Adaptive Emergency Scenery Video Communications using HEVC for Responsive Decision Support in Disaster Incidents

Z. Antoniou[1], *Member, IEEE*, A. S. Panayides[2], *Member, IEEE*, M. S. Pattichis[3], *Senior Member, IEEE,* S. Stavrou[4], E. Kyriacou[5], *Senior Member, IEEE,* A. Spanias[6], *Fellow, IEEE*, A. G. Constantinides[2], *Fellow, IEEE,* and C. S. Pattichis[1], *Senior Member, IEEE*

***Abstract*— This study proposes a unifying framework for m-Health video communication systems that provides for the joint optimization of video quality, bitrate demands, and encoding time. The framework is video modality and infrastructure independent and facilitates adaptation to the best available encoding mode that satisfies underlying technology and application imposed constraints. The scalability of the proposed algorithm is demonstrated using different HEVC encoding configurations and realistic modelling of 802.11x wireless infrastructure for emergency scenery and response videos. Extensive experimentation shows that a jointly optimal solution in the encoding time, bitrate, and video quality space is feasible.**

I. Introduction

Challenges associated with real-time wireless video communications for emergency response systems stem from satisfying often opposing constraints imposed by enabling technologies [1]. Primary technologies that are key to the success of such systems are video encoding, wireless transmission, and video quality assessment (VQA). Hence, the development of an efficient mobile-health (m-Health) video communication system for emergency telemedicine and response in disaster incidents, needs to address implications associated with parameterizing each component so that a jointly optimum solution is reached that both satisfies the application-specific and technology-imposed requirements.

The unparalleled growth of m-Health medical video communication systems over the past decade is driven by the vast array of applications that can provide significant improvements in patients' quality of care and hence life once deployed in standard clinical practice [1]-[3]. Such applications range from in-ambulance emergency video (trauma, ultrasound, etc.) communication for remote diagnosis and care [4]-[5], to mass population screening in developing countries of general cardiovascular diseases [6], [7] (performed by on-site ultrasound technicians and verified by remote medical experts), and robotic-assisted examination and surgery guided by ultrasound and ambient video communication. At the same time, real-time emergency scenery video in disaster incidents can significantly aid in assessing the current situation and facilitate responsive decision making and efficient triage support [2], [8].

Motivated by the aforementioned, m-Health video communication systems evolved to diagnostically driven systems, exploiting the common goal of maximizing the communicated video's diagnostic quality [1]. The approach is to adapt both the encoding and transmission process in a context-aware fashion, exploiting individual video properties toward enhancing compression efficiency (and hence quality) and reliability of (error-prone) wireless communication [4]-[9]. The limitation of these approaches lies in the observation that they are often video modality specific and/ or optimize a single objective (e.g. compression technology) while considering remaining objectives as black boxes (e.g. wireless infrastructure and VQA). Cross-layer methods address this issue by considering parameters from different layers (e.g. application and physical layers) but fail to accommodate scalability (in terms of video modality and enabling technologies) as they typically involve application-specific and technology-bounded parameters [10]-[12].

In this paper we propose a top-down, video modality and underlying technology independent, scalable approach. The approach is to seek a jointly optimum solution in the three-dimensional space dictated by the most prominent attributes that characterize a video's state and performance. Namely, the objective is to maximize video quality while minimizing bandwidth and encoding time. In vector form, this multi-objective optimization framework can be expressed as:

$$min\left(-Q(E_c), B_r(E_c), T(E_c)\right) \quad (1)$$

where Q stands for video quality (we use the negative sign on Q as our objective is to maximize video quality), B_r denotes the resulting video bitrate, and T denotes the total encoding time. The E_c argument corresponds to the employed encoding configurations resulting in Q, B_r, and T values.

The solution of equation (1) generates a Pareto front. The Pareto front is a set of encoding configurations so that no other configuration exists that simultaneously improves Q, B_r, and T values. In other words, an encoding configuration $E_{p,opt}$ is optimal if no other configuration $E_{p,other}$ exists that provides finer video quality, requires lower bandwidth demands, and is encoded in less time.

*Research supported by grant NSF AWD CNS-1422031.

[1]Z. Antoniou and C.S. Pattichis are with the Department of Computer Science, University of Cyprus, Nicosia, Cyprus (tel:+357-22892697, e-mail: {antoniou.zinonas, pattichi}@ucy.ac.cy).

[2]A. S. Panayides and A. G. Constantinides are with the Department of Electrical and Electronic Engineering, Imperial College, London, UK (e-mail: {a.panagidis, a.constantinides}@imperial.ac.uk).

[3]M.S. Pattichis is with the Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, USA (e-mail: pattichis@ece.unm.edu).

[4]S. Stavrou is with the Faculty of Pure and Applied Sciences, Open University of Cyprus, Nicosia, Cyprus (e-mail: stavros.stavrou@ouc.ac.cy).

[5]E. Kyriacou is with the Department of Computer Science and Engineering, Frederick University, Cyprus (e-mail: e.kyriacou@frederick.ac.cy).

[6]A. Spanias is with the School of Electrical, Computer, and Energy Engineering at Arizona State University e, USA (e-mail: spanias@asu.edu).

(a) Emergency Scenery (b) Emergency Response

Figure 1. (a) First responders (fire and rescue team) evacuating the injured after an earthquake (720x576, 25fps). (b) First responders (paramedics) provide on-site first aid to victims (1440x1080, 25 fps).

TABLE I. NEW HEVC CONFIGURATIONS FOR ULTRAFAST PRESET

Parameter	***Value***	***Parameter***	***Value***
Preset	Ultrafast	Frame Threads	1/2/3/4
GOP	B6,B4, ZL	SAO	On/Off
GOP structure	Open/Close	DBF	On/Off
QP	20,22,24,26,28,30, 32,34,36,38,40,42	Tuning	psnr, fastdecode, zerolatency
Total configurations per video		***1776***	

TABLE II. 802.11 UDP THROUGHPUT - MANHATTAN ENVIRONMENT

802.11	***Band. (GHz)***	***Chan. (MHz)***	***UDP Throughout***
a	5	20	5Mbps - 26.3Mbps
b	2	20	0.91Mbps – 6.23Mbps
g	2	20	5Mbps - 26.3Mbps
n	2	20	7 Mbps - 67.2Mbps
n	5	20	7 Mbps - 67.2Mbps
ac	5	20-160	7Mbps - 80Mbps (20MHz)

Guard Interval: 400, Aggregation Mode: A-MPDU. K-frames: 32 (for 802.11n and 802.11ac)

MAC: CSMA/CA, Appl. Bytes: 1470, Contention Window: 32, Single Input Single Output (SISO) Antennas (for all 802.11x protocols)

801.11ac: 14.5-172 Mbps (40MHz), 30.1-327 Mbps (80MHz), 60.5-32 Mbps (160MHz)

To demonstrate the efficiency and scalability of the proposed framework, we seek optimal solutions subject to realistic constraints imposed on video quality $Q \geq Q_{min}$, video bitrate $B_r \leq B_{r,max}$, and encoding time $T \leq T_{max}$. Naturally, video quality is application-specific (different requirements for emergency scenery, trauma, ultrasound videos, etc.), maximum bandwidth is imposed by the wireless network (different data rates in 3G, 4G (and beyond), 802.11x, 802.16x, etc.), and encoding time is purpose and device oriented (real-time applications, device capabilities).

In this paper we employ the emerging High Efficiency Video Coding (HEVC) standard. We use the open-source x265 implementation that provides the fastest implementation of HEVC to date and investigate different encodings that aim to generate a denser Pareto front. The latter allows wider adaptation options toward satisfying the constraints that are likely to change during a real-time streaming session, especially during in-ambulance treatment involving a moving vehicle, or in disaster incidents where communication infrastructure is pushed to the limits. In terms of wireless infrastructure, realistic IEEE 802.11x simulation is performed documenting data rates variations due to signal strength attenuation. The investigated application scenario is disaster incident response involving emergency scenery video and on-site and in-ambulance patient treatment by paramedical staff.

II. EXPERIMENTAL SETUP

A. Material

For testing the proposed framework, two videos were used, acquired during real training events of the Cyprus Fire and Rescue Team and the Cyprus Association of Emergency and Prehospital Care in 2014. The first, high-resolution video 720x576@25fps, documents first-responders that evacuate the injured following an earthquake and fire incident. The 2nd, high-definition (HD) video 1440x1080@25fps, involves on-site prehospital care of injured people, before dispatching with an ambulance to the nearest hospital.

B. Video Compression

The open-source x265 implementation of HEVC software details ten encoding presets that involve different encoding parameters aiming at different encoding speeds [13]. In this paper, we adopt the ultrafast preset based on previous experimentation [8]. We then vary encoding parameters to construct new ultrafast preset instances that will provide a dense Pareto front allowing efficient adaptation. It is important to note here that the HEVC standard, in contrast to previous standards, defines a single main profile that integrates all encoding options. Hence, the resulting bitstream is compatible with all devices that integrate an HEVC decoder. The new encoding configuration instances are depicted in Table I, and include different number of B frames used at a Group of Pictures (GOP) level, open and closed GOP structures, while both sample adaptive offset (SAO) and deblocking filter (DBF) options are switched on and off. Faster encoding is pursued via increasing the number of parallel threads. Other tuning parameters include PSNR, fast decode, and zero latency, while quantization parameter (QP) is varied between 20 and 42. A detailed insight on the performance and complexity of the afore-mentioned parameters, some new in HEVC standard, is given in [14].

C. Wireless Infrastructure: 802.11 Simulation Environment

Realistic modelling of emergency telemedicine scenarios where an ambulance traverses via a typical city environment to the nearest hospital premises requires precise modelling of the available upload data rates. Toward this direction, a typical Manhattan environment is considered in this series of experiments, where buildings are arranged in blocks. The simulation area is approximately 0.5km x 0.5km with each building block having dimensions of 40m x 40m and a height of 15m (5 story building). A 16m dual carriage way and typical pedestrian pavement is assumed between building blocks, while both building blocks and streets are assumed to be constructed by bricks and concrete, respectively [17].

In an emergency incident, either 802.11 wireless broadband services are present or deployed on demand in a microcell or a macrocell setup. To investigate the 802.11 a/b/g/n/ac system performance at 2.4 GHz and 5.2 GHz frequency bands, an electromagnetic ray tracing simulator [20] was employed (TruNET) and an 802.11 UDP throughout performance module was used to characterize the system's throughput in the area of interest. The simulator is capable of modeling the 3D radio coverage and 802.11 throughput performance of any outdoor or indoor environment. It calculates the contributions from reflected, transmitted, or diffracted contributions from buildings, building edges, and rooftops, and utilizes the estimated radio maps and the input parameters of Table II to compute the UDP throughput. In this particular scenario, the Base Station is set at 6m above the buildings and transmits 1W of power. The vehicular user devices have the same specifications.

```
function findParetoFront(Rsl, PF3D)
  Input: Rsl as csv files that contain: PSNR, FPS, QP and BPS
  Output: PF3D that contains 3D Pareto Front points
        Find Pareto Front optimal points
        for eachCfg in Rsl:
            find 3D Pareto Optimal points that satisfy maximum
            PSNR and FPS and minimize BPS as PF3D.
        end for
end function
```

(a) Pareto front computation (offline)

```
function findParetoFrontConstr(Rsl, OM, OMVAL, PF3DC)
  Input: Rsl as csv files that contain: PSNR, FPS, QP and BPS.
  OM as Optimal Mode. Can take values (maxFPS, minBPS, maxVQT)
  OMVAL as Optimal Mode's constraints in terms of FPS, BPS, PSNR
  Output: PF3D of 3D Pareto Front points that satisfy input constraints

    Find Pareto Front optimal points that meet the defined constraints
    for eachOM in OM:
        for eachOMVAL in OMVAL:
            for eachCfg in Rsl:
      find 3D Pareto Optimal points that satisfy maximum PSNR and
      FPS and minimum BPS to meet eachOMVAL, as PF3DC
            end for
        end for
    end for
end function
```

(b) Pareto front points that satisfy optimization constraints

```
function VideoEncoding (V, Prm, Rsl)
  Input: video V, parameters in Prm
  Output: Rsl as csv files that contain: PSNR, FPS, QP and BPS
    Apply all possible configurations on the Ultrafast preset
    for eachPrm in Prm:
  Compress V using eachPrm and compute (PSNR, FPS, BPS) as Rsl
    end for
end function
```

(c) Video encoding using all available configurations

Figure 2. Pseudo code of Pareto front computation for all HEVC configurations and optimal points selection that satisfy constraints.

III. METHODOLOGY

The algorithm for the computation of the Pareto front based on the extended HEVC encoding instances is depicted in Fig. 2. Resulting values from all new configurations described in Table I are traversed in order to construct the 3D Pareto front. This is an offline process that is used to train the proposed framework so as to select the optimal encoding modes based on the given constraints. As discussed earlier, the new parameterized ultrafast configurations, are mapped to the three key objectives that characterize the video's state, namely video quality, bandwidth demands, and encoding time. These objectives are further used to define the three operation modes of the proposed framework [15]-[16]:

- **Maximum video quality mode**
  $$max\ Q\ subject\ to\ \left(B_r \leq B_{r,max}\right)\ and\ \left(T \leq T_{max}\right) \quad (2)$$
- **Minimum bitrate demands mode**
  $$min\ B_r\ subject\ to\ \left(Q \geq Q_{min}\right)\ and\ \left(T \leq T_{max}\right) \quad (3)$$
- **Minimum encoding time (or maximum FPS) mode**
  $$min\ T\ subject\ to\ \left(Q \geq Q_{min}\right)\ and\ \left(B_r \leq B_{r,max}\right) \quad (4)$$

or, $max\ FPS\ subject\ to\ \left(Q \geq Q_{min}\right)\ and\ \left(B_r \leq B_{r,max}\right)$(5)

Then, using the set of points in the Pareto front and the selected mode of operation, the algorithm described in Fig. 2 is used to select the optimal configuration that satisfies the application and technology imposed constraints. This adaptation process can be performed at any given GOP level. In other words, real-time adaptation at a GOP level may be triggered following a change in the defined constraints (i.e., a change in the available upload data rates or a sudden bitrate increase due to video content change in the video sequence).

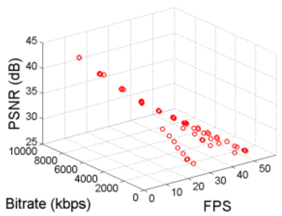
IV. RESULTS

For validating the proposed framework the two videos depicted in Fig. 1 were used. All encoding configurations run on a GNU/Linux 64-bit platform with 8GB RAM using QEMU Virtual CPU version (cpu64-rhel6) with 6 cores (6 threads) running at 2.26 GHz.

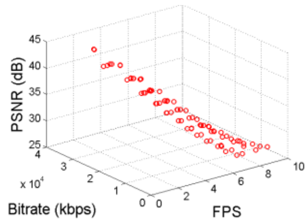
Resulting Pareto fronts for both videos are depicted in Fig. 3. An important observation is that for the HD video (see Fig. 3(b)), real-time encoding is not possible (all encoding rates are less than 25 fps which is the video's frame rate). As a result, in case HD video resolution is desired, equipment with more advanced capabilities should be employed. However, this is not the case for the high-resolution video, which is well accommodated by the employed platform.

Optimization for all modes of operation discussed in equations (2), (3), and (5) were investigated. These optimization examples are presented in Table III. Constraints for each optimization scenario have been generated using realistic constraints as discussed next. In the first example, the objective is to maximize the communicated video's quality, subject to real-time constraints (which translates to an encoding FPS (frames per second) rate higher than 25 fps) and maximum bitrate that does not exceed the minimum supported bitrate of 802.11b broadband service which is estimated at 0.91 Mbps as depicted in Table II (guaranteeing in this way unobstructed video transmission). The achieved optimization results show that all constraints have been met for this example, as the resulting configuration has an encoding rate ≥ 25 fps while it does not exceed the maximum allowed bitrate threshold. The highest possible video quality measured in Peak-Signal-to-Noise-Ratio (PSNR) is 33.78 dB. In case more stringent requirements are imposed, for example 3G upload data rates which extend up to 384 kbps [7], the maximum video quality is limited to 31.34 dB. The latter demonstrates the scalability of the proposed framework, abstracting infrastructure limitations to input constraints.

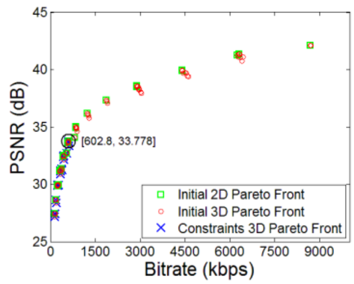
In the 2nd optimization mode, the objective is to minimize bandwidth demands subject to real-time encoding constraint (≥25 fps) and minimum video quality ≥ 32 dB. The lowest possible bitrate that satisfies both constraints is 440 kbps. Resulting PSNR value is 32.43 dB while encoding rate is well above the 25 fps threshold at 30.69 fps. In the last optimization mode the objective is to maximize the encoded rate measured in fps (minimize encoding speed) while conforming to 0.91 Mbps and 34 dB in terms of maximum bandwidth demands and minimum video quality, respectively. Both constraints are satisfied, as bitrate demands are less than the available data rate while PSNR value extends up to 34.61 dB. Despite both constraints being met however, the maximum achievable encoding rate does not qualify for real-time encoding at 23.46 fps. It is important to note here that in the typical optimization case where both constraints are met, at least one of the constraints is met marginally, with values close to the imposed thresholds.



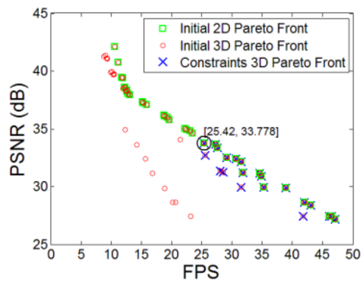
(a) Pareto front of emergency scenery video of Fig. 1(a).



(b) Pareto front of emergency scenery video of Fig. 1(b).



(c) 3D → 2D Pareto front of PSNR vs Bitrate



(d) 3D → 2D Pareto front of PSNR vs FPS

Figure 3. (a)-(b) 3D Pareto front of emergency scenery and first-responders on-site care videos of Figs. 1(a) and (b), respectively. (c) 3D → 2D Pareto front of PSNR vs Bitrate depicting optimal points that satisfy the 1st example optimization mode of Table III. (d) 3D → 2D Pareto front of PSNR vs FPS depicting optimal points that satisfy the 1st example optimization mode of Table III.

TABLE III. MODE OPTIMIZATION EXAMPLES FOR MAX VIDEO QUALITY, MIN BITRATE, AND MAX FPS (MIN ENCODING TIME) MODES

Mode	*QP*	*GOP*	*Frame Threads*	*FPS*	*Bitrate (kbps)*	*PSNR (dB)*
Max. Video Quality	32	B4	4	25.42	602.8	**33.78**
	constraints			≥25	≤910	
	36	ZL	1	28.04	382.6	**31.34**
	constraints			≥25	≤384	
Min. Bitrate	34	B4	4	30.69	**440.1**	32.43
	constraints			≥25		≥32
Max. FPS	30	B4	3	**23.46**	889.9	34.61
	constraints				≤910	≥34

V. CONCLUDING REMARKS

A unifying framework that provides for the joint optimization of video quality, bitrate, and encoding time has been proposed. Experimentation involved the emerging HEVC standard and realistic modelling of 802.11x wireless service using emergency scenery videos acquired by live training events. On-going work involves investigating additional HEVC encodings toward generating a denser Pareto front and different wireless topologies and wireless infrastructure for emergency telemedicine. The latter will allow demonstrating on-the-fly adaptation of the encoding parameters to varying network parameters. Moreover, experimenting with an already constructed dataset of different emergency incident response videos will allow the fine-tuning of the employed encoding parameters to cater for different video content and motions before pilot live deployments.

REFERENCES

[1] A. S. Panayides, M. S. Pattichis, and C. S. Pattichis, "Mobile-Health Systems Use Diagnostically Driven Medical Video Technologies [Life Sciences]," *IEEE Signal Processing Magazine,* vol.30, no.6, pp.163-172, Nov. 2013.

[2] E. Kyriacou, M.S. Pattichis, C.S. Pattichis, A. Panayides, and A. Pitsillides, "m-Health e-Emergency Systems: Current Status and Future Directions," *IEEE Antennas and Propagation Magazine*, vol. 49, no.1, pp. 216-231, Feb. 2007.

[3] A. S. Panayides, M.S. Pattichis, A.G. Constantinides, C.S. Pattichis, "M-health medical video communication systems: An overview of design approaches and recent advances," *in Proc. of IEEE EMBC'13*, Osaka, Japan, 3-7 July 2013, pp.7253-7256.

[4] A. Panayides, Z. Antoniou, Y. Mylonas, M. S. Pattichis, A. Pitsillides, and C. S. Pattichis, "High-Resolution, Low-delay, and Error-resilient Medical Ultrasound Video Communication Using H.264/AVC Over Mobile WiMAX Networks," *IEEE Journal of Biomedical and Health Informatics*, vol.17, no.3, pp.619-628, May 2013.

[5] A. Panayides, I. Eleftheriou, and M. Pantziaris, "Open-Source Telemedicine Platform for Wireless Medical Video Communication," *International Journal of Telemedicine and Applications*, vol. 2013.

[6] A. Panayides , Z. Antoniou , V. Barberis , M. Pattichis , C. Pattichis and E. Kyriacou "Abdominal aortic aneurysm medical video transmission", *Proc. IEEE-EMBS Int. Conf. Biomed. Health Inf.*, Hong Kong, China, Jan. 2-7, 2012, pp.679 -682.

[7] A. Panayides, M.S. Pattichis, C.S. Pattichis, C.P. Loizou, M. Pantziaris, and A. Pitsillides, "Atherosclerotic Plaque Ultrasound Video Encoding, Wireless Transmission, and Quality Assessment Using H.264", *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 3, pp.387-397, May 2011.

[8] A. S. Panayides, A. G. Constantinides, M. S. Pattichis, E. Kyriacou, C. S. Pattichis, "Adaptive real-time HEVC encoding of emergency scenery video," in Proc of MobiHealth'14, Greece, Nov. 2014, pp.217.

[9] E. Cavero, A. Alesanco, and J. Garcia, "Enhanced protocol for real time transmission of echocardiograms over wireless channels," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 11, pp.3212-3220, Nov. 2012.

[10] A. Alinejad, N. Philip, R. Istepanian, "Cross Layer Ultrasound Video Streaming over Mobile WiMAX and HSUPA Networks," *IEEE Trans. Inf. Technol. Biomed.*, vol. 16, no. 1, pp. 31-39, Jan 2012.

[11] R.S.H. Istepanian, N. Y. Philip, and M.G. Martini, "Medical QoS Provision based on Reinforcement Learning in Ultrasound Streaming over 3.5G Wireless Systems," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 4, May, 2009, pp. 566-574.

[12] M. G. Martini, R. S. H. Istepanian, M. Mazzotti, and N. Philip, "Robust multi-layer control for enhanced wireless tele-medical video streaming," *IEEE Trans. Mobile Comput.*, vol. 9(1),pp.5-16, Jan 2010.

[13] MulticoreWare Inc, "X265 software documentation," https://x265.readthedocs.org/en/default/, 2015.

[14] F. Bossen, B. Bross, K. Suhring, and D. Flynn, "HEVC Complexity and Implementation Analysis," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22(12), pp. 1685–1696, Dec 2012.

[15] Y. Jiang and M. Pattichis, "Dynamically reconfigurable dct architectures based on bitrate, power, and image quality considerations," *in Proc of ICIP'12*, Sept 2012, pp. 2465–2468.

[16] Y. Jiang and M. S. Pattichis, "A dynamically reconfigurable architecture system for time-varying image constraints (drastic) for motion jpeg," *Journal of Real-Time Image Processing*, pp.1–17, 2014.

[17] S. Stavrou and S.|R. Saunders, 'Review of constitutive parameters of building materials', IEE Antennas and Propagation, 2003. (ICAP 2003). Twelfth International Conference on Antennas and Propagation, Volume 1, 31 March-3 April 2003, pp.211 – 215.

[18] M. Yang, S. Stavrou, A.K., Brown, 'Hybrid Ray-tracing model for radio wave propagation through periodic building structures', IET Proceedings on Microwaves, Antennas and Propagation, 2011 February, pp.340-348.