ISM2015 - IEEE International Symposium on Multimedia, Miami, FL, USA, 14-16 Dec 2015

Go Green with EnVI: The Energy-Video Index

Oche Ejembi
School of Computer Science
University of St Andrews, Fife, UK
Email: ooe@st-andrews.ac.uk

Saleem N. Bhatti
School of Computer Science
University of St Andrews, Fife, UK
Email: saleem@st-andrews.ac.uk

***Abstract*—Video is the most prevalent traffic type on the Internet today. Significant research has been done on measuring user's Quality of Experience (QoE) through different metrics. We take the position that energy use must be incorporated into quality metrics for digital video. We present our novel, energy-aware QoE metric for video, the *Energy-Video Index (EnVI)*. We present our EnVI measurements from the playback of a diverse set of online videos. We observe that 4K-UHD (2160p) video can use ∼30% more energy on a client device compared to HD (720p), and up to ∼600% more network bandwidth than FHD (1080p), without significant improvement in objective QoE measurements.**

I. Introduction

Studies on the energy usage and carbon footprint of ICT are increasingly important as the use of ICT grows globally. Projections from a study by Lambert et al. [1] suggest that the use of ICT (communication networks, client devices and data centres) would require over 1100 TWh of electricity in 2015, or 5.5% of all the electricity consumed globally. This usage grows at a rate of 6% annually, compared to the overall annual growth rate of 3% for electricity use globally across all sectors. Meanwhile, video is the most prevalent traffic type on the Internet today: up to 79% of all traffic [2].

With the prevalence and continued growth of video usage, a significant amount of work has focused on studying and quantifying the *quality* of video. The Quality of Service (QoS) requirements of networks to support video are now well understood. However, it is generally agreed that there is no single, universal metric for measuring video *Quality of Experience (QoE)* [3]. Various QoE metrics are used, and QoE is highly dependant on the context of measurement. For example, there are several subjective measurements based on real human users: ratings such as Double Stimulus Impairment Scale (DSIS) and Absolute Category Rating (ACR) [4]. Additionally, objective evaluations for video quality, such as Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Metric (SSIM), may be used to generate reproducible results, relatively quickly and cheaply.

We take the position that *'greenness'* – energy usage – should be used as a quality metric for video. It is important that the energy usage if video is better understood as its use increases. We demonstrate an empirical approach, and present measurements of energy- usage and QoE of video playback (decoding). We also discuss the Energy Video-Index (EnVI) of quality, our objective metric which takes energy into consideration for the comparison of video quality.

II. Energy as a QoE metric for video

'Greenness', or sustainability, has been proposed as an additional non-functional requirement for modern ICT systems [5]. Penzenstadler et al. [6] authors take the position that sustainability or greenness must become a prominent feature in future software engineering practices, much like safety and security are today.

The same argument can be applied for both the software (e.g. software codecs, video players etc.) and hardware components (CPUs, GPUs, networking devices, storage etc.) of video systems, which are now very widely used. As the technology and infrastructure capable of supporting video continues to mature, the demand for video by consumers and its supply by providers will continue to grow [7]: for example, wide-scale introduction of 4K Ultra High Definition (UHD) video. In this paper, we show how 4K video can consume up to 30% more energy at the client than HD (720p), and up to 6 times more network bandwidth than Full HD (1080p), for only a fractional improvement in *objective* QoE (in Section IV).

The proposed 'Green Metadata' standard draft from the Moving Pictures Experts Group (MPEG) discusses how metadata from encoding, decoding, streaming and selecting video may be leveraged towards savings in energy [8]. Our work is complimentary to those efforts, as comprehensive measurement and benchmarking of existing software and systems is needed to generate and fully utilise this generated meta-data.

III. Energy benchmarking for video

Various stakeholders may have different motivations and incentives to be energy-aware when using video. For instance, mobile device users may want to conserve their battery. On the other hand, network administrators may want to conserve network resources, and ensure fairness for other users and applications. Video-on-Demand (VoD) providers will want to keep operating expenditure (OPEX) low while providing excellent service. However, to enable energy-awareness, there is a need for all parties to gain both quantitative and qualitative insight into the energy usage of diverse video *configurations*. In this context, a *configuration* is a set of properties a video streaming session could have e.g. codec type, picture size, software library, device type, network protocol etc.

Currently, there are no software tools or frameworks that offer this level of flexibility and specificity in measurement of video. We are developing a cross-platform, extensible, energy benchmarking and measurement tool for video. To the best of our knowledge, there are very few benchmarking tools in existence that use video playback as the workload for benchmark measurements. We highlight three existing tools, Microsoft Assessment Console [1], Phoronix Test Suite [2] and PCMark [3] that are relatively well known. However, none of these benchmark tools focus specifically video, but rather try to assess a full suite of performance indicators – of which video playback happens to be one.

A. The Energy-Video Index (EnVI)

In our previous work [9], we introduced the early concepts of a simple, intuitive, energy-based quality metric for video. We denoted the amount of energy required to playback one second of a given video as P_{dv} (with units of J/s_v – Joule per second of video – which is analogous to a Watt). From measured values for P_{dv}, we can then calculate E_v, the normalised energy metric. This is derived via a linear transformation into the range [0,1] as shown in Equation 1.

$$P_{dv} = \frac{\text{energy usage for video decoding}}{t_v} = P_a - P_{idle} \quad (1)$$

$$E_{dv} = \frac{P_{idle}}{P_a} \quad (2)$$

$$= \frac{P_{idle}}{P_{dv} + P_{idle}} \quad (3)$$

with P_a, P_{dv}, P_{idle} and t_v are as defined in Table I. Note that the denominator for Eqn. 3 is simply the mean power measured during the decoding process.

TABLE I: Definitions for E_{dv} (Eqn. (3)).

P_{dv}	mean energy usage per second during video decoding [J/s_v]
P_a	measured mean *system* power usage during video encoding or decoding [W]
P_{idle}	mean system power usage when idle [W]
t_v	the duration of the video sample [s]

We have defined a metric which incorporates energy usage of a video and any measure of quality [9] – which we refer to as the Energy-Video Index (EnVI) of quality, Q_M, as defined in Eqn. 4. This combines several quality metrics, including energy usage (E_{dv}), as a single value. This then makes it easy and convenient to apply the metric as an input parameter into optimisation functions or dynamic adaptation algorithms, e.g. [10], which uses SSIM scores as input.

$$Q = \alpha E_{dv} + (1-\alpha) Q_v \quad (4)$$
$$0 \leq Q_v \leq 1 \;\;,\;\; 0 \leq E_{dv} \leq 1$$
$$0 \leq Q \leq 1 \;\;,\;\; 0 \leq \alpha \leq 1$$

where each of Q, Q_v and E_{dv} have a value of zero to signify a 'worst' level, and a value of one to signify the 'best' level, with the meanings summarised in Table II. Q_v is a normalised

[1]https://msdn.microsoft.com/hh825384.aspx
[2]http://www.phoronix-test-suite.com/?k=documentation
[3]http://www.futuremark.com/benchmarks/pcmark

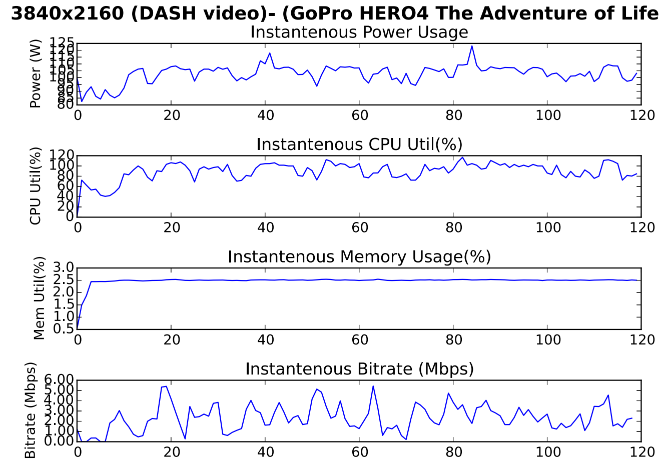


Fig. 1: Output for a single benchmarking session, showing a section of measurements for power, CPU, memory and ingress network usage.

QoE metric. Any standard QoE metric may be used for Q_v as long as it is possible to apply a linear transform so it is in the range [0,1]. The SSIM metric for instance, naturally exists in this form, so it is easy to use without any modification.

TABLE II: Definitions for Q (Eqn. (4)).

Q_v	a quality metric derived from a standard quality measure for a video stream
E_{dv}	a normalised energy metric derived from the measured energy usage for a video playback (decoding) session
α	weight (use 0.5 for equal weighting in this case)

IV. Experiments

We describe various experiments we have performed to show the utility of EnVI. These experiments also present interesting insights towards supporting green video, as well as potential energy-QoE trade-offs that could be made for adaptation to energy energy-efficient video playback.

A. Methodology

We ran our benchmark on two identical hardware configurations but with different operating systems installed. These configurations are summarised in Table III. For the workloads, we selected 10 popular videos on YouTube that are available in the 4K UHD (3096x2160 – 2160p) resolution. We found that YouTube makes such videos available in up to 32 different formats, identified by a 2 or 3 digit value referred to as an *itag*. For the purposes of this experiment, he have only used itags corresponding to the following picture sizes: 320p, 480p, 720p, 1080p and 2160p – of which there are 16 different configurations available. We benchmarked the first 120s of each of these videos. We have found from experience that this duration of video is sufficient to characterise the resource usage of the video. Additionally, a significant proportion of YouTube videos are less than 200 seconds in length [11].

TABLE III: Summary of the hardware and software used in our evaluation.

OS	Ubuntu Linux 14.04, Microsoft Windows 8.1
Workstation	Intel X99 chipset, Intel i7-4920 CPU @ 4GHz, 16GB DDR4 DRAM
GPU	Nvidia GTX 960, 2GB
Video application	libVLC version 2.2.2
Codecs	H.264 (libx264 version r2538) VP8/VP9 (libvpx version 1.4.0.8)
Power meter	Voltcraft VC870 (with USB interface)

B. Results

In summary, our results show that, as might be expected, videos with bigger picture sizes do use more energy for playback. However, the additional energy usage is not proportional to the difference in picture size, and the larger picture size does not necessarily give a higher *objective* QoE (using SSIM).

1) Single Run Results: During a benchmarking session, our benchmarking tool captures various measurements for the video workload. At the end of a single benchmarking run, these measurements are collated, and our tool is able to automatically generate graphical visualisations of measurements taken for that run as a time-series. This makes it possible to examine the points within the video that consumed the most or the least amount of resources.

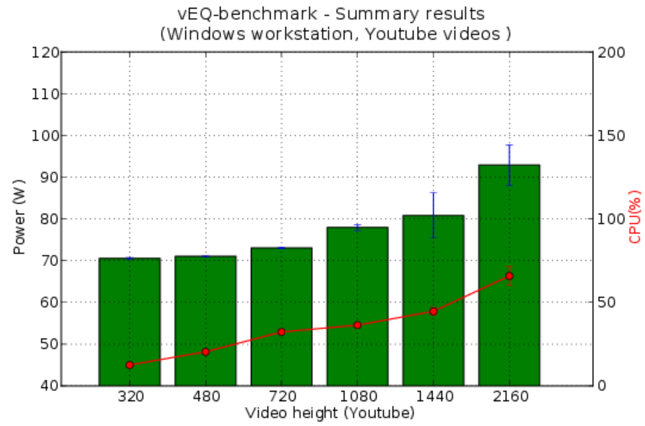
2) Summary Results: In Figure 2 are aggregate values of all 10 video samples. There are are no real surprising results considering the picture sizes: larger picture sizes consume more CPU power, more energy and more network capacity. However, it is interesting to see the *relative* differences between picture sizes (left to right across the columns in each of Figures 2a and 2b), and also between Linux and Windows 8.1 (for the same columns in Figures 2a and 2b).

3) Comparing HD, FHD and UHD formats: In Figure 3, we show a more fine-grained result for High Definition (720p), Full High Definition (1080p) and Ultra High Definition (2160p or 4K-UHD) across the 10 videos used in the experiment. We can observe the expected trend, as the larger formats require more energy for playback. However, this difference is significant, e.g. for the 'Samsung Galaxy S' video[4]', we observe a difference of $26J/s_v$ or 36% (HD compared to UHD) and $18J/s_v$ or 22% (HD compared to FHD).

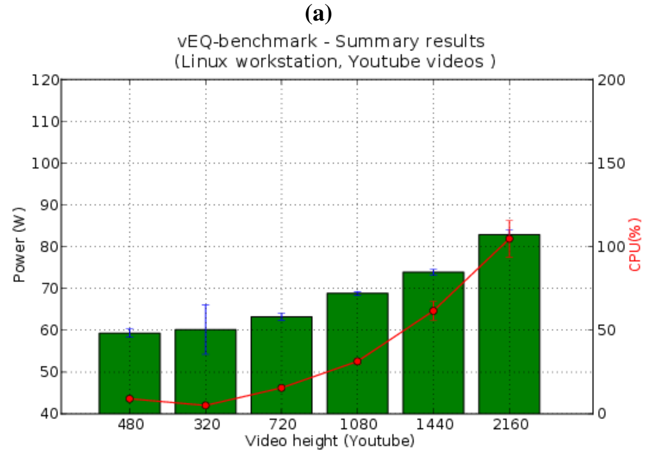
4) Quality Assessment: We measured the (relative) objective quality of the test videos, by using the FFMPEG tiny_ssim tool[5]. We performed an objective quality assessment using the well known Y-SSIM metric for the 10 YouTube videos used in our experiment. Y-SSIM is often described as having a very good approximation to human perception. We have made our SSIM measurements *relative* to the best quality of video available (i.e. the 4K video at the highest bitrate corresponding to an itag value of 272). We observed that all the videos show a very strong similarity to the best quality video (>0.95). These results indicate that most human viewers might not see a strong difference in these videos. Of course, in reality there would be several other factors and considerations to be made, which the SSIM metric does not capture. These could include screen size, distance of the viewer from the screen, brightness and ergonomics (amongst others). These human factors are often taken into consideration during standardised, subjective

[4]https://www.youtube.com/watch?v=onXpKXbnbE0

[5]https://ffmpeg.org/ffmpeg-filters.html#ssim



(a)



(b)

Fig. 2: Figures 2a and 2b show summary values of system power usage by YouTube videos during the benchmarking exercise for the the Windows and Linux machine configurations described in Table III. The results are grouped as the video height (vertical picture size), as the widths of YouTube videos can vary. (The line joining points is a visual aid only.)

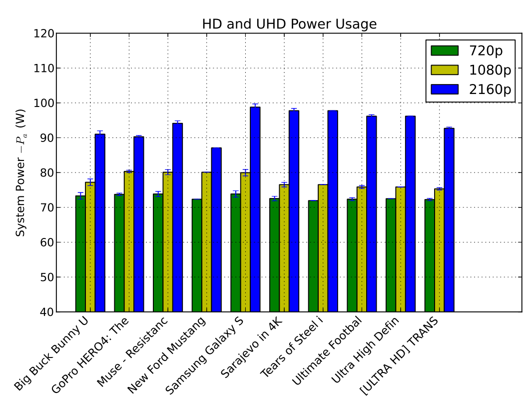


Fig. 3: System power usage (P_a) video playback on the Windows machine. Each bar is the mean of 10 runs with 95% confidence intervals shown.

measurements, e.g. [4].

5) EnVI: Energy-Video Index: Having measured the differences in energy usage and objective quality for the videos, it is then possible for us to objectively evaluate the trade-off between quality and energy for our sample videos. From Section III-A, given the *running* idle power of a system being benchmarked (P_{idle}), and the power usage during playback of a given video (P_a), we can calculate our normalised, measurement-based energy metric E_v, using Eqn. 3, and combine this with an objective quality metric, such as SSIM, to get an EnVI value (Q) using Eqn. 4.

We demonstrate this in n Figure 4, with an example video from our experiment, 'Go Pro 4K', evaluated on Windows 8.1. For each mean value of P_{dv}, we calculated the E_{dv} value, using Eqn. 3. Then, using the SSIM value for that format as the value for Q_v, we calculate the EnVI value, Q using Eqn. 4. To show the potential impact of user preferences, we use varying weights (α), which allows weighting between energy and QoE. For $\alpha = 0$, only the video quality metric, SSIM, is taken into consideration), and for $\alpha = 1$, only our energy metric, E_{dv}, is under consideration. Thus, we are able to observe the trade-off 'envelope' between objective quality and energy with E_{dv}. For $\alpha = 0.25$, the EnVI value remains relatively constant regardless of picture size, and so offers a 'locus of control' for improving energy efficiency, by trading off a small (possibly negligible) amount of picture quality (e.g. compared to $\alpha = 0$).

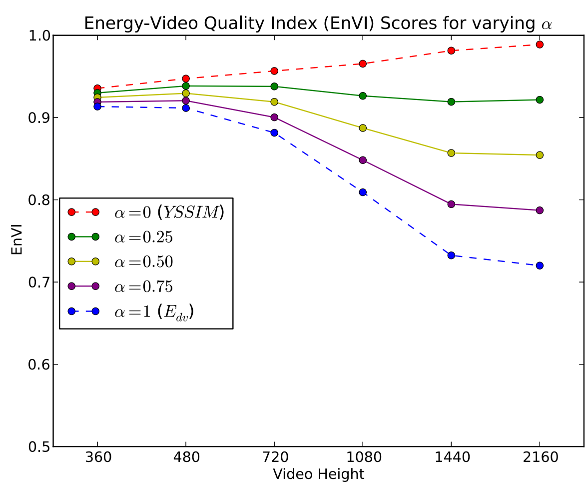


Fig. 4: Example Energy-Video Index (EnVI) value for a single YouTube video ('Go Pro 4K'). Values on each coloured line represent the Q score at various values of α. The lines are a visual aid and do not infer interpolation to intermediate picture sizes. For a single line (i.e. a given value of α), where the value of EnVI is closest to 1, that corresponding picture size is the optimum trade-off between QoE and energy.

Indeed, if suitable incentives can be found to enable and encourage appropriate selection of video streams by users, service providers and/or network providers could save a significant amount of energy [12].

V. Conclusion and Future Work

We discussed the need for awareness of the energy usage of digital video delivered over the Internet. We then presented our Energy Video Index (EnVI) as a combined measure of video quality of experience (QoE) and energy usage. We followed this with an evaluation of EnVI, through results gathered from our experimental test-bed. We observed that 4K-UHD video consumes over 30% more energy, and up to 6 times more data than full HD (1080p), for a marginal increase in objective quality. We also discussed our energy metric for video. We showed how this metric can be combined with existing QoE metrics for video to produce a combined score – the EnVI – which considers both the QoE and the energy usage of video. This could be employed by a variety of stakeholders for energy-aware usage and adaptation of video streams.

For future work, we intend to use extend our work into an open-source tool, and to use that tool to perform more analysis on the energy profiles of various video configurations, as well as for mobile devices. We will also investigate how our methodology can enable energy-aware adaptation of video streaming. Additionally, we hope to deploy the tool publicly and solicit the help of the community towards building a comprehensive dataset of energy usage information for various video titles, codecs, operating systems and hardware, towards achieving green video usage.

Acknowledgements

This work was partly supported by the IU-ATC project, funded by grant EP/J016756/1 from the Engineering and Physical Sciences Research Council (EPSRC). Oche Ejembi was funded by the Scottish Informatics and Computer Science Alliance (SICSA).

References

[1] W. V. Heddeghem, S. Lambert, B. Lannoo, D. Colle, M. Pickavet, and P. Demeester, "Trends in worldwide ICT electricity consumption from 2007 to 2012," *Computer Communications*, 2014.

[2] Cisco Systems, "Cisco Visual Networking Index 2014," Jun 2014. [Online]. Available: http://goo.gl/3gZy0z

[3] A. Balachandran, V. Sekar, A. Akella, S. Seshan, I. Stoica, and H. Zhang, "A Quest for an Internet Video Quality-of-Experience Metric," in *Proc. 11th ACM Workshop on Hot Topics in Networks*. ACM, 2012.

[4] "ITU-T Recommendation P.910: Subjective video quality assessment methods for multimedia applications," International Telecommunication Union, Tech. Rep., April 2008.

[5] A. B. Bener, M. Morisio, A. Miranskyy, and S. A. Kocak, "Green it and green software," *Software, IEEE*, October 2014.

[6] B. Penzenstadler, A. Raturi, D. Richardson, and B. Tomlinson, "Safety, Security, Now Sustainability: The Nonfunctional Requirement for the 21st Century," *Software, IEEE*, May 2014.

[7] M. Hazas, "Society pushes to go faster, but data binges carry environmental costs," *The Conversation*, January 2015.

[8] F. Fernandes, X. Ducloux, Z. Ma, E. Faramarzi, P. Gendron, and J. Wen, "The Green Metadata Standard for Energy-Efficient Video Consumption," *MultiMedia, IEEE*, Jan 2015.

[9] O. Ejembi and S. N. Bhatti, "Help save the planet: Please do adjust your picture," in *MM2014 - 22nd ACM Intl. Conf. on Multimedia*, Nov 2014.

[10] P. Georgopoulos, Y. Elkhatib, M. Broadbent, M. Mu, and N. Race, "Towards Network-wide QoE Fairness Using Openflow-assisted Adaptive Video Streaming," in *Proc. 2013 ACM SIGCOMM Workshop on Future Human-centric Multimedia Networking*, Aug 2013.

[11] X. Che, B. Ip, and L. Lin, "A Survey of Current YouTube Video Characteristics," *MultiMedia, IEEE*, vol. 22, no. 2, pp. 56–63, Apr 2015.

[12] O. Ejembi and S. N. Bhatti, "Client-side Energy Costs of Video Streaming," in *GreenCom 2015 - 11th IEEE Intl. Conf. on Green Computing and Communications*, Dec 2015.