Performance of in-network processing for visual analysis in wireless sensor networks

Hussein Al-Zubaidy, György Dán, Viktoria Fodor
School of Electrical Engineering, KTH Royal Institute of Technology, Stockholm, Sweden.
E-mail: {hzubaidy, gyuri, vjfodor}@kth.se.

***Abstract*—Nodes in a sensor network are traditionally used for sensing and data forwarding. However, with the increase of their computational capability, they can be used for in-network data processing, leading to a potential increase of the quality of the networked applications as well as the network lifetime. Visual analysis in sensor networks is a prominent example where the processing power of the network nodes needs to be leveraged to meet the frame rate and the processing delay requirements of common visual analysis applications. The modeling of the end-to-end performance for such networks is, however, challenging, because in-network processing violates the flow conservation law, which is the basis for most queuing analysis. In this work we propose to solve this methodological challenge through appropriately scaling the arrival and the service processes, and we develop probabilistic performance bounds using stochastic network calculus. We use the developed model to determine the main performance bottlenecks of networked visual processing. Our numerical results show that an end-to-end delay of 2-3 frame length is obtained with violation probability in the order of 10^{-6}. Simulation shows that the obtained bounds overestimates the end-to-end delay by no more than 10%.**

I. Introduction

Due to the advances of hardware design, the memory and the computational capability of wireless sensor nodes has increased significantly in recent years. As a consequence, cheap sensors are not only capable of sensing and of information transmission, but can contribute to the processing of sensed data. The resulting in-network data processing is a promising approach for extending the network lifetime and for improving application performance, optimizing not only the path of data transmission across the network, but also the location of data processing.

Wireless visual sensor networks (VSNs) give a prominent example where in-network processing can have significant advantage. VSNs are built from low–cost, battery operated cameras and sensor nodes, and capture and process visual information from the surrounding environment to support applications like surveillance, tracking and traffic monitoring. They differ from traditional sensor networks considering both the amount of collected and transmitted information and the computational complexity of processing the information. Visual sensors capture visual information, e.g., images or sequences of video frames. This rich information should be collected and processed, usually within a delay limit, to allow timely visual analysis. As collecting the raw image data would require too high transmission rates, recent works propose to perform the processing at the nodes of the sensor network, such that only the information required for decision making, a so called *set of visual features*, is sent to the intended receiver, the sink node of the sensor network [1] [2]. Performing the processing at the sensor nodes results in a major reduction in the dimensionality of the data to be transmitted and hence the required transmission resources.

ISBN 978-3-901882-68-5

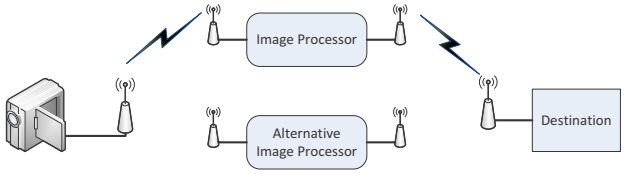

Fig. 1. Wireless visual sensor network with distributed image processing.

Extracting features from images requires, however, complex computations, and consequently feature extraction at the low–cost camera nodes would lead both to high processing delays and to fast depletion of the energy resources of the camera nodes. As camera nodes need to be parameterized and calibrated in situ in order to serve complex analysis tasks, it is advisable to decrease their energy consumption and thus the frequency of replacement, and utilize the energy resources of the easy to deploy network nodes, as shown in Fig. 1 [3] [4].

While in-network processing seems to be a promising solution, it still requires the camera node to transmit the image content and it introduces processing delay at the network nodes. If transmission and processing need to be finished within a delay limit, the sensor network needs to be appropriately dimensioned, and dimensioning requires simple yet powerful system performance models.

The main methodological challenge when modeling in-network processing systems is traffic morphism at the image processing nodes: the departing traffic from an image processing node is a fraction of the traffic arriving at that node. This violates the premise for any traditional stable queuing system analysis based on the flow conservation law, and thus it requires adaptation and extension of the existing performance evaluation methodology in order to model in-network processing.

In this paper we show that stochastic network calculus [5] can be used to address this challenge by appropriately scaling the arrival and service processes. We provide a methodology

to compute end–to–end probabilistic performance bounds for visual sensor networks with intermediate processing nodes. We adopt a system–theoretic stochastic network calculus approach [6], developed to model the end-to-end performance of data transmission over multiple network nodes and capitalize on recent results [7] that provide a service model for wireless fading channels. Based on the analytic model we evaluate how the networking environment and the transmission and processing resources affect the end–to–end performance of networked visual processing. The proposed approach provides analytically quantifiable performance that exposes trends and trade-offs of the investigated wireless VSN. The analyzed model is suitable to study other emerging systems, e.g., measure-analyze-control in cyber-physical systems.

The rest of the paper is organized as follows. Sec. II presents related work. Sec. III presents background regarding the problem and the methodology used. Sec. IV presents our model. The main results are given in Sec. V. Numerical examples are discussed in Sec. VI. Sec. VII concludes the paper.

II. Related Work

The challenge of networked visual analysis is addressed in [8], [9], defining feature extraction schemes with low computational complexity. To decrease the transmission bandwidth requirements, [10], [11] propose lossy image coding schemes optimized for descriptor extraction, while [12]–[14] give solutions to decrease the number and the size of the descriptors to be transmitted. In [15] the number and the quantization level of the considered descriptors are jointly optimized to maximize the accuracy of the recognition, subject to energy and bandwidth constraints. In [16], the authors suggest that, in the case of video sequences, the candidate descriptor locations be chosen based on motion prediction and only these areas are processed and transmitted, in order to decrease the transmission requirements of feature extraction. In [17], [18] intra- and inter-frame coding of descriptors is proposed to decrease the transmission requirements.

Our work is motivated by recent results on the expected transmission and processing load of visual analysis in sensor networks [2], [15]. Measurements in [2] demonstrate that processing at the camera or at the sink node of the VSN leads to significant delays, and thus distributed processing is necessary for real-time applications. The experiments in [3] and [15] determined the characteristics of the processing delay in the processing nodes. Their experimental results show that the delay grows linearly as the image size as well as the number of detected descriptors increase.

Since its conception, $(\min, +)$ network calculus has been successfully applied to provide deterministic, and later on probabilistic performance bounds for wired networks. Recently, there have been many attempts to use a similar approach to analyze the performance of wireless networks, e.g., [19]–[22]. These efforts highlighted the difficulties of modeling and analyzing wireless networks. The communication link in wireless networks is prone to noise and fading which results in randomly varying received SNR, and hence randomly varying channel capacity. This proved to be the single most challenging aspect of developing wireless network calculus. Most of the existing work in this area work around this difficulty by assuming an abstracted finite-state Markov channel (FSMC) model of the underlying fading channel [24]. However, the complexity of the Markov channel model analysis limits the applicability of this approach in multi–hop wireless networks analysis with more than few state FSMC model and more than two hops. Flow transformation due to loss, dynamic routing or retransmissions was studied in [25] by defining a (virtual) scaling element to compensate for the lost/added traffic to the original flow. This technique is based on an earlier work [26] that provided such scaling element in deterministic settings.

In this work, we follow an alternative approach that was proposed by Al-Zubaidy et al [7], where a wireless network calculus based on the $(\min, \times)$ dioid algebra was developed. The main premise for this approach is that the channel capacity, and hence the offered service, of fading channels is related to the instantaneous received SNR through the logarithmic function as given by the Shannon capacity function, $C(\gamma) = \log(1+\gamma)$. Hence, an equivalent representation of the channel capacity in an isomorphic transform domain, obtained using the exponential function, would be $e^{C(\gamma)} = 1 + \gamma$. This simplifies the otherwise cumbersome computations of the end–to–end performance bounds.

III. Background

A. Visual Feature Extraction

Visual feature descriptor computation for an image involves two major processing steps: feature detection and feature descriptor extraction. First, a blob or an edge detector is used to identify features, which are regions in the image for which the filter response exceeds a pre-specified detection threshold [8], [9]. The desirable number of features is defined by the image processing application. As it is shown in [27], the number of detected features can be kept nearly constant by tuning the detection threshold dynamically according to the image content. After all features are detected, a fixed size feature descriptor is extracted for each of them from the image patches.

The time required for the feature detection is linear to the image size and to the number of features detected [15] [3]. The time for the extraction of a feature descriptor is roughly constant, which again gives a linear increase of the extraction time in the number of features.

B. $(\min, \times)$ Network Calculus

Network calculus in general aims at deriving end-to-end performance bounds in terms of backlog (that is, data waiting or under transmission in the network) and end–to–end delay, based on cumulative functions of arrivals, services and departures. The methodology we use considers fluid–flow traffic that is infinitely divisible, and operates in discrete–time domain, where time slots are denoted by $t \in \{0, 1, \ldots\}$ and slot

duration $\Delta t = 1$. Let the cumulative arrivals, service and departures to a node k during the time interval $[\tau, t)$ be denoted by the real–valued non–negative bivariate processes $A_k(\tau,t), S_k(\tau,t)$ and $D_k(\tau,t)$ respectively. These processes are non-decreasing in t with $A_k(t,t) = S_k(t,t) = D_k(t,t) = 0$ and $A_k(0,t) \leq D_k(0,t)$ for all t.

Different from the typical $(\min,+)$ network calculus, $(\min,\times)$ network calculus transforms the problem into an alternative domain, called SNR domain, where the SNR service process[1] $(\mathcal{S}_i)$ is obtained by taking the exponent of the original service process, i.e., $\mathcal{S}_i = e^{S_i}$. A network element i is referred to as *dynamic SNR server*, if it offers a service $\mathcal{S}_i$ that satisfies the following input–output inequality [6]

$$\mathcal{D}(0,t) \geq \mathcal{A} \otimes \mathcal{S}_i(0,t)\,. \tag{1}$$

The $(\min,\times)$ convolution and deconvolution are respectively defined for any two SNR processes $\mathcal{X}_1(\tau,t)$ and $\mathcal{X}_2(\tau,t)$ as

$$\mathcal{X}_1 \otimes \mathcal{X}_2(\tau,t) \stackrel{\triangle}{=} \inf_{\tau \leq u \leq t} \left\{ \mathcal{X}_1(\tau,u) \cdot \mathcal{X}_2(u,t) \right\},$$

$$\mathcal{X}_1 \oslash \mathcal{X}_2(\tau,t) \stackrel{\triangle}{=} \sup_{u \leq \tau} \left\{ \frac{\mathcal{X}_1(u,t)}{\mathcal{X}_2(u,\tau)} \right\}.$$

The total network backlog in the SNR domain at any time $t \geq 0$ is given by

$$\mathcal{B}(t) = e^{B(t)} = \frac{\mathcal{A}(0,t)}{\mathcal{D}(0,t)} \tag{2}$$

and the end–to–end delay is given by

$$\mathcal{W}(t) = W(t) = \inf\left\{ w \geq 0 : \frac{\mathcal{A}(0,t)}{\mathcal{D}(0,t+w)} \leq 1 \right\}. \tag{3}$$

A main advantage of using network calculus is that one can obtain the network service process by concatenating the service processes for all nodes along a path. The network SNR service process $\mathcal{S}_{\rm net}$ is given by [7]

$$\mathcal{S}_{\rm net}(\tau,t) = \mathcal{S}_1 \otimes \mathcal{S}_2 \otimes \mathcal{S}_3(\tau,t)\,. \tag{4}$$

The computation of the $(\min,\times)$ convolution and deconvolution operation is not straight forward since they involve product and quotients of random processes. Thus, an exact solution for Eq. (4) may not be feasible. However, bounds on these operations are possible. Probabilistic performance bounds for a multi-hop wireless network are then given by the following theorem, due to [7].

Theorem 1. *For a multi–hop wireless network with an SNR arrival process $\mathcal{A}$ and a network SNR service process $\mathcal{S}_{\rm net}$, the end–to–end probabilistic performance bounds for a violation probability $\varepsilon > 0$ are given by*

- OUTPUT BURSTINESS: $Pr\big(D(\tau,t) > d^{\varepsilon}_{\rm net}\big) \leq \varepsilon$, *where*

$$d^{\varepsilon}_{\rm net}(\tau,t) = \inf_{s>0}\left\{ \frac{1}{s}\big(\log \mathsf{M}_{\rm net}(s,\tau,t) - \log\varepsilon\big)\right\};$$

- BACKLOG: $Pr\big(B(t) > b^{\varepsilon}_{\rm net}\big) \leq \varepsilon$, *where*

$$b^{\varepsilon}_{\rm net} = \inf_{s>0}\left\{ \frac{1}{s}\big(\log \mathsf{M}_{\rm net}(s,t,t) - \log\varepsilon\big)\right\};$$

- DELAY: $Pr\big(W(t) > w^{\varepsilon}_{\rm net}\big) \leq \varepsilon$, *where $w^{\varepsilon}_{\rm net}$ is the smallest w satisfying*

$$\inf_{s>0}\Big\{ \mathsf{M}_{\rm net}(s,t+w,t)\Big\} \leq \varepsilon\,,$$

where the function $\mathsf{M}_{\rm net}(s,\tau,t)$ is given by

$$\mathsf{M}_{\rm net}(s,\tau,t) = \sum_{u=0}^{\min(\tau,t)} \mathcal{M}_{\mathcal{A}}(1+s,u,t)\,\mathcal{M}_{\mathcal{S}_{\rm net}}(1-s,u,\tau) \tag{5}$$

and $\mathcal{M}_Z(s) = E[Z^{s-1}]$ is the Mellin transform of the nonnegative random variable Z, for any complex valued s and when the expectation exists [23].

Theorem 1 states that probabilistic performance bounds can be obtained in terms of the Mellin transforms of the SNR arrival and the network SNR service processes. However, the network SNR service process is given by the $(\min,\times)$ convolution of all the SNR service processes along the multi–hop path as given by Eq. (4). Computing the exact Mellin transform for $\mathcal{S}_{\rm net}$ may not be possible. Instead, we will use the following lemma to obtain a bound on the Mellin transform of the network SNR service process when the SNR servers are independent.

Lemma 1. *For $s < 1$, let $\mathcal{S}_1(\tau,t)$ and $\mathcal{S}_2(\tau,t)$ be two independent SNR service processes. The Mellin transform of $\mathcal{S}_1 \otimes \mathcal{S}_2(\tau,t)$ is bounded by*

$$\mathcal{M}_{\mathcal{S}_1\otimes\mathcal{S}_2}(s,\tau,t) \leq \sum_{u=\tau}^{t} \mathcal{M}_{\mathcal{S}_1}(s,\tau,u)\cdot\mathcal{M}_{\mathcal{S}_2}(s,u,t)\,. \tag{6}$$

Lemma 1 suggests that the Mellin transform of the $(\min,\times)$ convolution of two independent service processes is bounded by a function of their Mellin transforms. Analogous bound can be obtained for the $(\min,\times)$ deconvolution. The proof of Lemma 1 can be found in [7].

The above results can be extended for the case when the two SNR service processes are dependent. Then Lemma 2 below can be used to obtain a bound on the $(\min,\times)$ convolution.

Lemma 2. *Let $\mathcal{S}_1(\tau,t)$ and $\mathcal{S}_2(\tau,t)$ be two dependent SNR service processes. The Mellin transform of $\mathcal{S}_1 \otimes \mathcal{S}_2(\tau,t)$ for any $s<1$ and any $p>1$ and $1/p+1/q=1$, is bounded by*

$$\mathcal{M}_{\mathcal{S}_1\otimes\mathcal{S}_2}(s,\tau,t) \leq \sum_{u=\tau}^{t}\Big(\mathcal{M}_{\mathcal{S}_1}(1-p+sp,\tau,u)\Big)^{\frac{1}{p}}\cdot\Big(\mathcal{M}_{\mathcal{S}_2}(1-q+sq,u,t)\Big)^{\frac{1}{q}}. \tag{7}$$

Proof. We use the fact that the function $f(x) = x^{s-1}$ is decreasing for $s < 1$. Using the definition of the Mellin

[1]We use the calligraphic upper–case letters to represent traffic and service processes in the SNR domain and to distinguish them from their bit domain (where traffic and service are measured in bits) counterparts.

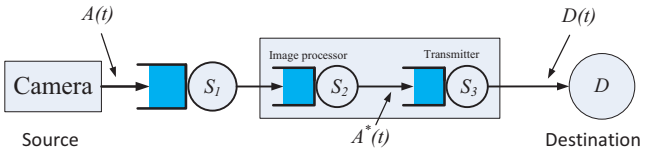

Fig. 2. System model description.

transform, for any $s < 1$, we have

$$\mathcal{M}_{\mathcal{S}_1 \otimes \mathcal{S}_2}(s,\tau,t) = E\Big[\Big(\inf_{\tau \le u \le t}\{\mathcal{S}_1(\tau,u)\cdot \mathcal{S}_2(u,t)\}\Big)^{s-1}\Big]$$
$$= E\Big[\sup_{\tau \le u \le t}\Big\{(\mathcal{S}_1(\tau,u))^{s-1}\cdot(\mathcal{S}_2(u,t))^{s-1}\Big\}\Big]$$
$$\le \sum_{u=\tau}^{t}\Big(E\Big[(\mathcal{S}_1(\tau,u))^{(s-1)p}\Big]\Big)^{\frac{1}{p}}\Big(E\Big[(\mathcal{S}_2(u,t))^{(s-1)q}\Big]\Big)^{\frac{1}{q}}.$$

In the last step, we used the union bound and non-negativity of the service processes $\mathcal{S}_1$ and $\mathcal{S}_2$, then applied Hölder's inequality to bound the expectation of the products. The Lemma then follows from the Mellin transform definition. □

It is worth noting that due to applying Hölder's inequality, the bound obtained in Lemma 2 may be loose, but this may be the best possible bound one can get without further information regarding the correlation of the two SNR service processes. The bound can be tightened by optimizing over the value of p.

IV. System Model

We consider a visual sensor network as shown in Fig. 1, consisting of a single camera node, a processing node and a sink node, which is the destination of the information transmission. The camera node captures a sequence of frames, and sends the frames to the image processing node in the network. Once a frame is *completely received* by the processing node, it computes a constant N feature descriptors. The feature descriptor data (not the original image) are then forwarded to the sink via a wireless channel. The computational power and energy resources of all nodes are limited as well as their communication bandwidth.

Fig. 2 shows a queuing model for the wireless visual sensor network. Server S_1 represents the wireless channel from the camera to the processing node, S_2 represents the image processing, and S_3 the wireless transmission from the image processor to the destination We assume that the servers have infinite buffer. In this paper we consider the specific case with one image processor. The extension of several parallel processors is possible in the same framework and will be the topic of our future work. The extension for multihop wireless transmission follows directly from [7].

A. Arrival and Service Curves

The model we develop is a discrete–time fluid model. We denote time slots by $t \in \{0,1,\ldots\}$ and the slot duration is Δt. The slot duration is chosen small enough that block fading is a reasonable model for the wireless channels. We denote by γ_t the instantaneous SNR during slot t, which remains unchanged during a time slot and changes independently in subsequent time slots. This is a reasonable fading model that is used extensively in the literature to model many systems including orthogonal frequency division multiplexing (OFDM) and frequency hopping spread spectrum (FHSS) systems. In this work we consider a Rayleigh fading model for the wireless channels in the system, but the approach works for other fading models as well. In the following we describe the arrival and service curves at the three servers.

Server 1: We denote by n the number of frames per time slot generated by the camera, and by r the number of bits per frame. This corresponds to a periodic source with period $\frac{1}{n}$. Then the cumulative arrival process is bounded by

$$A_1(\tau,t) \le \lceil n(t-\tau)\rceil r \le r + nr(t-\tau) := E_1(t-\tau)\,, \quad (8)$$

where $\lceil x\rceil$ is the smallest integer larger than x. Eq. (8) states that the camera output is bounded by an affine arrival envelope E_1 with a burst of size r bits and a rate nr bits per second.

This arrival process is transmitted over the first wireless fading channel. The wireless channel provides time varying service equivalent to the instantaneous channel capacity, which can be determined as a function of the instantaneous signal–to–noise ratio (SNR) at the receiver. Using the block fading model and the instantaneous SNR γ_t at the receiver, the instantaneous channel capacity is given by $C(\gamma_t) = W\log(1+\gamma_t)$ bits/s/Hz, where W is the channel bandwidth, and the cumulative service offered to the incoming traffic is given by

$$S_1(\tau,t) = S(\tau,t) = W\sum_{u=\tau}^{t-1}\log(1+\gamma_u)\,. \quad (9)$$

Server 2: The arrival process to Server 2, the image processing node, is the departure process of Server 1, i.e., $A_2 = D_1$. The image processing node processes the frames in order, and it only starts feature detection on a frame once the frame is completely received. It thus introduces a latency, denoted by T_c which is the time required for a transmitted frame to be received successfully by the image processing node. Recall that the channel gain varies randomly and hence it offers randomly varying service, thus T_c is also random. Recalling that r bits per frame are generated by the camera, transmitting a single frame at the channel capacity rate, when the instantaneous SNR at the receiver is given by γ_i, will take

$$T_c = \min\{x \ge 0 : \sum_{i=1}^{x} C(\gamma_i) \ge r\} \quad (10)$$

seconds, where $C(\gamma)$ is the Shannon channel capacity limit, $\gamma_i = \bar{\gamma}|h+i|^2$, $\bar{\gamma}$ is the average received SNR and $|h_i|$ is the instantaneous fading channel gain at time slot i. Then we have

$$Pr(T_c \le x) \quad = \quad Pr(\sum_{i=1}^{x} C(\gamma_i) \ge r)$$

which can be approximated using the central limit theorem by a Gaussian distribution, $\mathcal{N}(\mu_C x, \sigma_C\sqrt{x})$, where μ_C and σ_C are the mean and standard deviation of the channel capacity

respectively. The probability mass function for T_c is then estimated as follows

$$f_{T_c}(t) = F_{\mathcal{N}}(r, \mu_C(t-1), \sigma_C\sqrt{t-1}) - F_{\mathcal{N}}(r, \mu_C t, \sigma_C\sqrt{t}), \quad (11)$$

where $F_{\mathcal{N}}(y, \mu, \sigma)$ is the distribution function for a Gaussian random variable with mean μ and a standard deviation σ evaluated at y.

Once a frame is received, the image processing node starts *feature detection*. The data remains in the buffer during detection and does not depart the node until feature extraction begins. The detection time θ depends on the input image size, on the number N of detected features per frame and on the processing speed. As all these parameters are constant, we can model the feature detection time θ as deterministic.

Detection is followed by *feature extraction*, which produces N descriptors, one at a time, and a total of r^* bits per frame. The feature extraction service rate ρ, measured in frames per time slot, depends on the processing speed and on N, the number of detected features per frame, and therefore can be considered constant. Observe that while the feature detection time θ is constant, in case of feature extraction it is the rate ρ that is constant.

Consequently, the image processing node offers a rate–latency service curve to the incoming traffic A_2 with rate ρr^* and latency $T_c + \theta$

$$S_2^*(\tau, t) = \rho r^* \cdot (t - \tau - (T_c + \theta)) \cdot 1_{\{t-\tau \geq T_c+\theta\}}, \quad (12)$$

where, $1_{\{.\}}$ is the indicator function.

Server 3: The arrival to Server 3, the second wireless link, is $A_3 = D_2$. The service curve of Server 3 is similar to that of Server 1,

$$S_3(\tau, t) = S(\tau, t), \quad (13)$$

and the departure process of this server is the departure process of the system.

B. Scaling for Flow Conservation

It is important to observe that due to in-network processing the flow conservation law does not hold: the image processor receives r bits per frame, but after feature extraction it transmits N feature descriptors per frame only, at a total rate of r^* bits per frame. Therefore,

$$A_3(\tau, t) = D_2(\tau, t) \leq E_3^*(t - \tau) = \phi E_3(t - \tau),$$

where $\phi = r^*/r$ is the *scaling factor*, E_3 is an envelope for the arrivals at server 3 if the frame size remain unchanged (i.e., r bits per frame) and the asterisk denote the new frame size r^* after feature extraction by server 2.

As Network Calculus builds on the flow conservation law, we need to ensure that flow is conserved in the model of the VSN despite scaling. Therefore, we have to scale down the traffic at the ingress A_1, together with the servers that operate on it (S_1 and S_2) by the scaling factor ϕ. The arrival function to the scaled version of the system is $A_1^* = \phi A_1$ and the service at Servers 1 and 2 is $S_1^* = \phi S_1$ and $S_2^* = \phi S_2$, respectively. For Server 3 we have $S_3^* = S_3$.

The implication of scaling is that the backlog bound obtained for the scaled model has to be scaled back to get the actual backlog for the real system. The delay bound computation is, however, not affected by scaling since both the arrival and service are scaled by the same factor.

V. Performance Bounds for Wireless Visual Sensor Networks

In this section, we will apply results from $(\min, \times)$ network calculus to obtain probabilistic bounds, considering the scaled version of the model. While scaling does not affect the delay bound, we scale then back the final results to derive the backlog bound of the original system.

A. SNR Arrival and Service Processes

To proceed with the analysis we have to transform the arrival and service processes presented in Section IV to the SNR domain. Recall that the service process for Server 1 in the bit domain in the scaled version of the model is $S_1^* = \phi S_1$, where S_1 is described by Eq. 9. Its SNR domain counterpart is described by the logarithm–free form

$$\mathcal{S}_1(\tau, t) = \prod_{u=\tau}^{t-1} (1 + \gamma_u)^{\phi W}. \quad (14)$$

Let $g_1(\gamma_u) = (1+\gamma_u)^{\phi W} = (1+\bar{\gamma}_u|h_u|^2)^{\phi W}$, where $|h_u|^2$ is the fading channel gain. Then

$$Pr(g_1(\gamma) \leq x) \quad = \quad Pr(|h|^2 \leq \frac{x^{\frac{1}{\phi W}} - 1}{\bar{\gamma}}).$$

Assuming Rayleigh fading, i.e., $|h|^2 \sim Exp(1)$ we have

$$Pr(g_1(\gamma) \leq x) \quad = \quad 1 - e^{-(\frac{x^{\frac{1}{\phi W}} - 1}{\bar{\gamma}})}. \quad (15)$$

Similarly, from Eq. (9) we obtain for server 3

$$\mathcal{S}_3(\tau, t) = \prod_{u=\tau}^{t-1} g_3(\gamma_u) = \prod_{u=\tau}^{t-1} (1 + \gamma_u)^{\frac{1}{W}}, \quad (16)$$

where $g_3(\gamma_u)$ has the distribution

$$Pr(g_3(\gamma) \leq x) \quad = \quad 1 - e^{-(\frac{x^{\frac{1}{W}} - 1}{\bar{\gamma}})}.$$

For the image processing server with service process S_2^* given by Eq. (12) we obtain

$$\mathcal{S}_2(\tau, t) = e^{S_2^*(\tau,t)} = e^{\phi\rho r \cdot (t-\tau-(T_c+\theta)) \cdot 1_{\{t-\tau \geq T_c+\theta\}}}. \quad (17)$$

For the traffic characterization, from Eq. (8), the scaled version of the arrival is given by $A_1^* = \phi A_1$. Then the SNR arrival process $\mathcal{A}$ is characterized by

$$\mathcal{A}(\tau, t) \leq e^{\phi E_1(\tau,t)} = e^{\phi r + \phi n r(t-\tau)}. \quad (18)$$

Note that an upper bound on the SNR arrival process is sufficient to obtain an upper bound on the network performance.

B. Mellin Transforms for the Arrival and Service Elements

In order to compute the desired probabilistic performance bounds in Theorem 1 we need to obtain the Mellin transform for the relevant SNR arrival and service process in the network.

Using the SNR arrival process description from Eq. (18), we have for any $s > 1$

$$\mathcal{M}_{\mathcal{A}}(s,\tau,t) \leq e^{(s-1)(\phi r+\phi nr(t-\tau))}\,. \tag{19}$$

1) Mellin transform for $\mathcal{S}_i$: From Eq. (14), the SNR service process for server 1 is given in terms of the function $g_1(\gamma)$ as

$$\mathcal{S}_1(\tau,t) = \prod_{u=\tau}^{t-1} g_1(\gamma)\,,$$

where the distribution of $g_1(\gamma)$, assuming i.i.d. Rayleigh fading, is given by Eq. (15). Hence, the Mellin transform of $g_1(\gamma)$ is given by

$$\begin{aligned}\mathcal{M}_{g_1(\gamma)}(s) &= E[(g_1(\gamma))^{s-1}]\\ &= e^{\frac{1}{\bar\gamma}}\bar\gamma^{\phi W(s-1)}\Gamma(1+\phi W(s-1),\frac{1}{\bar\gamma})\,,\end{aligned} \tag{20}$$

where, $\Gamma(a,b)$ is the incomplete Gamma function, i.e., $\Gamma(a,b) = \int_b^\infty x^{a-1}e^{-x}dx$. Then using the i.i.d. assumption and the product property of the Mellin transform [23] we obtain

$$\mathcal{M}_{\mathcal{S}_1}(s,\tau,t) = \underbrace{\Big(e^{\frac{1}{\bar\gamma}}\bar\gamma^{\phi W(s-1)}\Gamma\big(1+\phi W(s-1),\frac{1}{\bar\gamma}\big)\Big)}_{\alpha_1(s-1)}^{t-\tau}. \tag{21}$$

The SNR service process for server 2 is given by Eq. (17). For a fixed sample value t_c of the random variable T_c the Mellin transform for the conditional SNR service is given by

$$\mathcal{M}_{\mathcal{S}_2}(s,\tau,t) = e^{\phi\rho r(s-1)(t-\tau-(t_c+\theta))^+}\,, \tag{22}$$

where, $(X)^+ = \max(0,X)$.

The Mellin transform for $\mathcal{S}_3$ can be computed similarly to $\mathcal{S}_1$ and is given by

$$\mathcal{M}_{\mathcal{S}_3}(s,\tau,t) = \underbrace{\Big(e^{\frac{1}{\bar\gamma}}\bar\gamma^{W(s-1)}\Gamma\big(1+W(s-1),\frac{1}{\bar\gamma}\big)\Big)}_{\alpha_2(s-1)}^{t-\tau}, \tag{23}$$

where we omit the other arguments of α_i to simplify notation.

Now that we have a complete description of all service processes, we can compute a bound on the Mellin transform of the network SNR service process using the server concatenation property, Eq. (4), and Lemma 1 as follows

$$\begin{aligned}\mathcal{M}_{\mathcal{S}_{\rm net}}(s,\tau,t) &\leq \sum_{u_1=\tau}^{t}\mathcal{M}_{\mathcal{S}_1}(s,\tau,u_1)\mathcal{M}_{\mathcal{S}_2\otimes\mathcal{S}_3}(s,u_1,t)\\ &= \sum_{u_1=\tau}^{t}\mathcal{M}_{\mathcal{S}_1}(s,\tau,u_1)\sum_{u_2=u_1}^{t}\mathcal{M}_{\mathcal{S}_2}(s,u_1,u_2)\mathcal{M}_{\mathcal{S}_3}(s,u_2,t)\\ &= \sum_{u_1=\tau}^{t}\sum_{u_2=u_1}^{t}\mathcal{M}_{\mathcal{S}_1}(s,\tau,u_1)\mathcal{M}_{\mathcal{S}_2}(s,u_1,u_2)\mathcal{M}_{\mathcal{S}_3}(s,u_2,t)\,.\end{aligned} \tag{24}$$

Next we apply Theorem 1 to compute probabilistic delay and backlog bounds for the developed model.

C. Probabilistic Delay Bound

To compute a probabilistic delay bound for the system under investigation, we first evaluate the function $\mathsf{M}_{\rm net}(s,t+w,t)$ from Eq. (5) for $s \geq 0$, as given by the following lemma. We let $\alpha_i = \alpha_i(-s)$ to improve readability.

Lemma 3. *For the network in Fig. 2 and $s \geq 0$, the function $\mathsf{M}_{\rm net}(s,t+w,t)$, conditional on $T_c = t_c$, as $t \to \infty$ satisfies the following inequality*

$$\begin{aligned}\mathsf{M}_{\rm net}(s,t+w,t) \leq\ & \frac{\alpha_1^w e^{s\phi r}\left(\frac{1-(\frac{\alpha_2}{\alpha_1})^{t_c+\theta+1}}{1-\frac{\alpha_2}{\alpha_1}} - \frac{1-(\frac{1}{\alpha_1})^{t_c+\theta+1}}{\alpha_2(1-\frac{1}{\alpha_1})}\right)}{(1-\alpha_2^{-1})(1-\alpha_1 e^{-s\phi nr})}\\ &+\frac{e^{s\phi r}\left(\frac{1}{1-\alpha_1 e^{s\phi nr}} - \frac{\alpha_2}{\alpha_1(1-\alpha_2 e^{s\phi nr})}\right)\left(\frac{\alpha_2^{t_c+\theta+1}-1}{1-\frac{1}{\alpha_2}} + \frac{e^{-s\phi\rho r}}{1-\frac{e^{-s\phi\rho r}}{\alpha_2}}\right)}{(1-\frac{\alpha_2}{\alpha_1})e^{-s\phi nr(t_c+\theta-w+1)}}\\ &-\frac{e^{s\phi r}e^{-s\phi r(\rho-n)(t_c+\theta-w+1)}\left(\frac{1}{1-\alpha_1 e^{s\phi nr}} - \frac{e^{-s\phi\rho r}}{\alpha_1(1-e^{-s\phi r(\rho-n)})}\right)}{\alpha_2(1-\frac{e^{-s\phi\rho r}}{\alpha_1})(1-\frac{e^{-s\phi\rho r}}{\alpha_2})e^{-s\phi\rho r(t_c+\theta-w-1)}}\end{aligned}$$

whenever the following conditions are met

$$\begin{aligned}&\alpha_1\cdot e^{s\phi nr} < 1, \quad \alpha_2\cdot e^{s\phi nr} < 1\,,\\ &e^{-s\phi r(\rho-n)} < 1 \Longrightarrow n < \rho\,.\end{aligned} \tag{25}$$

Proof. Using Theorem 1 we compute

$$\begin{aligned}\mathsf{M}_{\rm net}(s,t+w,t) &= \sum_{v=0}^{t}\mathcal{M}_{\mathcal{A}}(1+s,v,t)\,\mathcal{M}_{\mathcal{S}_{\rm net}}(1-s,v,t+w)\\ &\leq e^{s\phi r}\sum_{v=0}^{t}\sum_{u_1=v}^{t+w}\sum_{u_2=u_1}^{t+w} e^{s\phi rn(t-v)}\alpha_1^{u_1-v}\\ &\quad\cdot e^{-s\phi\rho r(u_2-u_1-(t_c+\theta))^+}\alpha_2^{t+w-u_2}\,.\end{aligned}$$

Letting $v = t-v$, $u_1 = t-u_1$, $u_2 = t-u_2$ then letting $u_2 = u_1 - u_2$ and after rearranging the product, we arrive at

$$\begin{aligned}\mathsf{M}_{\rm net}(s,t+w,t) \leq\ & e^{s\phi r}\sum_{v=0}^{t}\sum_{u_1=-w}^{v}\sum_{u_2=0}^{u_1+w} e^{s\phi nrv}\alpha_1^{v-u_1}\\ &\cdot e^{-s\phi\rho r(u_2-(t_c+\theta))^+}\alpha_2^{u_1+w-u_2}\\ =\ & \alpha_2^w e^{s\phi r}\sum_{v=0}^{t}\sum_{u_1=-w}^{v}(\alpha_1 e^{s\phi nr})^v\left(\frac{\alpha_2}{\alpha_1}\right)^{u_1}\\ &\cdot\left[\sum_{u_2=0}^{\min(u_1+w,t_c+\theta)}\left(\frac{1}{\alpha_2}\right)^{u_2} + e^{s\phi\rho r(t_c+\theta)}\sum_{u_2=t_c+\theta+1}^{u_1+w}\left(\frac{e^{-s\phi\rho r}}{\alpha_2}\right)^{u_2}\right]\\ \leq\ & \alpha_2^w e^{s\phi r}\left[\sum_{v=0}^{t}\sum_{u_1=-w}^{t_c+\theta-w}(\alpha_1 e^{s\phi nr})^v\left(\frac{\alpha_2}{\alpha_1}\right)^{u_1}\sum_{u_2=0}^{u_1+w}\left(\frac{1}{\alpha_2}\right)^{u_2}\right.\\ &+\sum_{v=t_c+\theta-w+1}^{t}\ \sum_{u_1=t_c+\theta-w+1}^{v}\left(\alpha_1 e^{s\phi nr}\right)^v\left(\frac{\alpha_2}{\alpha_1}\right)^{u_1}\\ &\left.\cdot\left(\sum_{u_2=0}^{t_c+\theta}\left(\frac{1}{\alpha_2}\right)^{u_2} + e^{s\phi\rho r(t_c+\theta)}\sum_{u_2=t_c+\theta+1}^{u_1+w}\left(\frac{e^{-s\phi\rho r}}{\alpha_2}\right)^{u_2}\right)\right],\end{aligned}$$

where in the second step of the derivation, we use the fact that $e^{-s\phi\rho r(u_2-(t_c+\theta))^+}=1$, when $u_2 \leq t_c+\theta$ and $e^{-s\phi\rho r(u_2-(t_c+\theta))}$, otherwise. In the last step, we decompose the three sums and bound the first resulting term by extending the sum over u_2 to $t_c+\theta-w$ to simplify the resulting expression. This may loosen the obtained upper bound by a small margin.

Lemma 3 follows by letting $t\to\infty$ and evaluating the geometric sums successively. When the convergence conditions in Eq. 25 are satisfied, then the geometric progressions above converge as $t\to\infty$. □

The three convergence conditions in Eq. 25 define conditions for stable operation of the system. Indeed, the first two conditions are satisfied if the service provided by the two wireless channels, at the first hop and last hop respectively, is higher than the offered load. The last condition reduces to $n<\rho$, i.e., the frame rate is less than the intermediate image processor service rate (in frames per second).

The function $\mathsf{M}_{\rm net}(s,t+w,t)$ in Lemma 3 above[2] is obtained conditional on $T_c=t_c$ and for any $s>0$. Marginalizing over T_c gives the desired delay bound violation probability according to Theorem 1.

The marginal delay bound violation probability is given by

$$\begin{aligned}Pr\big(W(t)>w^{\varepsilon}_{\rm net}\big)&=\int_0^{\infty}Pr\big(W(t)>w^{\varepsilon}_{\rm net}|T_c=x\big)Pr(T_c=x)dx\\&\leq\int_0^{\infty}\inf_{s>0}\{\mathsf{M}_{\rm net}(s,t+w^{\varepsilon}_{\rm net},t)\}Pr(T_c=x)\,dx\,.\end{aligned}\tag{26}$$

The desired probabilistic delay bound is obtained by equating the right hand side of Eq. (26) to $\varepsilon^{w}_{\rm net}$ and optimizing over s. As a result, a bound on the violation probability of the end-to-end delay is given by

$$\varepsilon^{w}_{\rm net}=\min\left(1,\int_0^{\infty}\inf_{s>0}\{\mathsf{M}_{\rm net}(s,t+w^{\varepsilon}_{\rm net},t)\}Pr(T_c=x)dx\right),\tag{27}$$

where we used the fact that a probability is always bounded by 1.

Eq. (27) can be solved numerically to obtain the desired delay bound for a given violation probability $\varepsilon^{w}_{\rm net}$.

D. Probabilistic Backlog Bound

According to Theorem 1, the probabilistic backlog bound can be obtained by computing the function $\mathsf{M}_{\rm net}(s,t,t)$. This function is given by Lemma 3 when substituting $w=0$ as follows

$$\begin{aligned}\mathsf{M}_{\rm net}(s,t,t)\leq{}&\frac{e^{s\phi r}\left(\frac{1-(\frac{\alpha_2}{\alpha_1})^{t_c+\theta+1}}{1-\frac{\alpha_2}{\alpha_1}}-\frac{1-(\frac{1}{\alpha_1})^{t_c+\theta+1}}{\alpha_2(1-\frac{1}{\alpha_1})}\right)}{(1-\frac{1}{\alpha_2})(1-\alpha_1e^{-s\phi nr})}\\&+\frac{e^{s\phi r}\left(\frac{1}{1-\alpha_1e^{s\phi nr}}-\frac{\alpha_2}{\alpha_1(1-\alpha_2e^{s\phi nr})}\right)\left(\frac{\alpha_2^{t_c+\theta+1}-1}{1-\frac{1}{\alpha_2}}+\frac{e^{-s\phi\rho r}}{1-\frac{e^{-s\phi\rho r}}{\alpha_2}}\right)}{(1-\frac{\alpha_2}{\alpha_1})e^{-s\phi nr(t_c+\theta+1)}}\\&-\frac{e^{s\phi r}e^{-s\phi r(\rho-n)(t_c+\theta+1)}\left(\frac{1}{1-\alpha_1e^{s\phi nr}}-\frac{e^{-s\phi\rho r}}{\alpha_1(1-e^{-s\phi r(\rho-n))}}\right)}{\alpha_2(1-\frac{e^{-s\phi\rho r}}{\alpha_1})(1-\frac{e^{-s\phi\rho r}}{\alpha_2})e^{-s\phi\rho r(t_c+\theta-1)}}\,.\end{aligned}\tag{28}$$

The probabilistic backlog bound b^{ε}, conditioned on $T_c=t_c$ is given by

$$Pr\big(B(t)>b_{\rm net}|T_c=t_c\big)\leq e^{-s\,b_{\rm net}}\mathsf{M}_{\rm net}(s,t,t)$$

where $\mathsf{M}_{\rm net}(s,t,t)$ is given by Eq. (28). Then the marginal violation probability is given by

$$\begin{aligned}Pr\big(B(t)>b^{\varepsilon}_{\rm net}\big)&=\int_0^{\infty}Pr\big(B(t)>b^{\varepsilon}_{\rm net}|T_c=x\big)Pr(T_c=x)\,dx\\&\leq e^{-s\,b_{\rm net}}\int_0^{\infty}\mathsf{M}_{\rm net}(s,t,t)\cdot Pr(T_c=x)\,dx\,,\end{aligned}\tag{29}$$

for all $s>0$, where the distribution of T_c is given by Eq. (11). For a given violation probability $\varepsilon^{b}_{\rm net}$, equating the right hand side of Eq. (29) to $\varepsilon^{b}_{\rm net}$ and optimizing over $s>0$, the backlog bound is obtained as follows

$$b^{\varepsilon}_{\rm net}=\inf_{s>0}\left\{\frac{1}{s}\left[\log\int_0^{\infty}\mathsf{M}_{\rm net}(s,t,t)Pr(T_c=x)\,dx-\log\varepsilon^{b}_{\rm net}\right]\right\}.\tag{30}$$

By substituting Eq. (28) in Eq. (29), one can obtain the desired tail distribution for any backlog bound $b^{\varepsilon}_{\rm net}$, or in Eq. (30) to obtain the backlog bound for any violation probability $\varepsilon^{b}_{\rm net}$. However, this step may only be performed numerically due to the complexity of this expression. Note that this is the backlog bound for the scaled (by a factor of ϕ) model as we discussed in Sec. IV. The backlog bound for the original system is given by $b^{\varepsilon}_{\rm net}/\phi$ for $\phi\leq 1$. When $\phi<<1$, one expects most of the backlog to be at the first or the second server. The third server will see traffic that is reduced (significantly when $\phi<<1$) in intensity, and thus it will have a small backlog compared to the other two servers.

VI. Results and Discussion

In this section we evaluate the performance of wireless sensor network with in-network visual processing based on the analytic model that we derived and compare these results to simulation. We consider the three hop wireless visual sensor network with an intermediate image processing node as in Fig. 2. For this example, we assume a constant frame rate n and image size of $r=1.6$ Mb/frame generated by the visual sensor node. Pixel information and descriptors are transmitted over wireless links. The bandwidth of the wireless channel is

[2]Here, for simplicity, we do not show explicitly the dependence of the function $\mathsf{M}_{\rm net}$ on T_c.

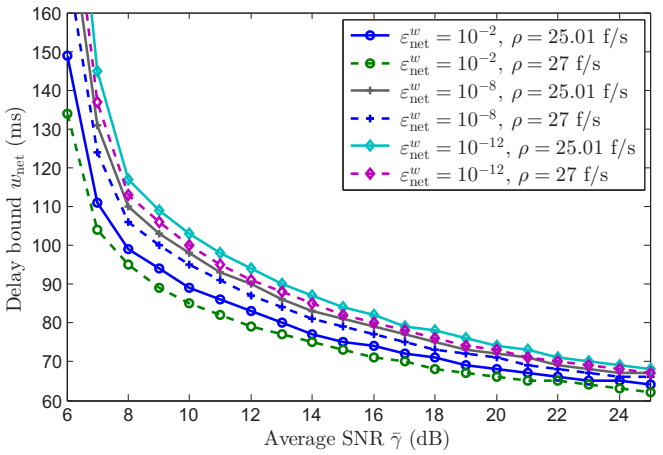

Fig. 3. End-to-end delay bound ($w^{\varepsilon}_{\text{net}}$) vs. average SNR ($\bar{\gamma}$) of wireless visual sensor network, for different image processing service rates (ρ) and different bound violation probability ($\varepsilon^w_{\text{net}}$), with $W = 22$ MHz, $r = 1.6$ Mb/frame, $\phi = 0.25$ and $n = 25$ frames/s.

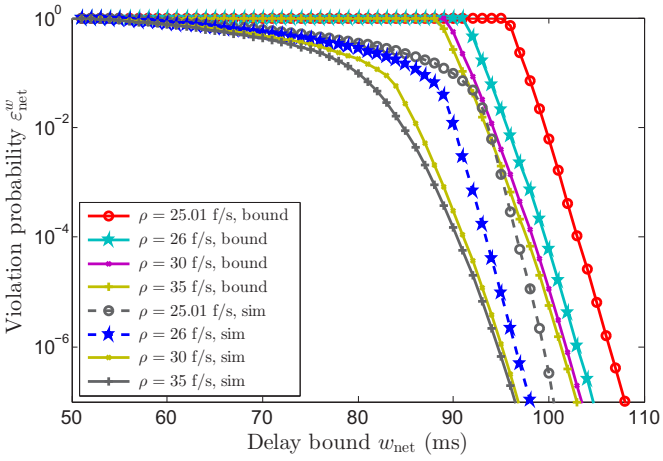

Fig. 4. Violation probability ($\varepsilon^w_{\text{net}}$) vs. delay bound ($w_{\text{net}}$) of wireless visual sensor network, for different image processing service rate (ρ), with $W = 22$ MHz, $r = 1.6$ Mb/frame, $\phi = 0.25$, $\bar{\gamma} = 8$ dB and $n = 25$ frames/s.

$W = 22$ MHz, which is typical of the IEEE 802.11 standard, and the average received SNR is $\bar{\gamma}$. For the intermediate image processor, the extraction ratio is chosen to be $\phi = r^*/r = 0.25$ and the detection time $\theta = 10$ ms. The probabilistic bounds are obtained for violation probabilities between 10^{-2} to 10^{-12}. We direct our attention to the delay bound since the two computed bounds have similar characteristics and will result in similar conclusions.

We also present simulation results for the network in Fig. 2. For the simulation, we model the arrival by a periodic source with period $\frac{1}{n}$ seconds and average $\frac{r}{n}$ bits/seconds. The service offered by the wireless channels is equal to their instantaneous channel capacity. We use a fluid flow arrival and service model. Time is slotted with intervals of 1 millisecond and we run the simulation for 10^{10} time slots.

Fig. 3 shows the analytic delay bound for increasing average

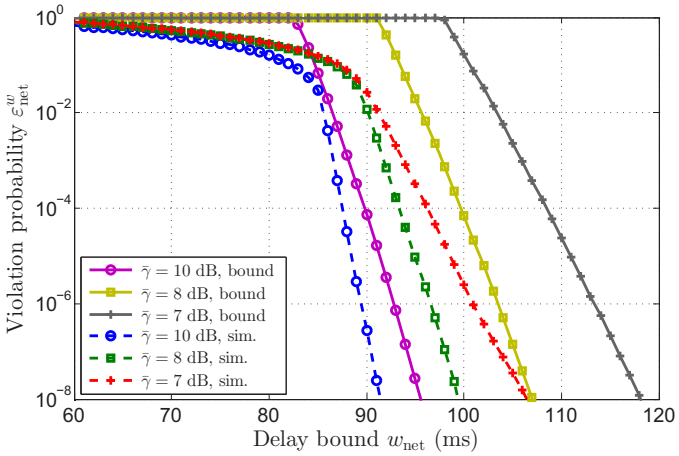

Fig. 5. Violation probability ($\varepsilon^w_{\text{net}}$) vs. delay bound ($w_{\text{net}}$) of wireless visual sensor network, for different average SNR ($\bar{\gamma}$), with $W = 22$ MHz, $r = 1.6$ Mb/frame, $\phi = 0.25$, $\rho = 26$ frames/s and $n = 25$ frames/s.

SNR of the wireless channel, for $n = 25$ frames/s, for different values of the feature extraction rate ρ at the intermediate image processing node and different delay bound violation probability $\varepsilon^w_{\text{net}}$. The graph shows that for higher $\bar{\gamma}$ the delay decreases slightly with increasing $\bar{\gamma}$, but the delay bound increases significantly at low SNR values, between 6 and 8 dB, where the system stability limit is reached. The system stability limit is reached also at $\rho = 25.01 f/s$, but the effect of low ρ value is less prominent compared to that of low $\bar{\gamma}$. The figure also shows that a smaller delay can be achieved if larger violation probability is tolerated. In any case, however, the delay bound stays at around 2 frame lengths even when the channel quality is good and the processing speed is high, which shows a significant effect of the in-network processing on the end-to-end performance.

Fig. 4 and Fig. 5 show the delay bound violation probability versus the delay bound for different values of image processing rate ρ and average SNR $\bar{\gamma}$ respectively, for $n = 25$, and compares these bounds to simulation results. The results show that the violation probability decays exponentially in w_{net}, and can be decreased significantly even with a limited increase of the delay. The quantitative result gives useful insight into the quality of service expected of this network, as it shows that a violation probability in the order of 10^{-6} can be achieved for a delay of 2-3 frame interarrival times even for low $\bar{\gamma}$.

The simulation results show that the obtained delay bounds reflect the general trends of the simulated system, for all ρ and $\bar{\gamma}$ values. They also show that the bounds overestimate the simulated end-to-end delay by at most 10% which is very reasonable considering the complexity of the problem, the analytic complexity in obtaining these performance measures using other methods, e.g., queuing theory and the required simulation time when evaluating rare violation events.

Fig. 6 shows the delay bound violation probability versus the feature detection time θ for different values of $\bar{\gamma}$, ρ and w_{net}. The results show that under stable operation the

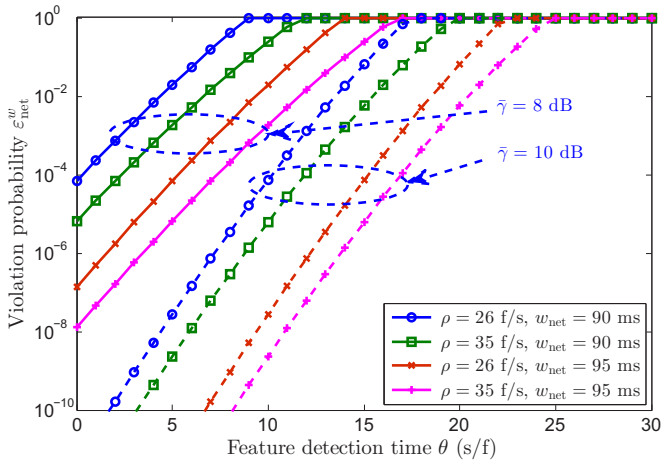

Fig. 6. Violation probability ($\varepsilon^w_{\rm net}$) of wireless visual sensor network vs. feature detection time (θ), for different delay bounds ($w_{\rm net}$), image processing service rate (ρ) and $\bar\gamma$, with $W = 22$ MHz, $r = 1.6$ Mb/frame, $\phi = 0.25$ and $n = 25$ frames/s.

violation probability increases almost exponentially in θ, and the rate depends mainly on $\bar\gamma$. It also shows similar trends for increasing $\bar\gamma$ as well as ρ. The relative position of the different curves in Fig. 6 suggests that once in a stable operation regime, increasing one of the service element's average service rate, e.g., $\bar\gamma$ or ρ, may efficiently compensate for the degradation in the other one.

VII. Conclusion

We presented a performance analysis for wireless sensor networks with in-network image processing capability that takes into consideration wireless fading channel properties as well as the in-network image processor's properties. The analysis extends a newly developed stochastic network calculus for wireless fading channels with the ability to model processing nodes, through appropriate scaling of the arrival processes and the service times. Our analysis highlights various trends and tradeoffs in the performance of visual sensor networks, which can be used to enhance quality of service and mitigate network variability. Most importantly, the analysis shows that for a reasonable set of network parameters, a moderate end-to-end delay can be achieved with very low violation probability. The model can be extended to capture processing nodes in series and in parallel, this subject of our ongoing work. The proposed method could be utilized to model other in-network processing systems with distributed computing, sensing or control.

References

[1] A. Redondi, L. Baroffio, M. Cesana, and M. Tagliasacchi, "Compress-then-analyze vs. analyze-then-compress: Two paradigms for image analysis in visual sensor networks," in *IEEE Int. Workshop on Multimedia Signal Processing (MMSP)*, Sep. 2013.

[2] A. Redondi, L. Baroffio, A. Canclini, C. M., and T. M., "A visual sensor network for object recognition: Testbed realization," *IEEE/EURASIP Conf. on Digital Signal Processing (DSP)*, Jul. 2013.

[3] E. Eriksson, G. Dan, and V. Fodor, "Real-Time Distributed Visual Feature Extraction from Video in Sensor Networks," in *IEEE International Conference on Distributed Computing in Sensor Systems (DCOSS)*, May 2014.

[4] M. A. Khan, G. Dan, and V. Fodor, "Characterization of SURF interest point distribution for visual processing in sensor networks," in *Proc. of IEEE/EURASIP Conf. on Digital Signal Processing (DSP)*, Jul. 2013.

[5] Y. Jiang and Y. Liu. *Stochastic network calculus*. Springer, 2008.

[6] C.-S. Chang. *Performance guarantees in communication networks*. Springer Verlag, 2000.

[7] H. Al-Zubaidy, J. Liebeherr, and A. Burchard. A (min, x) network calculus for multi-hop fading channels. In *Proc. IEEE Infocom*, 2013.

[8] M. C. Stefan Leutenegger and R. Y. Siegwart, "Brisk: Binary robust invariant scalable keypoints," in *IEEE International Conference on Computer Vision (ICCV)*, 2011.

[9] E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105–119, Jan. 2010.

[10] L.-Y. Duan, X. Liu, J. Chen, T. Huang, and W. Gao, "Optimizing jpeg quantization table for low bit rate mobile visual search," in *IEEE Visual Communications and Image Processing (VCIP)*, Nov 2012.

[11] J. Chao, H. Chen, and E. Steinbach, "On the design of a novel JPEG quantization table for improved feature detection performance," in *Proc. of IEEE International Conf. on Image Processing (ICIP)*, Sep 2013.

[12] V. R. Chandrasekhar, S. S. Tsai, G. Takacs, D. M. Chen, N.-M. Cheung, Y. Reznik, R. Vedantham, R. Grzeszczuk, and B. Girod, "Low latency image retrieval with progressive transmission of CHoG descriptors," in *ACM Multimedia workshop on Mobile cloud media computing*, 2010.

[13] H. Jegou, M. Douze, and C. Schmid, "Product quantization for nearest neighbor search," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, no. 1, pp. 117–128, Jan 2011.

[14] A. Redondi, L. Baroffio, J. Ascenso, M. Cesana, and M. Tagliasacchi, "Rate-accuracy optimization of binary descriptors," in *IEEE International Conference on Image Processing*, 2013.

[15] A. Redondi, M. Cesana, and M. Tagliasacchi, "Rate-accuracy optimization in visual wireless sensor networks," in *Prof. of the IEEE ICIP*, 2012.

[16] D.-N. Ta, W.-C. Chen, N. Gelfand, and K. Pulli, "SURFTrac: Efficient tracking and continuous object recognition using local feature descriptors," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition,*, Jun. 2009.

[17] G. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (hevc) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668, 2012.

[18] L. Baroffio, M. Cesana, A. Redondi, S. Tubaro, and M. Tagliasacchi, "Coding video sequences of visual features," in *Proc. of IEEE Intl. Conf. on Image Processing (ICIP)*, 2013.

[19] F. Ciucu, "Non-asymptotic capacity and delay analysis of mobile wireless networks," in *Proc. of ACM Sigmetrics*, pp. 359–360, June 2011.

[20] M. Fidler, "A network calculus approach to probabilistic quality of service analysis of fading channels," in *Proc. IEEE Globecom*, pp. 1–6, Nov. 2006.

[21] K. Mahmood, A. Rizk, and Y. Jiang, "On the flow-level delay of a spatial multiplexing MIMO wireless channel," in *Proc. IEEE ICC*, pp. 1–6, June 2011.

[22] G. Verticale and P. Giacomazzi, "An analytical expression for service curves of fading channels," in *Proc. IEEE Globecom*, pp. 635–640, Nov. 2009.

[23] B. Davies, *Integral transforms and their applications*, Springer-Verlag, NY, 1978.

[24] M. Fidler, "An end-to-end probabilistic network calculus with moment generating functions," in *Proc. of the IEEE IWQoS*, pp. 261–270, June 2006.

[25] F. Ciucu, J. Schmitt, and H. Wang, "On expressing networks with flow transformations in convolution-form," in *Proc. of IEEE INFOCOM*, pp.1979-1987, April 2011.

[26] M. Fidler and J.B. Schmitt, "On the way to a distributed systems calculus: an end-to-end network calculus with data scaling," *ACM SIGMETRICS Perform. Eval.*, pp.287-298, June 2006.

[27] E. Eriksson, G. Dán, and V. Fodor, "Prediction-based load control and balancing for feature extraction in visual sensor networks," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2014.