Simulation Modelling Practice and Theory 87 (2018) 120–137

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat

SenseVid : A traffic trace based tool for QoE Video transmission assessment dedicated to Wireless Video Sensor Networks

Moufida Maimour*,[a],[b]

[a] Université de Lorraine, CRAN, UMR 7039 Campus Sciences, BP 70239, Vandoeuvre-lès-Nancy Cedex 54506, France
[b] CNRS, CRAN, UMR 7039, France

ARTICLE INFO

Keywords:
Wireless Video Sensor Networks (WVSN)
Video Quality Evaluation
SenseVid
QoS
QoE
Energy

ABSTRACT

Video applications are being a key component to enhance traditional wireless sensor networks (WSN) applications. As a result, sensor network researchers need adequate and easy to use tools to assess the performances of their proposals. The existing transmission evaluation tools either consider video sequences along with codecs that are unsuitable for WSNs or make use of low cost compression methods for still images without inter-frame coding required for efficient video transmission. In this paper, we present SenseVid, an open source video transmission and evaluation tool that considers WSN specific characteristics. Besides low energy intra-frame compression based on fast pruned discrete cosine transform (DCT), a low complexity inter-frame encoding is provided to allow efficient support of video flows. A configurable fine-grain energy model is provided where both video capture and encoding cost are accounted for on a per frame basis. Video traffic differentiation based on priority levels is also provided. SenseVid adopts the video traffic traces approach, allowing its use in any simulation or real testbed environment. Using SenseVid, the user is able to reconstruct the received video considering lost packets during its transmission as well as estimating the achieved quality of service (QoS) and quality of experience (QoE).

1. Introduction

The emergence of low-cost and low-power visual modules [1–3] fostered the development of Wireless Video Sensor Networks (WVSN) and paved the way to a plethora of applications mainly in monitoring and surveillance fields. WVSNs generate unique challenging problems as the capture, the encoding and the transmission of video flows are resource-hungry while the underlying network is constrained in terms of energy, bandwidth, processing and storage means. As a result, a significant effort has to be made in the design and development of new algorithms and protocols where the network limited resources are saved while providing a good quality of service (QoS) to the end user.

Network researchers usually make use of discrete event network simulators such as ns2 [4] and OMNeT++ [5] to develop and evaluate their algorithms and protocols. Compared to a real WSN testbed, network simulators are good means to easily test protocols in a large scale with very limited cost. However, with respect to WVSN, pure simulation testing may not accurately reproduce real-life scenarios and may not provide appropriate metrics to assess proposed protocols and/or video encoding techniques. Metrics such as packet loss ratio, throughput, delay and jitter remain pure network related metrics that provide only limited insight into the video quality perceived by a human and do not reflect its associated subjective factors. One step forward to achieve more accurate

* Correspondence to Université de Lorraine, CRAN, UMR 7039 Campus Sciences, BP 70239, Vandoeuvre-lès-Nancy Cedex 54506, France.
E-mail address: moufida.maimour@univ-lorraine.fr.

https://doi.org/10.1016/j.simpat.2018.06.006
Received 30 January 2018; Received in revised form 18 May 2018; Accepted 29 June 2018
Available online 30 June 2018
1569-190X/

assessment of network proposals is to be able to estimate the received image/video quality. Human perception can be evaluated using Quality of Experience (QoE) metrics such as structural similarity index (SSIM) [6]. This requires simulated experiments to be performed using actual encoders, decoders and video data. EvalVid [7] made it possible by providing a set of tools, that allows researchers to carry out simulations of real video sequences transmission. EvalVid takes the video traffic trace approach to characterise a video sequence where trace files with limited size are produced rather than using actual video bit stream [8].

In the context of WVSN, Authors of [9] proposed to integrate EvalVid along with WiSE-MNet [10] in the Castalia simulator [11]. However, standard video coding techniques based on motion estimation algorithm such as MPEG-4, H.263 or H.264 supported by EvalVid are not suitable for sensor nodes [12]. Even JPEG's low complexity still image compression algorithm is not very beneficial in terms of power consumption. This is mainly due to the discrete cosine transform (DCT) stage which consumes at least 60% of the whole power encoder [13]. Therefore, efficient compression algorithms, in terms of power dissipation, are mandatory to handle video flows. To fit WSN needs, WVSN model [14] integrates the low-cost compression method proposed in [15] into the Castalia simulator source code. However, the image bit stream adopted makes the WVSN model usable only with the Castalia simulation environment. Sim-LIT [16] is another tool targeted to image transmission designed to assess *block interleaving* algorithms using the TOSSIM simulation environment [17]. Both WVSN model and Sim-LIT are inefficient for video transport since they are only concerned with intra-coded still images. EvalVSN [18] based on Matlab, adopts the traffic trace method along with an inter-frame coding to consider video flows. However, it suffers from a heavy execution time as well as poor performances due mainly to the adopted inter-frame coding and the substantial length of the generated traffic traces.

The existing transmission evaluation tools either consider video sequences along with codecs that are unsuitable for WSN or make use of low cost compression methods for still images without inter-frame coding required for efficient video transmission. In this paper, we propose *SenseVid* (publicly available at [19]), a video transmission and evaluation tool targeted to WVSN by implementing low complexity inter-frame coding as well as reduced cost DCT intra-frame coding. SenseVid adopts the video traffic trace approach, allowing its use in any simulation or real testbed environment provided that a trace-based transmission application is available. Currently, we provide an application module for the Contiki operating system [20] along with tools and scripts to automate the evaluation process. Other similar application modules can easily be developed for other environments. We implemented the PSNR (Peak Signal to Noise Ratio) and the SSIM (Structural SIMilarity) [6] as QoE metrics to assess the quality of both the encoded video before transmission and the received video with respect to the transmitted one. Finally, a configurable fine-grain energy model is provided where both the cost of the capture and the encoding of the video are accounted for on a per frame basis.

The remainder of this paper is organised as follows. Section 2 overviews the exiting tools targeted to image or video transmission and evaluation along with some background on DCT-based compression algorithms that are suitable to WVSN. The SenseVid architecture and its main functionalities are presented in Section 3 while Section 4 summarises main numerical results. Section 5 gives use scenarios in a simulation environment (cooja [21]) as well as in a real sensor testbed (IoT-LAB [22]). Section 6 concludes the paper with some future directions.

2. Background and related work

EvalVid [7] is a popular framework used in the network research community to evaluate video transmission through QoE metrics such as PSNR in addition to other classic QoS metrics like delay, jitter and loss rate. Evalvid takes the traffic trace approach where, based on an input video and a given codec, a video trace file is generated. Each line of the trace file gives mainly, for each frame, its type and its size. A sender trace file is also generated with mainly the transmission time and the size of packets to be sent. The sender trace file can be used by any simulator or real application to emulate the transmission of a given video. Based on the receiver trace file that gives the list of the received packets, Evalvid rebuilds the received video and computes the achieved QoS and QoE metrics. Application modules that allow the use of EvalVid trace files with ns2 and NS3 have been proposed in [23] and [24] respectively.

Since EvalVid was developed to be used in less constrained networks, no energy model is provided. Moreover, the supported encoders, namely, MPEG-4, H.263 or H.264 are resource-hungry and as a result are not adapted to constrained networks such as WSN. These codecs concentrate the most computational complexity at the encoder while video sensors have neither sufficient processing power nor enough energy to perform complex compression algorithms. Nevertheless, compression is still needed to reduce the amount of transmitted data and thus save bandwidth, transmission time and energy dissipation. Hence, efficient compression algorithms, in terms of both processing and energy requirements, are mandatory to handle video transmission in WSN. Compared to lossless compression, lossy one is more suitable to WSN since it allows higher compression ratios which consume less communication resources. Transform-based compression algorithms such as DCT-based ones are generally preferred over non-transform-based ones due the fact that their encoder is less complex [25].

The most popular compression scheme based on DCT is JPEG [26]. The image is divided into small 8×8 blocks on which the DCT is applied to separate the high and low frequency information. To compress the image, a quantisation is applied on each DCT block to discard the coefficients with the least information (highest frequency components). A zigzag scan is then applied to reorder coefficients from lower to higher frequencies. Finally, an entropy coding is applied. It generally consists in a run-length encoding (RLE) followed by a Huffman or arithmetic coding. The DCT transform results in a low memory usage since it operates on small blocks and allows acceptable compression ratios. However, as reported in [13], DCT computation consumes at least 60% of the whole power encoder. Even more, authors of [27] have shown that compressing, using JPEG, and transmitting an image consumes more energy than simply transmitting the uncompressed image.

Research effort has gone into the development of methods to decrease the DCT computational complexity and thus reducing its energy footprint. *Fast DCT* computation algorithms mainly eliminate redundant operations through factorising the DCT matrix [28]