury K (2007) Wireless multimedia sensor networks: A survey, vol 51
4. Alaoui-Fdili O, Coudoux F, Fakhri Y, Corlay P, Aboutajdine D (2013) Energy efficient adaptive video compression scheme for WWSNs. In: The European signal processing conference. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6811596>
5. Alaoui-Fdili O, Fakhri Y, Corlay P, Coudoux FX, Aboutajdine D (2014) Energy consumption analysis and modelling of a h.264/AVC intra-only based encoder dedicated to WWSNs. In: The IEEE international conference on image processing, pp 1189–1193. doi:[10.1109/ICIP.2014.7025237](https://doi.org/10.1109/ICIP.2014.7025237)
6. Boluk PS, Baydere S, Harmanci AE (2011) Perceptual quality-based image communication service framework for wireless sensor networks. *Wirel Commun Mob Comput* 14(1):1–18. doi:[10.1002/wcm.1218](https://doi.org/10.1002/wcm.1218)
7. Costa D, Guedes L, Vasques F, Portugal P (2014) Relevance-based partial reliability in wireless sensor networks. *EURASIP J Wirel Commun Netw* 2014(1):142. doi:[10.1186/1687-1499-2014-142](https://doi.org/10.1186/1687-1499-2014-142)
8. Costa DG, Guedes LA (2011) A survey on multimedia-based cross-layer optimization in visual sensor networks. *Sensors* 11(5):5439–5468. doi:[10.3390/s110505439](https://doi.org/10.3390/s110505439)
9. Costa DG, Guedes LA (2012) A discrete wavelet transform (dwt)-based energy-efficient selective retransmission mechanism for wireless image sensor networks. *Journal of Sensor and Actuator Networks* 1(1):3–35. doi:[10.3390/jsan1010003](https://doi.org/10.3390/jsan1010003)
10. Felemban E, Chang-Gun L, Ekici E (2006) MMSPEED: multipath multi-speed protocol for QoS guarantee of reliability and timeliness in wireless sensor networks. *IEEE Trans Mob Comput* 5(6):738–754. doi:[10.1109/TMC.2006.79](https://doi.org/10.1109/TMC.2006.79)

11. Felemban E, Sheikh AA, Manzoor MA (2014) Improving response time in time critical visual sensor network applications. *Ad Hoc Netw* 23:65–79. doi:[10.1016/j.adhoc.2014.06.003](https://doi.org/10.1016/j.adhoc.2014.06.003)
12. He Y, Lee I, Guan L (2009) Distributed algorithms for network lifetime maximization in wireless visual sensor networks. *IEEE Trans Circuits Syst Video Technol* 19(5):704–718. doi:[10.1109/TCSVT.2009.2017411](https://doi.org/10.1109/TCSVT.2009.2017411)
13. He Z, Eggert J, Cheng W, Zhao X, Millsbaugh J, Moll R, Beringer J, Sartwell J (2008) Energy-aware portable video communication system design for wildlife activity monitoring. *IEEE Circuits Syst Mag* 8(2):25–37. doi:[10.1109/MCAS.2008.923007](https://doi.org/10.1109/MCAS.2008.923007)
14. He Z, Liang Y, Chen L, Ahmad I, Wu D (2005) Power-rate-distortion analysis for wireless video communication under energy constraints. *IEEE Trans Circuits Syst Video Technol* 15(5):645–658. doi:[10.1109/TCSVT.2005.846433](https://doi.org/10.1109/TCSVT.2005.846433)
15. Heinzelman W, Chandrakasan A, Balakrishnan H (2000) Energy-efficient communication protocol for wireless microsensor networks. In: The annual Hawaii international conference on system sciences. doi:[10.1109/HICSS.2000.926982](https://doi.org/10.1109/HICSS.2000.926982)
16. IEEE std 802.11e/D3.3 draft supplement to IEEE standard for telecommunications and information exchange between systems – LAN/MAN specific requirements. part 11: Wireless lan medium access control (MAC) and physical layer (PHY) (2002). http://www.ece.virginia.edu/~mv/edu/el604/references/P802.11E-D1_3.pdf
17. Intel core2 duo processor. <http://support.intel.co.jp/pressroom/kits/core2duo/>
18. ITU-T RECOMMENDATION P (1999) Subjective video quality assessment methods for multimedia applications
19. Lambert P, Neve WD, Dhondt Y, de Walle RV (2006) Flexible macroblock ordering in H.264/AVC. *J Vis Commun Image Represent* 17(2):358–375. doi:[10.1016/j.jvcir.2005.05.008](https://doi.org/10.1016/j.jvcir.2005.05.008) Introduction: Special Issue on emerging H.264/AVC video coding standard
20. LILIN IPD2220ES. http://www.tsys.com/pdf/RFID/IPD2220ES-EN_TTS_v4.pdf
21. Lu X, Wang Y, Erkip E (2002) Power efficient H.263 video transmission over wireless channels. In: The international conference on image processing, vol 1, pp 533–536. doi:[10.1109/ICIP.2002.1038078](https://doi.org/10.1109/ICIP.2002.1038078)
22. Macit M, Gungor VC, Tuna G (2014) Comparison of QoS-aware single-path vs. multi-path routing protocols for image transmission in wireless multimedia sensor networks. *Ad Hoc Netw* 19:132–141. doi:[10.1016/j.adhoc.2014.02.008](https://doi.org/10.1016/j.adhoc.2014.02.008)
23. Politis I, Tsagkaropoulos M, Dagiuklas T, Kotsopoulos S (2008) Power efficient video multipath transmission over wireless multimedia sensor networks. *Mob Netw Appl* :274–284. doi:[10.1007/s11036-008-0061-5](https://doi.org/10.1007/s11036-008-0061-5)
24. Rui D, Pu W, Akyildiz I (2012) Correlation-aware QoS routing with differential coding for wireless video sensor networks. *IEEE Trans Multimed* 14(5):1469–1479. doi:[10.1109/TMM.2012.2194992](https://doi.org/10.1109/TMM.2012.2194992)
25. Sahin D, Gungor VC, Kocak T, Tuna G (2014) Quality-of-service differentiation in single-path and multi-path routing for wireless sensor network-based smart grid applications. *Ad Hoc Netw* 22:43–60. doi:[10.1016/j.adhoc.2014.05.005](https://doi.org/10.1016/j.adhoc.2014.05.005)
26. Shah G, Liang W, Akan O (2012) Cross-layer framework for QoS support in wireless multimedia sensor networks. *IEEE Trans Multimed* 14(5):1442–1455. doi:[10.1109/TMM.2012.2196510](https://doi.org/10.1109/TMM.2012.2196510)
27. Sobeih A, Hou J, Kung L, Li N, Zhang H, Chen W, Tyan H, Lim H (2006) J-sim: a simulation and emulation environment for wireless sensor networks. *IEEE Wirel Commun* 13(4):104–119. doi:[10.1109/MWC.2006.1678171](https://doi.org/10.1109/MWC.2006.1678171)
28. Soro S, Heinzelman W (2009) A survey of visual sensor networks *Adv Multimed*. doi:[10.1155/2009/640386](https://doi.org/10.1155/2009/640386)
29. Zhai F, Eisenberg Y, Pappas T, Berry R, Katsaggelos A (2006) Rate-distortion optimized hybrid error control for real-time packetized video transmission. *IEEE Trans Image Process* 15(1):40–53. doi:[10.1109/TIP.2005.860353](https://doi.org/10.1109/TIP.2005.860353)
30. Zou J, Xiong H, Li C, Zhang R, He Z (2011) Lifetime and distortion optimization with joint source/channel rate adaptation and network coding-based error control in wireless video sensor networks. *IEEE Trans Veh Technol* 60(3):1182–1194. doi:[10.1109/TVT.2011.2111425](https://doi.org/10.1109/TVT.2011.2111425)



Othmane Alaoui-Fdili received the M.S degree in 2010 in Telecommunications and Computer Sciences from Mohammed V University (UMV), Rabat, Morocco. He received his PhD from both LRIT-CNRST (URAC 29) laboratory of UMV, and the Department of Opto-Acosuto-Electronics of the Institute of Electronics, Microelectronics and Nanotechnologies (UMR 8520) of the University of Valenciennes, France. He joined the Ecole Supérieure de Technologie de Safi, Université Cadi Ayyad, as an Associate Professor in June 2017. His research interests include routing, video compression, energy efficiency and cross-layer design for wireless sensor networks.



François-Xavier Coudoux received the M.S. and Ph.D. degrees in electrical engineering from the University of Valenciennes, Valenciennes, France, in 1991 and 1994, respectively. Since 2004, he has been a Professor in the Department of Opto-Acousto-Electronics of the Institute of Electronics, Microelectronics, and Nanotechnologies, Valenciennes, France (UMR 8520). His research interests include telecommunications, multimedia delivery over wired and wireless networks, image quality, and adaptive video processing.



Youssef Fakhri received his Bachelor's Degree (License *és* Sciences) in Electronic Physics in 2001 and his Master's Degree (DESA) in Computer and Telecommunication from the Faculty of Sciences, University Mohammed V, Rabat, Morocco, in 2003 where he developed his Master's Project at the ICI company, Morocco. He received a PhD in 2007 from the University Mohammed V, Rabat, Morocco in collaboration with the Polytechnic University of Catalonia (UPC), Spain. He joined the Faculty of Sciences of Kénitra, Department of Computer Science and Mathematics, Ibn Tofail University, Morocco, as an Associate Professor in Mars 2009, he is the head of Networks and Telecommunications research team at LARIT, Associate Researcher at the Laboratory for Research in Computing and Telecommunications (LRIT) in the Faculty of Sciences of Rabat, and Member of Pole of Competences STIC Morocco. His current research interests include QoS in wireless communication, WSN, wireless.



Patrick Corlay received the Ph.D. degree in electrical engineering from the University of Valenciennes, Valenciennes, France, in 1994. He is an Assistant Professor in the Department of Opto- Acousto-Electronics of the Institute of Electronics, Microelectronics, and Nanotechnologies, Valenciennes, France (UMR 8520). His current research interests are in telecommunications, digital transmission over wired networks (power line, asynchronous digital subscriber line), channel estimation, and robust MPEG2, and H.264 AVC/SVC video transmission.



Driss Aboutajdine received the Doctorat de 3me Cycle and the Doctorat d'Etat-esSciences degrees in signal processing from the Mohammed V University, Rabat, Morocco, in 1980 and 1985, respectively. He joined Mohammed V University, Rabat, Morocco, in 1978, first as an assistant professor, then as an associate professor in 1985, and full Professor since 1990. Over 35 years, he developed teaching and research activities covering various topics of signal and image processing, wireless communication and pattern recognition which allow him to advise more than 50 Ph.D. theses and publish over 200 journal papers and conference communications. He succeeded to found and manage since 1993 and 2001 respectively the LRIT Laboratory and the Centre of Excellence of Information & Communication Technology (pole de competences STIC) which gathers more than 30 research laboratories from all Moroccan universities and including industrial partners as well. Prof. Aboutajdine has organized and chaired several international conferences and workshops. He was elected member of the Moroccan Hassan II Academy of Science and technology on 2006 and fellow of the TWAS academy of sciences on 2007. As an IEEE Senior Member, he cofounded the IEEE Morocco Section in 2005 and he is chairing the Signal Processing chapter he founded in December 2010. He is currently the director of the CNRST research center in Morocco, Rabat.