Energy Consumption Of Visual Sensor Networks: Impact Of Spatio-Temporal Coverage Based On Single-Hop Topologies

Alessandro Redondi[1], Dujdow Buranapanichkit[2], Matteo Cesana[1], Marco Tagliasacchi[1] and Yiannis Andreopoulos[3]

[1]Electronic, Information and Bioengineering Department, Politecnico di Milano, Italy
`{redondi,cesana,tagliasa}@elet.polimi.it`
[2]Department of Electrical Engineering, Prince of Songkla University, Thailand
`dujdow.b@psu.ac.th`
[3]Electronic and Electrical Engineering Department, University College London, UK
`i.andreopoluos@ucl.ac.uk`

Abstract. Wireless visual sensor networks (VSNs) are expected to play a major role in future IEEE 802.15.4 personal area networks (PAN) under recently-established collision-free medium access control (MAC) protocols. In such environments, the trade-off between the number of camera sensors to deploy (spatial coverage) and the frame rate to use for each camera sensor (temporal coverage) plays a major role in the VSN energy consumption. In this paper, we address this aspect for *single-hop* VSNs, i.e. networks comprising independent and identical wireless visual sensor nodes connected to a collection node via a star topology. We derive analytic results for the energy-optimal spatio-temporal coverage parameters of such VSNs under *a-priori* known bounds for the minimum frame rate per sensor and the minimum and maximum possible number of nodes to deploy. Our results are parametric to the probability density function characterizing the data-production rate per node and the energy consumption parameters of the system of interest. Experimental results using TelosB motes under: a collision-free transmission protocol, the IEEE 802.15.4 PAN physical layer (CC2420 transceiver) and Monte-Carlo–generated data sets, reveal that our analytic results are within 7% of the energy consumption measurements for a wide range of settings. In addition, results obtained via a multimedia subsystem performing visual feature extraction in video frames show that the optimal spatio-temporal settings derived by the proposed framework allow for up to 48% of reduction of energy consumption in comparison to ad-hoc settings. As such, our analytic modeling is useful for early-stage studies of possible VSN deployments under collision-free MAC protocols prior to costly and time-consuming experiments in the field.

1 Introduction

The integration of low-power wireless networking technologies such as IEEE 802.15.4-enabled transceivers [1] with inexpensive camera hardware [2] has enabled the development of the so-called *visual sensor networks* (VSNs). VSNs

can be thought of as networks of wireless devices capable of sensing multimedia content [3], such as still images and video, audio, depth maps, etc. Via the recent provisioning of an all-IPv6 network layer under 6LoWPAN and the emergence of collision-free low-power medium access control (MAC) protocols, such as the time slotted channel hopping (TSCH) of IEEE 802.15.4e-2012 [4], VSNs are expected to play a major role in the evolution of the Internet-of-Things (IoT) paradigm [5].

1.1   Review of Visual Sensor Networks

An increasing number of VSN solutions were proposed recently with a focus on new transmission protocols allowing for high-bandwidth collision-free communications [6,4] and in-network visual processing techniques.Most of the proposed VSN solutions can be abstracted as two tightly-coupled subsystems: a multimedia processor board and a low-power radio subsystem [2], interconnected via a push model. A cluster of such identical nodes can be organized into a VSN comprising a star topology that can operate in collision-free steady-state mode as illustrated in the example of Figure 1, with the consumption rate of each node being s bits for each interval of T seconds that the VSN remains active, or $\frac{s}{T}$ bits-per-second (bps). Within each node, the multimedia subsystem is responsible for acquiring images, processing them and pushing the processed data to the radio subsystem. For example, in the most typical application scenario for VSNs, the multimedia subsystem would acquire each image, compress it into a JPEG bitstream (e.g., using an MJPEG codec) and push the JPEG bitstream to the radio subsystem [2]. The latter transmits the processed data stream to the low-power border router (LPBR) [5], and eventually to a remote destination, which, under the IoT paradigm, could be any IPv6 Internet address.

Multimedia processing subsystem: The frame rate under which each VSN camera is operating, i.e. each node's temporal coverage, is controlling the frequency of the push operations. At the same time, the multimedia processing task itself (e.g., JPEG compression) controls the volume of bits pushed to the radio subsystem within each frame's duration.

Communications subsystem: The number of sensors in the (single-hop) star topology, i.e. the VSN's spatial coverage (Figure 1), controls the bandwidth available to each sensor (i.e. its average transmission rate) under a collision-free MAC protocol [6,7,4]. Thus, there is a fundamental tradeoff between the spatial and temporal coverage in a network: high frame rate leads to high bandwidth requirement per transmitter, which in turn decreases the number of sensors that can be accommodated with the same LPBR. Conversely, dense spatial coverage via the use of a large number of visual sensors decreases the available bandwidth per sensor and this in turn reduces the achievable frame rate per sensor in order to maintain tight coupling.
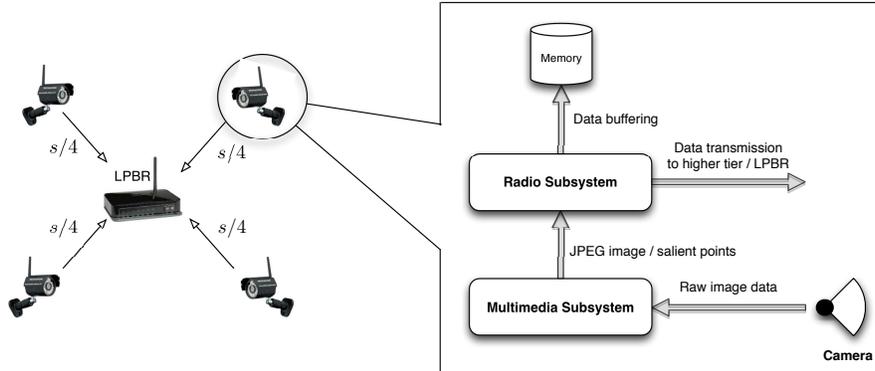
Fig. 1. Single-hop star topology in a visual sensor network, where every visual sensor (video camera) has its own spatial coverage, with s indicating the bits consumed by each node within each active interval of T seconds. Each camera node comprises two subsystems, which are illustrated in the figure expansion. If required, each node can buffer parts of its data stream for later transmission.

Overall system perspective – energy consumption: Like traditional wireless sensor networks, VSN nodes are battery operated. Hence, energy consumption plays a crucial role in the design of a VSN, especially for those applications where a VSN is required to operate for days or even weeks without external power supply. In the last few years, several works have addressed the problem of lifetime maximization in VSNs: depending on the research area, solutions are available for energy-aware protocols and cross-layer optimization [4], application tradeoffs [5] and deployment strategies [2]. While existing work addresses transmission, scheduling and protocol design aiming for energy efficiency, it does not consider the impact of the spatio-temporal coverage in the energy consumption of VSNs. This is precisely the focus of this paper.

1.2 Contribution

In this paper, we derive analytic results concerning energy-aware VSN deployments under the push model of Figure 1. Specifically, we are interested in the link of the aforementioned spatio-temporal tradeoff with the incurred energy consumption under well known probability density functions modeling the pushed data volume of image and video applications, such as intra/inter-frame video coding and visual features extraction and transmission. We focus on the widely used case of single-hop VSNs comprising identical sensors connected to the LPBR via a star topology and we derive an analytic model that captures the expected energy consumption in function of: *(i)* the number of visual sensors deployed, *(ii)* the frame rate used by each camera sensor and *(iii)* the statistical characterization of the bitstream data volume produced by each sensor after on-board image analysis or compression. The extrema of the derived energy consumption

function are then analytically derived in order to provide closed-form expressions for the minimum energy consumption of each case under consideration. The derived analytic results are experimentally validated with a VSN performing visual feature extraction and transmission to the LPBR.

2 Utilized System Model and Its Expected Energy Consumption

In the first four subsections we introduce the components of the utilized system model and the corresponding nomenclature. The key notations are summarized in Table 1, along with the practical settings used in the experiments of the paper.

2.1 Communication and System Infrastructure

Consider a wireless visual sensor network comprising a star topology. The network consists of n independent and identical camera nodes that process visual data and transmit multimedia streams to the LPBR. The MAC layer of the network is operating under a collision-free time-division (or time-frequency division) multiple access protocol [6,7,4], so that simultaneous transmissions (self-inflicted collisions) are avoided. Let $\frac{s}{T}$ bps be the average consumption rate of the LPBR over the VSN active interval of T seconds.

Within each node, the multimedia and radio subsystems work in parallel: while the multimedia system acquires and processes data corresponding to the current video frame, the radio subsystem transmits the multimedia stream stemming from the processing of the previous video frame(s). Examples of VSN applications that fit the communications model illustrated in Figure 1 are: multi-camera JPEG compression and transmission of video bitstreams [8], multi-sensor visual features extraction [9] and transmission, multi-camera compression and transmission for object recognition, etc.

2.2 Active Time Interval and Delay Tolerance

Given that applications based on visual sensor networks are subject to severe bandwidth requirements, it may not be possible to transmit the entirety of each multimedia stream within the transmission opportunities corresponding to the duration of one video frame. In such a case, buffering to on-board memory is required. This means that the application must tolerate certain delay until all multimedia streams are received by the LPBR. This delay is controlled by the chosen value of T and can be tuned to fit the constraints imposed by each deployment scenario.

After T seconds, each sensor stops gathering new data, completes the transmission of any data that may exist in its buffer and goes into suspension mode until the occurrence of the next active time interval. For example, setting $T = 5$ s indicates that the sensors are actively gathering and processing visual data for five seconds, complete any remaining data transmissions after that, and then

suspend their activity until being reactivated. The VSN activation can either be event-driven (e.g., when activity or motion is detected) or periodic, with a certain duty cycle [2,4].

2.3 Spatio-Temporal Coverage and Statistical Characterization of Data Transmission Volume per VSN Node

We consider that the VSN is established under the following two application constraints:

- *spatial coverage bounds*; the number of deployed nodes, n, is upper- and lower-bounded, i.e., $N_{\min} \leq n \leq N_{\max}$
- *temporal coverage lower bound*; the total frame acquisitions, k, within a pre-defined time interval, T, is lower-bounded, i.e., $k \geq K_{\min}$

The bounds of the spatio-temporal coverage stem from application specifics, such as: the cost of installing and maintaining visual sensors, the minimum and maximum spatial coverage required for the area to be monitored, and the minimum frame rate that allows for visual data gathering and analysis with sufficient temporal resolution for the application.

Since the multimedia subsystem of each visual sensor produces varying amounts of data depending on the monitored events and the specifics of the visual analysis and processing under consideration, the data stream volume produced by each camera in such multimedia applications is a non-deterministic quantity. We thus model the data volume produced when each camera processes k frames via a random variable (RV) $\mathcal{X}_k$, with marginal probability density function (PDF) $P(\chi_k)$, $\mathcal{X}_k \backsim P(\chi_k)$.

2.4 Energy Consumption Penalties

Following the push model illustrated in Figure 1, each node performs the following operations during the active time interval T:

1. *Acquisition:* A low-power camera sensor acquires k frames, each incurring a Joule (J) of energy expenditure. Hence, the energy consumed during T s is ka J.
2. *Processing and transmission:* Each captured video frame is processed with a CPU-intensive algorithm, realized by the multimedia subsystem. Each frame processing produces, on average, r bits (b) for transmission. These bits are pushed to the radio subsystem, which in turn transmits them to the base station. Let g J be the average energy required for processing and producing one bit of information and j the average energy required to transmit it to the LPBR. The average energy consumed for processing and transmission within the active interval is hence $(g + j) \int_0^\infty \chi_k P(\chi_k) d\chi = (g + j)\mathbb{E}[\mathcal{X}_k]$ J, with $\mathbb{E}[\mathcal{X}_k]$ bits comprising the statistical expectation of the data volume corresponding to k frames.

Table 1. Visual sensor energy and bitrate parameters, including settings used in this paper.

Parameter	Description	Unit	Value
Radio Subsystem (TelosB)			
n, $N_{\min}$, $N_{\max}$	Number of nodes, min/max setting	–	$N_{\min} = 2$, $N_{\max} = 16$
$\frac{s}{T}$	Data consumption rate	kbps	144
j	Transmission energy per bit	J/b	2.197×10^{-7}
b	Beaconing/idling energy per bit	J/b	1.902×10^{-7}
p	Buffering energy per bit	J/b	2.861×10^{-7}
Multimedia Subsystem ASIC [10][2]			
k, $K_{\min}$	Frames captured within T s, min. setting	–	$K_{\min} = 2T$
a	Acquisition energy per frame	J	4.000×10^{-3}
g	Processing energy per bit (visual features)	J/b	1.907×10^{-8}

3. *Buffering and Idling:* As shown in Figure 1, the sensor network consists of n sensor nodes that communicate with the LPBR, which has pre-defined consumption rate of $\frac{s}{T}$ bps. Under balanced coupling, each sensor node can transmit $\frac{s}{n}$ bits during the analysis time interval of T s. We thus identify two cases: if the amount of data generated by the processing phase is greater than $\frac{s}{n}$, then the sensor node has to buffer the remaining data in a high-power, typically off-chip, memory. Letting p J be the energy cost of storing one bit of information, the energy spent for buffering during the active time interval is: $p \int_{\frac{s}{n}}^{\infty} (\chi_k - \frac{s}{n}) P(\chi_k) d\chi_k$ J. Conversely, if the data generated is less than $\frac{s}{n}$, the sensor node enters an "idle" state, where b J is consumed for beaconing and other synchronization operations corresponding to the duration of the transmission of one bit. The energy spent during the idle mode of the analysis time interval is thus: $b \int_{0}^{\frac{s}{n}} (\frac{s}{n} - \chi_k) P(\chi_k) d\chi_k$ J.

2.5 Expected Energy Consumption

Summing all contributions 1~3 described in the previous subsection, the energy consumption of the coupled system, E_{c}, over the time interval T is:

$$E_{\mathrm{c}}(n, k) = ka + (g + j)\mathbb{E}[\mathcal{X}_k] + p \int_{\frac{s}{n}}^{\infty} (\chi_k - \frac{s}{n}) P(\chi_k) d\chi_k + b \int_{0}^{\frac{s}{n}} (\frac{s}{n} - \chi_k) P(\chi_k) d\chi_k. \tag{1}$$

Adding and subtracting $p \int_{0}^{\frac{s}{n}} (\chi_k - \frac{s}{n}) P(\chi_k) d\chi_k$ to (1) leads to:

$$E_{\mathrm{c}}(n, k) = ka + (p + g + j)\mathbb{E}[\mathcal{X}_k] - \frac{ps}{n} + (b + p) \int_{0}^{\frac{s}{n}} (\frac{s}{n} - \chi_k) P(\chi_k) d\chi_k. \tag{2}$$

The last equation forms the basis for the analytic exploration of the minimum energy consumption under the knowledge of the marginal PDF characterizing the data production and transmission process.

3 Analytic Derivation of Minimum Energy Consumption

Our objective is to derive the spatio-temporal parameters minimizing $E_c(n, k)$ in (2) subject to the spatio-temporal constraints defined in Section 2, that is:

$$\{n^\star, k^\star\} = \arg \min_{\forall n,k} E_c(n, k), \tag{3}$$

with

$$N_{\min} \leq n \leq N_{\max} \text{ and } k \geq K_{\min} \tag{4}$$

and $\{n^\star, k^\star\}$ the values deriving the minimum energy consumption.

In the following, we consider two different marginal distributions for $P(\chi_k)$ and derive the choice for n and k that minimizes the energy consumption, while ensuring the conditions imposed by the application constraints are met. While our analysis is assuming that n and k are continuous variables, once the $\{n^\star, k^\star\}$ values are derived, they can be discretized to the sets $\{\lfloor n^\star \rfloor, \lfloor k^\star \rfloor\}$, $\{\lceil n^\star \rceil, \lceil k^\star \rceil\}$ $\{\lceil n^\star \rceil, \lfloor k^\star \rfloor\}$ $\{\lfloor n^\star \rfloor, \lceil k^\star \rceil\}$ [if all four satisfy the constraints of (4)] in order to check which discrete pair of values derives the minimum energy consumption in (2). This is because: *(i)* the energy functions under consideration are continuous and differentiable; and *(ii)* we shall show that, for the data size PDFs under consideration, a *unique* minimum is found for (2) that is parametric to the setting of the temporal constraint ($K_{\min}$). As such, the analysis on the continuous variable space can be directly mapped onto the discrete variable set under the aforementioned discretization.

3.1 Illustrative Case: Uniform Distribution

When one has limited or no knowledge about the cumulative data transmitted by each VSN node during T, one can assume that $P(\chi_k)$ is uniform over the interval $[0, 2kr]$. That is,

$$P(\chi_k) = \begin{cases} \frac{1}{2kr} & 0 \leq \chi_k \leq 2kr \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

with $\mathbb{E}_U[\mathcal{X}_k] = kr$ corresponding to the mean value of the data transmitted by a node that produces k frames of r bits each on average. Using (5) in (2) leads to:

$$E_{c,U}(n, k) = k[a + r(p + j + g)] - \frac{ps}{n} + \frac{s^2(b + p)}{4n^2 kr}. \tag{6}$$

To obtain the solution to (3) under the energy consumption given by (6), one can search for critical points of $E_{\mathrm{c,U}}$. Imposing that the derivatives of $E_{\mathrm{c,U}}$ with respect to n and k are both equal to zero leads to:

$$\begin{cases} \frac{\partial E_{\mathrm{c,U}}}{\partial n} = \frac{ps}{n^2} - \frac{s^2(b+p)}{2n^3kr} = 0 \\ \frac{\partial E_{\mathrm{c,U}}}{\partial k} = a + r(p+j+g) - \frac{s^2(b+p)}{4n^2k^2r} = 0 \end{cases} \qquad (7)$$

It can be shown that the solution of (7) requires $a < 0$ (the detailed derivation is omitted due to space limitations). However, this is not physically feasible since a is the energy cost to acquire one frame. Hence, under the physical constraints of the problem, *there is no single (global) solution $\{n^\star, k^\star\} \in \mathbb{R} \times \mathbb{R}$ to (3) in its unconstrained form, i.e. when one ignores the constraints of (4)*. However, we may look at one or the other direction individually (i.e., along n or k) in order to find a local or global minimum for that particular direction and then choose for the other direction the value that minimizes (3) under the spatio-temporal constraints of (4). Subsequently, we can identify if the derived minima are unique under the imposed constraints and whether the entire region of support of the energy function has been covered by the derived solutions. These are investigated in the following.

n-direction We examine the function $E_{\mathrm{c,U}}$ along the plane $k = \bar{k}$, $\bar{k} \geq K_{\min}$, and analyze $E_{\mathrm{c,U}}(n, \bar{k})$ which is now a function of n only. It is straightforward to show by first-derivative analysis that the only candidate extremum or inflection point of $E_{\mathrm{c,U}}(n, \bar{k})$ is $n_{0,\mathrm{U}} = \frac{\beta_{\mathrm{U}}}{\bar{k}}$, with

$$\beta_{\mathrm{U}} = \frac{s(b+p)}{2pr} \qquad (8)$$

defined as a "system-specific" parameter (which depends on average bits transmitted and the energy penalty rates). This candidate extremum or inflection point is valid under the assumption that: $N_{\min} \leq n_{0,\mathrm{U}} \leq N_{\max}$, i.e. that the candidate point falls within the predefined spatial constraints of (4). Furthermore, we find that $\left. \frac{d^2 E_{\mathrm{c,U}}(n, \bar{k})}{dn^2} \right|_{n=n_{0,\mathrm{U}}} > 0$, which demonstrates that $n_{0,\mathrm{U}}$ is a local minimum. Given that local extrema must alternate within the region of support of a continuous and differentiable function, the boundary points ($n = N_{\min}$ and $n = N_{\max}$) cannot be local minima. Thus, $n_{0,\mathrm{U}}$ is the global minimum of $E_{\mathrm{c,U}}(n, \bar{k})$ within $N_{\min} \leq n \leq N_{\max}$.

Having derived the global minimum of $E_{\mathrm{c,U}}(n, \bar{k})$ along an arbitrary plane $k = \bar{k}$, $\bar{k} \geq K_{\min}$, we can now attempt to find the value of k, $k \geq K_{\min}$, that minimizes the energy function. Evaluating $E_{\mathrm{c,U}}(n, k)$ on $n = n_{0,\mathrm{U}}$ we obtain:

$$E_{\mathrm{c,U}}(n_{0,\mathrm{U}}, k) = k \left[a + r \left[(p+j) - \frac{p^2}{b+p} + g \right] \right]. \qquad (9)$$

Evidently, the value of k minimizing (9) is the minimum allowable, i.e. $k = K_{\min}$. Thus, the solution minimizing (3) in the n-direction is $\mathcal{S}_{n_{0,\mathrm{U}}} = \left(\frac{\beta_{\mathrm{U}}}{K_{\min}}, K_{\min} \right)$. This solution holds under the constraint:

$$N_{\min} \leq \frac{\beta_{\mathrm{U}}}{K_{\min}} \leq N_{\max}. \tag{10}$$

k-direction Similarly, we cut $E_{\mathrm{c,U}}(n,\,k)$ along the plane $n = \bar{n}$, $N_{\min} \leq \bar{n} \leq N_{\max}$, and minimize $E_{\mathrm{c,U}}(\bar{n},\,k)$ which is now function of k only. Following the steps presented earlier, we can show by first and second derivative analysis that the global minimum of $E_{\mathrm{c,U}}(\bar{n},\,k)$ occurs at $k_{0,\mathrm{U}} = \frac{\gamma_{\mathrm{U}}}{\bar{n}}$ with the "system-specific" parameter γ_{U} defined as:

$$\gamma_{\mathrm{U}} = \frac{s}{2} \sqrt{\frac{b+p}{r\left[a + r(p+j+g)\right]}}. \tag{11}$$

The global minimum of $k_{0,\mathrm{U}}$ given above holds under the assumption that $k_{0,\mathrm{U}} \geq K_{\min}$ due the predefined temporal constraint of (4). Having derived the global minimum of $E_{\mathrm{c,U}}(\bar{n},\,k)$ along an arbitrary plane $n = \bar{n}$, $N_{\min} \leq \bar{n} \leq N_{\max}$, we can now attempt to find the value of n, $N_{\min} \leq n \leq N_{\max}$, that minimizes the energy function. Evaluating $E_{\mathrm{c,U}}(n,\,k)$ on $k = k_{0,\mathrm{U}}$ we obtain:

$$E_{\mathrm{c,U}}(n,\,k_{0,\mathrm{U}}) = \frac{1}{n}\left[a + r(p+j+g)\gamma_{\mathrm{U}} - ps + \frac{s^2(b+p)}{4r\gamma_{\mathrm{U}}}\right] \tag{12}$$

Evidently, the value of n minimizing (12) is the maximum allowable, i.e. $n = N_{\max}$. Hence, the solution when attempting to minimize (12) in the k-direction under the constraints of (4) is $\mathcal{S}_{k_{0,\mathrm{U}}} = \left(N_{\max},\,\frac{\gamma_{\mathrm{U}}}{N_{\max}}\right)$, under the constraint:

$$K_{\min} \leq \frac{\gamma_{\mathrm{U}}}{N_{\max}}. \tag{13}$$

Uniqueness of solution and solution when (10) and (13) do not hold: Starting from (10), with a few straightforward manipulations we reach $\frac{\beta_{\mathrm{U}}}{N_{\max}} \leq K_{\min} \leq \frac{\beta_{\mathrm{U}}}{N_{\min}}$, with β_{U} defined in (8). The second constraint for $K_{\min}$ is provided by (13). It can be shown that $\beta_{\mathrm{U}} > \gamma_{\mathrm{U}}$ (derivation omitted due to page limitation), which demonstrates that *the constraints of the two established solutions are non-overlapping.* Thus, the solutions $\mathcal{S}_{n_{0,\mathrm{U}}}$ and $\mathcal{S}_{k_{0,\mathrm{U}}}$ are unique within their respective bounds for $K_{\min}$.

To establish the optimal solutions when *neither* of these two constraints is satisfied, we have to analyze what happens when $\frac{\gamma_{\mathrm{U}}}{N_{\max}} < K_{\min} < \frac{\beta_{\mathrm{U}}}{N_{\max}}$ or $K_{\min} > \frac{\beta_{\mathrm{U}}}{N_{\min}}$, as neither $\mathcal{S}_{n_{0,\mathrm{U}}}$ nor $\mathcal{S}_{k_{0,\mathrm{U}}}$ are applicable in such cases. It is straightforward to show that $\frac{\partial E_{\mathrm{c,U}}}{\partial n}$ and $\frac{\partial E_{\mathrm{c,U}}}{\partial k}$ are never zero within these intervals. Hence, the solution we are looking for must lie on one of the two boundary points: $(N_{\min}, K_{\min})$ or $(N_{\max}, K_{\min})$.

Let's focus on the case of $K_{\min} \in \left(\frac{\gamma_{\mathrm{U}}}{N_{\max}},\,\frac{\beta_{\mathrm{U}}}{N_{\max}}\right)$ and evaluate $E_{\mathrm{c,U}}(n,\,k)$ on the boundary plane $n = N_{\max}$. Since $E_{\mathrm{c}}(N_{\max}, k)$ is monotonically increasing for

$k > \frac{\gamma_U}{N_{\max}}$ the optimal point is $k = K_{\min}$, which leads to the solution $\mathcal{S}_{\max\min} = (N_{\max}, K_{\min})$. Similarly, let's look at the k direction by evaluating the energy function on the $k = K_{\min}$ plane. Now $n_{0,U} = \frac{\beta_U}{K_{\min}}$ is larger than $N_{\max}$ and is thus not admissible. Since $E_{c,U}(n, K_{\min})$ is decreasing for $n < n_{0,U}$, the optimal point is $n = N_{\max}$, which also leads to the solution $\mathcal{S}_{\max\min} = (N_{\max}, K_{\min})$. Hence we conclude that when $K_{\min} \in \left(\frac{\gamma_U}{N_{\max}}, \frac{\beta_U}{N_{\max}} \right)$, the optimal solution is $\mathcal{S}_{\max\min} = (N_{\max}, K_{\min})$. Finally, when $K_{\min} > \frac{\beta_U}{N_{\min}}$, following a similar analysis we reach that the optimal solution is $\mathcal{S}_{\min\min} = (N_{\min}, K_{\min})$.

Summarizing, when the data transmitted by each VSN node follows the Uniform distribution of (5), the set of solutions giving the minimum energy consumption in (3) under the spatio-temporal constraints of (4) is:

$$
\{n^\star, k^\star\}_U = \begin{cases} \left(N_{\max}, \frac{\gamma_U}{N_{\max}} \right) & \text{if } K_{\min} \leq \frac{\gamma_U}{N_{\max}} \\ (N_{\max}, K_{\min}) & \text{if } \frac{\gamma_U}{N_{\max}} < K_{\min} < \frac{\beta_U}{N_{\max}} \\ \left(\frac{\beta_U}{K_{\min}}, K_{\min} \right) & \text{if } \frac{\beta_U}{N_{\max}} \leq K_{\min} \leq \frac{\beta_U}{N_{\min}} \\ (N_{\min}, K_{\min}) & \text{if } K_{\min} > \frac{\beta_U}{N_{\min}} \end{cases} \tag{14}
$$

with β_U and γ_U defined by (8) and (11).

3.2 Pareto Distribution

We present a second example with the Pareto distribution. This distribution has been used, amongst others, to model the marginal data size distribution of TCP sessions that contain substantial number of small files and a few very large ones [11]. It will also be shown by the experimental results of this paper that it presents a good fit to multimedia traffic generated by visual features extraction algorithms and hence it warrants detailed study under the proposed VSN framework.

Consider $P(\chi_k)$ as the Pareto distribution with scale v and shape $\alpha > 1$:

$$
P(\chi_k) = \begin{cases} \alpha \frac{v^\alpha}{\chi_k^{\alpha+1}}, & \chi_k \geq v \\ 0, & \text{otherwise} \end{cases}. \tag{15}
$$

Setting $v = \frac{\alpha-1}{\alpha}kr$ leads to $\mathbb{E}_P[\mathcal{X}_k] = kr$, i.e., we match the expected transmission data volume to that of the Uniform distribution.

Under (15), the energy expression of (2) becomes:

$$
E_{c,P} = k\left[a + r(p + j + g)\right] + \frac{bs}{n} + (b + p)\left(\frac{v^\alpha n^{\alpha-1}}{s^{\alpha-1}(\alpha-1)} - \frac{\alpha v}{\alpha-1} \right). \tag{16}
$$

Following the same analysis as for the Uniform PDF, we conclude that, when the data transmitted by each VSN node follows the Pareto distribution of (15), the set of solutions giving the minimum energy consumption in (3) under the spatio-temporal constraints of (4) is:

$$\{n^\star, k^\star\}_\mathrm{P} = \begin{cases} \left(N_\mathrm{max}, \frac{\gamma_\mathrm{P}}{N_\mathrm{max}}\right) & \text{if } K_\mathrm{min} \leq \frac{\gamma_\mathrm{P}}{N_\mathrm{max}} \\ (N_\mathrm{max}, K_\mathrm{min}) & \text{if } \frac{\gamma_\mathrm{P}}{N_\mathrm{max}} < K_\mathrm{min} < \frac{\beta_\mathrm{P}}{N_\mathrm{max}} \\ \left(\frac{\beta_\mathrm{P}}{K_\mathrm{min}}, K_\mathrm{min}\right) & \text{if } \frac{\beta_\mathrm{P}}{N_\mathrm{max}} \leq K_\mathrm{min} \leq \frac{\beta_\mathrm{P}}{N_\mathrm{min}} \\ (N_\mathrm{min}, K_\mathrm{min}) & \text{if } K_\mathrm{min} > \frac{\beta_\mathrm{P}}{N_\mathrm{min}}, \end{cases} \tag{17}$$

with

$$\beta_\mathrm{P} = \frac{s\alpha}{r(\alpha - 1)} \left(\frac{b}{b+p}\right)^{\frac{1}{\alpha}}, \tag{18}$$

and

$$\gamma_\mathrm{P} = \frac{s\alpha}{r(\alpha - 1)} \left(\frac{r(b - j - g) - a}{r(b+p)}\right)^{\frac{1}{\alpha - 1}}. \tag{19}$$

The details of the derivation of (17) follow the same steps as the ones detailed for the Uniform distribution and are omitted for brevity of description.

3.3 Discussion

The results of this section can be used in practical applications to assess the impact of the spatio-temporal constraints and the data production and transmission process (as characterized by its marginal PDF) on the energy consumption of VSNs, under a variety of energy consumption rates for the radio and multimedia subsystems. For example, under particular technology specifications (i.e. given j, b, p, a and g parameters) and preset number of nodes and frames to capture within the activation time interval, one can determine the required energy in order to achieve the designated visual data gathering task. Similarly, under the proposed framework, one can determine the data production and transmission (marginal) PDFs that meet preset energy supply, spatio-temporal constraints and technology parameters (i.e. energy consumption per bit for each task). As a result, our proposed energy consumption model and the associated analytic results can be used in many ways for early exploration of system, network, and data production parameters in VSNs that match the design specifications of classes of application domains. Such application examples are given in Section 5.

4 Evaluation of the Analytic Results

For the radio subsystem of Figure 1, we used TelosB sensor nodes equipped with a 802.15.4-compliant CC2420 radio transceiver. Each TelosB runs the low-power Contiki 2.6 operating system and implements the open-source TFDMA protocol proposed recently [6] for time-synchronized multichannel communications with the LPBR. Given that the TFDMA protocol ensures collision-free packet transmissions by each node via application-layer adaptation of the transmission

slots based on a desynchronization mechanism [6], we enabled the low-power NullMAC and NullRDC options of the Contiki OS. This led to data consumption rate at the application layer of $\frac{s}{T} = 144$ kbps. Evidently the usage of the TFDMA protocol is only for illustration purposes and any other protocol ensuring collision-free communications by centralized or distributed slot allocation, such as the IEEE 802.15.4e-2006 GTS [7] or the IEEE 802.15.4e-2012 TSCH [4] can be used.

All energy measurements were performed using a Tektronix MDO4104-6 oscilloscope to capture the real-time current consumption at a high-tolerance 1 Ohm resistor placed in series with each TelosB node running the described operations. Under these operational settings, the average transmission cost per bit, g J/b, as well as the cost for beaconing, b J/b, and buffering, p J/b, were established experimentally; their values are shown at the top half of Table 1.

Although capable of simple processing tasks, the TelosB is not a multimedia platform. However, it can be attached via its integrated FTDI USB chip to a more powerful platform such as the BeagleBone [12], or integrated with a low power DSP processor, as done in the CITRIC project [2]. Here, we assume that the multimedia sybsystem is based on an application-specific integrated circuit (ASIC), such as the one proposed recently for energy-efficient visual feature extraction [10] in images. Obviously, the processing cost per bit is application and hardware dependent; in the following, we use g derived from Park *et al* [10] and reported at the bottom half of Table 1. Finally, concerning image acquisition, we considered the energy cost of acquiring a frame (a J) derived from the specifications of the OV7670 camera sensor, which is widely used in low-power visual sensor platforms [2] and is also reported in Table 1.

Under the settings described previously and shown in Table 1, our first goal is to validate the basic analytic expressions of Section 3, namely (6) and (16), with respect to the energy consumption measured when performing Monte-Carlo–based experiments combined with actual energy measurement. To this end, we simulated the multimedia data production process on each VSN node by: *(i)* artificially creating several sets of data size values according to the marginal PDFs of Section 3 via rejection sampling and *(ii)* setting the mean data size per video frame to $r = 5.2$ kbit. The sets containing data sizes are copied onto the read-only memory of each sensor node during deployment. At run time, each node fetches a new frame size from the preloaded set, produces artificial data according to it (akin to receiving the information from the multimedia subsystem) and transmits the information to the LPBR following the process described in the system model of Section 2. Depending on the frame size, the node can enter in idling/beaconing state or it can buffer the data exceeding the allocated TFDMA slots.

We report here energy measurements obtained under varying values of n and k. The chosen active time interval was set to be $T = 154$ s and, beyond measuring the accuracy of the model versus experiments, we also compared the theoretically-optimal values for k and n according to Section 3 with the ones producing the minimum energy consumption in the experiments. For the

reported experiments of Figure 2 and Table 2, the spatio-temporal constraints were: $N_{\min} = 2$, $N_{\max} = 16$ and $K_{\min} = 2T$ frames, i.e. two frames per second. All our reported measurements and the values for k are normalized to a one-second interval for easier interpretation of the results.

As one can see from Figure 2 and Table 2, the theoretical results match the experimental results for all the tested distributions, with the maximum percentile error between them limited to 6.12% and all the coefficients of determination between the experimental and the model points being above 0.995. In addition, the theoretically-obtained optimal values for $\{n^\star, k^\star\}$ from (14) and (17), are always in agreement with the experimentally-derived values that were found to offer the minimum energy consumption under the chosen spatio-temporal constraints. We have observed the same level of accuracy for the proposed model under a variety of data sizes (r), active time interval durations (T) and spatio-temporal constraints $(N_{\min}, N_{\max}$ and $K_{\min})$, but omit these repetitive experiments for brevity of exposition.

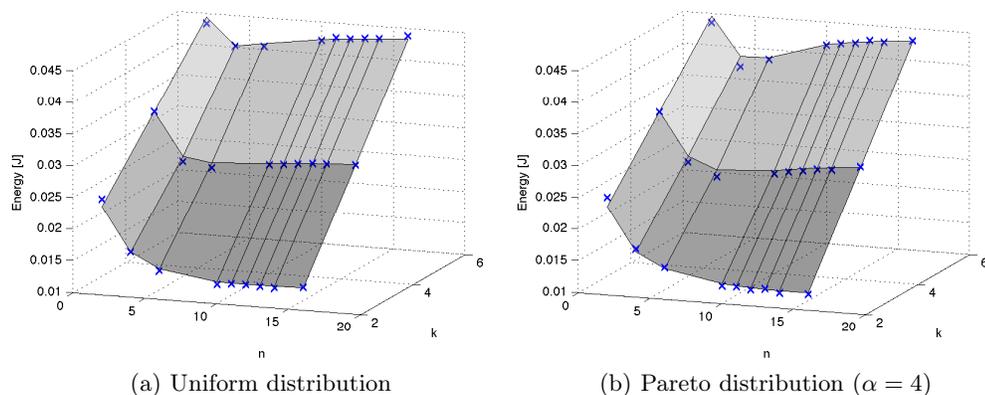(a) Uniform distribution　　　　　　(b) Pareto distribution ($\alpha = 4$)

Fig. 2. The grayscale surfaces show the energy consumption of a single camera sensor node in function of the number of frames per second and the total number of nodes. The blue crosses correspond to the value of the consumed energy as measured from the sensor network test-bed. All energy values and frames (k) are normalized to a one-second interval.

5　Application in Visual Features Extraction

In order to assess the proposed model against application deployments, we consider the extraction and transmission of local visual features for image analysis. This scenario represents a wide range of practical VSN-related deployments proposed recently [2,8]. In a nutshell, salient keypoints of an image are identified by

Table 2. Differences between theoretical and experimental results and the optimal value for the number of nodes and the frames-rates for the considered data transmission (marginal) PDFs under the settings of Figure 2.

PDF	Mean Error (%)	Max Error (%)	R^2	Theoretical optimum
Uniform	1.08	4.81	0.9987	$\{n^\star, k^\star\}_\mathrm{U} = \{12, 2\}$
Pareto ($\alpha = 4$)	1.64	6.05	0.9973	$\{n^\star, k^\star\}_\mathrm{P} = \{14, 2\}$

means of a *detector* algorithm, and the patch of pixels around each keypoint is encoded in a feature vector by a specialized *descriptor* algorithm. Here, we focus on corner-like local features, such as the ones detected by the FAST corner detector [9], which is optimized for fast and efficient detection on low-power devices. For what concerns the descriptor, several algorithms are available in the literature: here we assume to use the BRIEF descriptor, which produces a 64 bytes binary feature vector starting from intensity comparisons between pixels of the patch to be described, thus being particularly tailored for resource constrained devices. As input data, we considered the video sequences from the PETS2007 dataset[1], which are taken from an airport surveillance video system. The original resolution of all sequences is 768 x 576 pixels and the original frame rate is 25 frames-per-second. Similarly as before, the process is repeated for different video sequences and different frame rates (i.e. different values of k normalized to frames-per-second).

We repeated the experimental measurements described in Section 4 for this application scenario and under the same spatio-temporal constraints ($N_\mathrm{min} = 2$, $N_\mathrm{max} = 16$, $K_\mathrm{min} = 2T$, i.e. two frames per second), this time loading on the sensor network the traces of data sizes computed after the processing of the video sequences in the two cases and utilizing the energy parameters of Table 1. Then, we fitted[2] the energy measurements with the ones produced by the Pareto distribution. The best fit was obtained under parameters $\alpha = 4$, $v = kr$ and $r = 11.7$ kbit, as shown in Figure 3, with coefficient of determination value $R^2 \cong 0.96$. Similarly as before, all reported energy values and number of frames are normalized to a one-second interval for easier interpretation of the results.

Given the high accuracy of the Pareto-based energy model against the application results, we utilized the settings for the minimum energy consumption derived for the Pareto case [see (17)] to ascertain the energy saving that can be potentially achieved against arbitrary (ad-hoc) settings. As an example, in Table 3, we consider two different cases, characterized by different spatio-temporal constraints. For each case, we compare the optimal solution given by (17) with an ad-hoc "least-cost" solution that assumes values equal to the minimum spatio-temporal constraints (under the intuitive assumption that less nodes and less frames-per-second lead to smaller energy consumption). Evidently, the proposed

---

[1] http://pets2007.net

[2] Fitting is performed by matching the average data size r of each distribution to the average data size of the set of visual features.
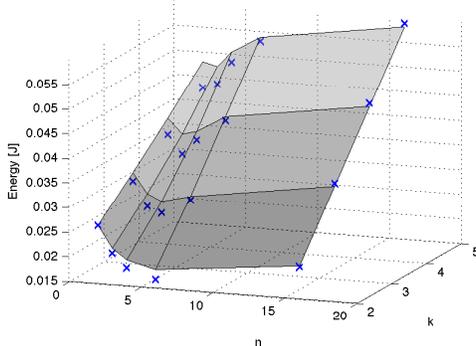
Fig. 3. The energy function for the considered application scenario. The grayscale surfaces represent the fitted energy function obtained with the Pareto PDF, while the blue crosses represent the experimental measurements. All energy values and frames (k) are normalized to a one-second interval.

Table 3. Energy consumption under varying spatio-temporal constraints with ad-hoc settings and with settings derived from the proposed Pareto model ($K_{\min}$, k and the energy values are normalized to a one-second interval).

	Spatio-temporal	Visual features extraction		
	Constraints	Ad-hoc deployment	Proposed approach	Gain
Case 1	$K_{\min} = 5$	$k_{\mathrm{adhoc}} = 5$	$k^{\star} = 5$	10.30%
	$N_{\min} = 3$	$n_{\mathrm{adhoc}} = 3$	$n^{\star} = 6$	
	$N_{\max} = 6$	$E_{\mathrm{c}} = 0.031$ J	$E_{\mathrm{c}} = 0.028$ J	
Case 2	$K_{\min} = 2$	$k_{\mathrm{adhoc}} = 2$	$k^{\star} = 2$	48.65%
	$N_{\min} = 2$	$n_{\mathrm{adhoc}} = 2$	$n^{\star} = 10$	
	$N_{\max} = 10$	$E_{\mathrm{c}} = 0.022$ J	$E_{\mathrm{c}} = 0.012$ J	

approach allows for significant energy savings, which can be more than 45% in comparison to the ad-hoc settings.

6 Conclusions

We proposed an analytic model for the energy consumption of wireless VSN arranged in a star-shaped topology under preset spatio-temporal constraints. We derived analytic conditions for the optimal spatio-temporal settings within the VSN for different PDFs characterizing the multimedia data volume to be transmitted by each node. Monte-Carlo experiments performed via an energy-measurement testbed revealed that the proposed model's accuracy is within 7% of the obtained energy consumption. Applying the model to the application scenario of visual features extraction demonstrated that substantial energy savings can be obtained via the proposed approach against ad-hoc settings for the spatio-temporal parameters of the VSN. As such, the proposed model can be

used for early-stage studies of VSNs to determine the best operational parameters prior to cumbersome and costly real-world deployment and testing. Future research directions include the extension of the proposed framework to multi-tier cluster-tree topologies as well as other multimedia traffic distributions.

Acknowledgments

References

1. G. Lu, B. Krishnamachari, and C. Raghavendra, "Performance evaluation of the IEEE 802.15.4 mac for low-rate low-power wireless networks," in *IEEE Internat. Conf. on Perform., Comp., and Comm.*. IEEE, 2004, pp. 701–706.
2. P. Chen, P. Ahammad, C. Boyer, S.-I. Huang, L. L., E. Lobaton, M. Meingast, S. Oh, S. Wang, P. Yan, A. Yang, C. Yeo, L.-C. Chang, J. D. Tygar, and S. Sastry, "CITRIC: A low-bandwidth wireless camera network platform," in *ACM/IEEE Int. Conf. on Distrib. Smart Cam., ICDSC*, Sept., pp. 1–10.
3. Y. Charfi and B. Canada, "Challenging issues in visual sensor networks," *IEEE Wireless Comm.*, vol. 16, no. 2, pp. 44–49, Apr. 2009.
4. A. Bachir *et al.*, "MAC essentials for wireless sensor networks," *IEEE Comm. Surv. Tut.*, vol. 12, no. 2, pp. 222 –248, quarter 2010.
5. J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Future Gen. Comp. Syst. J.*, vol. 29, no. 7, pp. 1645–1660, Sep. 2013.
6. D. Buranapanichkit and Y. Andreopoulos, "Distributed time-frequency division multiple access protocol for wireless sensor networks," *IEEE Wireless Comm. Letters*, vol. 1, no. 5, pp. 440–443, Oct. 2012.
7. A. Koubâa, M. Alves, and E. Tovar, "GTS allocation analysis in IEEE 802.15.4 for real-time wireless sensor networks," in *IEEE Proc. Internat. Par. and Distrib. Process. Symp., IPDPS*. IEEE, 2006, pp. 8–pp.
8. Y. Kwon and D. Shin, "The security monitoring system using IEEE 802.15.4 Protocol and CMOS Image Sensor," in *Proc. IEEE Internat. Conf. on New Trends in Inf. and Serv. Sci., NISS '09*. IEEE, 2009, pp. 1197–1202.
9. E. Rosten, R. Porter, and T. Drummond, "FASTER and better: A machine learning approach to corner detection," *IEEE Trans. Patt. Anal. and Machine Intel.*, vol. 32, pp. 105–119, 2010.
10. J.-S. Park, H.-E. Kim, and L.-S. Kim, "A 182mW 94.3 fps in full HD pattern-matching based image recognition accelerator for embedded vision system in 0.13 um CMOS technology," *IEEE Trans. on Circ. and Syst. for Video Technol.*, vol. 23, no. 5, pp. 832 – 845, May 2013.
11. K. Park, "On the relationship between file sizes, transport protocols, and self-similar network traffic," in *Proc. IEEE Internat. Conf. on Network Protocols, ICNP*, 1996, pp. 171–180.
12. M. Rachmadi *et al.*, "Adaptive traffic signal control system using camera sensor and embedded system," in *TENCON 2011 - 2011 IEEE Region 10 Conf.*. IEEE, 2011, pp. 1261–1265.