

Simple, Effective Rate Control for Video Distribution in Heterogeneous Intelligent Transportation System Networks

Ekasit Vorakitolan[†], Joseph P. Havlicek[†], Ronald D. Barnes[†], and Alan R. Stevenson[‡]

[†]*School of Electrical and Computer Engineering, University of Oklahoma, Norman, OK, USA*

[‡]*Technology Services Division, Oklahoma Department of Transportation, Oklahoma City, OK, USA*

Abstract—We introduce a simple video rate control strategy for reducing or eliminating decoder errors due to network errors and dropped packets on low bandwidth links in heterogeneous intelligent transportation system video networks. The proposed method is based on adaptively dropping frames prior to H.264 encoding. Performance of this strategy is assessed by evaluating three video quality metrics (PSNR, SSIM, and VQM) on actual traffic surveillance video that is transmitted across links with simulated network errors. We show that high spatial quality can be maintained at the expense of temporal resolution by reducing the effective frame rate as low as five frames per second, which is still sufficient for traffic network monitoring. Our solution is driven by several specialized constraints. Hence, while it is highly effective for intelligent transportation systems, this approach would not be suitable for more general video entertainment applications.

Keywords—video rate control, video quality assessment, intelligent transportation systems

I. INTRODUCTION

We propose a specialized rate control strategy to improve video quality in an application specific heterogeneous distribution network where some links are characterized by low bandwidth. Our solution is simple, works well, and should be applicable to other types of networks with similar constraints. We are mainly interested in the Oklahoma statewide Intelligent Transportation System (ITS) network [1], [2], which is a wide-area distributed, fault tolerant network that extends across the entire state connecting transportation related instrumentation and communication devices to ITS control consoles situated at over 100 stakeholder sites. The devices include microwave speed sensors, portable and highway overhead dynamic message signs (DMS), and roadway weather information systems (RWIS), all of which have relatively low bandwidth requirements, and hundreds of video surveillance cameras. The stakeholders include state and municipal transportation engineers and planners, emergency managers, Oklahoma Department of Transportation (ODOT) headquarters and field offices, dispatchers and 911 operations centers, and police, fire, EMS, and military agencies. These stakeholders use the video signals and other sensor information for situational awareness, to monitor the statewide surface transportation network, and to plan,

execute, and coordinate multi-agency incident responses. Ensuring high quality video distribution is absolutely essential to the overall effectiveness of the statewide intelligent transportation system.

The network backbone is Gigabit Ethernet (GigE). However, because of practical cost constraints and the inherently wide-area nature of the system, many stakeholder consoles do not have a physical connection to the backbone but are connected via lower bandwidth alternative technologies such as point-to-point microwave links, 900 MHz radio, CDMA modems, T-1 lines (1.544 Mbps), thin Ethernet (10 Mbps), or fast Ethernet (100 Mbps). During periods of heavy network traffic, these consoles can experience substantial video quality loss due to routers and switches dropping packets to enforce packet limits and avoid buffer overruns. Examples of such degraded video frames are shown in Fig. 4(a) and (b). For comparison, a temporally close frame that is nearly error-free is shown in Fig. 4(c). This presents a significant problem because heavy network traffic often occurs during serious incidents when delivery of high quality video is most critical from a public safety standpoint. In this paper, we propose a simple yet effective adaptive rate control strategy based on frame dropping and replication that can reduce the required video bitrate and avoid errors of the type shown in Fig. 4 while maintaining the aspects of quality that are most important for this application. More details about the problem and the associated constraints are given in Section II. The proposed solution is described in Section III followed by a preliminary performance evaluation with respect to three well known video quality metrics in Section IV.

II. BACKGROUND AND CONSTRAINTS

In block based video compression standards such as MPEG-4 and H.264, both of which are used in the Oklahoma ITS network, the encoder video quality control module sets the quantization parameter (QP) to balance quality in the spatial and temporal dimensions against compression ratio and video bitrate. To maintain video quality, QP should be adjusted based on the dynamic network conditions and video complexity. Standards like H.264 provide multiple modes for interframe prediction (temporal) and intraframe prediction (spatial). There is a circular relationship where

This work was supported in part by grants from the Oklahoma Department of Transportation and the Oklahoma Highway Safety Office.

QP is defined by the mean absolute residual (MAD) between the prediction macroblock (MB) and the current MB and the rate distortion optimization (RDO) algorithm selects the prediction mode based on QP, but the RDO result is required in order to calculate the MAD, resulting in a “chicken and egg” dilemma [3], [4]. Novel algorithms such as scalable rate control, MAD prediction, and SSIM-based frame skipping have been proposed to circumvent this dilemma [4]–[9].

However, the majority of existing research on video rate control algorithms has been focused on constraints that are driven primarily by the video entertainment industry where high quality in both the spatial and temporal dimensions is required. This is in contrast to ITS applications where spatial fidelity and resolution are critical but high temporal resolution is usually of less concern. In these applications, frame rates as low as 15 or even five frames per second (fps) are often adequate provided that good spatial resolution can be maintained.

What is more important in ITS applications is that, in addition to maintaining spatial quality, the implemented rate control strategy must be compatible with existing CODECS and display devices, both of which generally represent substantial sunk costs. The CODECS deployed for ITS applications must operate under harsh environmental conditions including extreme temperatures, moisture, and vibration. They are typically expensive and often implement only a subset of the full video coding standard functionality. In view of these constraints, we seek a rate control algorithm that avoids the need for complex, non-standard calculations and maintains an effective frame rate of 30 fps for compatibility with existing decoder hardware and display systems.

III. PROPOSED RATE CONTROL STRATEGY

An important observation about the video produced by ITS cameras is that it usually consists of moving vehicles against a background that is essentially stationary (see Fig. 4). A typical camera pole in the Oklahoma ITS has up to four fixed cameras that are rigidly mounted to provide stationary views of nearby roadways and a single pan-tilt-zoom camera that will be steered occasionally (e.g., for incident surveillance) but remains stationary most of the time. Thus, the video backgrounds are essentially constant up to wind shake. Therefore, a frame dropping scheme can dramatically reduce the number of positions that need to be encoded for each moving vehicle without incurring any appreciable loss in the substantial coding gains achieved on the background by MPEG-4 or H.264 interframe prediction.

Compatibility with equipment that is hardwired for 30 fps can then be maintained with minimal overhead by repeating each retained frame, since every macroblock in a repeated frame has a perfect match with a motion vector that is zero. This idea is illustrated in Fig. 1, where two out of every three frames are dropped and each retained frame is repeated three times. This scheme is straightforward to implement

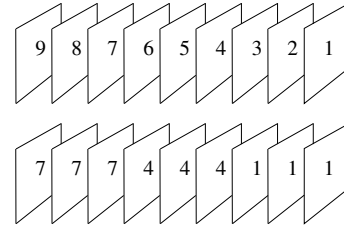


Figure 1. Rate control scheme for dropping and repeating frames. Top: original video stream at 30 fps. Bottom: the rate controller drops two out of every three frames. Each retained frame is repeated three times.

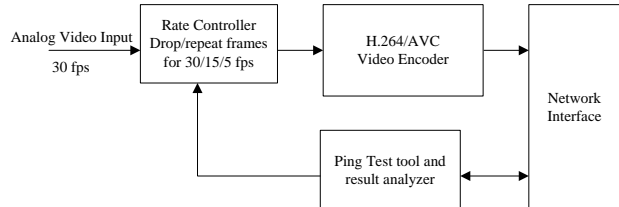


Figure 2. Block diagram of proposed video rate control strategy. The ping test module adjusts the rate controller. Dropped frames are replaced by copies of the retained frames prior to H.264 encoding.

in an inexpensive custom circuit module, does not require changing or modifying any video compression standard, and is interoperable with all currently deployed ITS equipment.

A block diagram of the proposed rate control strategy is shown in Fig. 2, where the scheme of Fig. 1 is implemented in the *Rate Controller* block. In practice, we have implemented three rates: no frame dropping (30 fps), drop one of every two frames (15 fps), and drop five of every six frames (5 fps). Closed loop control is implemented by incorporating a simple ping test module that sends ICMP packets to continuously monitor network traffic on the route to the client console and adjusts the number of frames dropped by the rate controller based on the results.

IV. EXPERIMENTS AND DISCUSSION

Performance of the proposed rate control strategy was evaluated using the experimental setup shown in Fig. 3. The open source WANem network simulator [10] was used to simulate network errors including random bit errors at 5%, 15%, 30%, and 50%, dropped packets at 5%-50%, duplicated packets at 5%-50%, randomly reordered packets at 5%-50%, random packet delays of 500 ms, random packet delays with jitter up to 300 ms, and random network disconnects on T-1, thin Ethernet (10 Mbps), and fast Ethernet (100 Mbps) links. The maximum router/switch buffer size (packet limit) was set to 10,000 bytes for T-1 links, 10,240 bytes for thin Ethernet links, and 1,000 bytes for fast Ethernet links.

The video server sourced real traffic surveillance video acquired from the Oklahoma ITS using the proposed rate control strategy as depicted in Fig. 2 at frame rates of 30, 15, and 5 fps for all three link types. In all cases, the

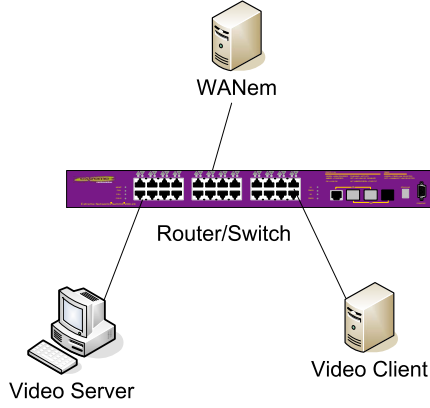


Figure 3. Experimental setup for video quality assessment.

dropped frames were replaced with copies of the retained frames as shown in Fig. 1. The SmartCapture USB H.264 hardware video encoder was used [11] with a resolution of 640×480 pixels, frame rate of 30 fps, video bitrate of 1.01 Mbps, and I-frame interval of 30 frames. Nine minutes of video was transmitted for each test case. The received video streams were captured at the client console and analyzed with the MSU video quality measurement tool (VQMT) version 3.0 [12].

Average values computed for the PSNR (dB), SSIM, and VQM [12], [13] video quality metrics are reported in Table I, where the same video sequences were input to the rate controller for each test. Better video quality is indicated by higher PSNR and SSIM values and by lower VQM values. Under heavy loading and network errors, it may be seen in almost every case that the proposed rate control strategy delivers the best performance at the lowest rate of 5 fps as predicted. Three typical decoded frames for the case of dropped packets over a thin Ethernet link are shown in Fig. 4. The frame in Fig. 4(a) was encoded without rate control and exhibits the type of video degradation that we have observed in the actual ITS network. The improvement in Fig. 4(b) was obtained by dropping one of every two frames at the rate controller and, as can be seen in Fig. 4(c), virtually all decoder errors are eliminated by dropping five of every six frames.

The video rate control strategy proposed in this paper is simple and effective for reducing the video bitrate in applications like intelligent transportation system video distribution networks where spatial quality must be assured but temporal resolution is less critical. Our approach is based on the idea of dropping frames to reduce the information bandwidth but replacing the dropped frames with copies of the retained frames to maintain compatibility with deployed video devices that may be hardwired for operation at 30 fps. This solution is specific to our application and would not be suitable for more general video entertainment applications.



Figure 4. Decoded video frames received through a network with dropped packets. (a) No rate control (30 fps). (b) Proposed rate control method at 15 fps. (c) Proposed method at 5 fps with virtually all decode errors eliminated.

A key assumption important to the success of this approach is that the video signals consist primarily of relatively small objects moving against a stationary background. Similar constraints apply in other applications such as video traffic signal enforcement [14] and weather monitoring [15] where

Table I
VIDEO QUALITY MEASURED BETWEEN ORIGINAL AND RECEIVED VIDEO STREAMS FOR VARIOUS NETWORK ERRORS. PSNR IS GIVEN IN dB.

| | | T-1 (5fps) | T-1 (15fps) | T-1 (30fps) | thinE (5fps) | thinE (15fps) | thinE (30fps) | fastE (5fps) | fastE (15fps) | fastE (30fps) |
|------------|------|----------------|----------------|-------------|----------------|---------------|---------------|----------------|---------------|---------------|
| Random | PSNR | 20.3584 | 19.6885 | 19.3964 | 20.2144 | 19.3156 | 19.2852 | 19.8713 | 18.9832 | 18.3290 |
| TX | SSIM | 0.8474 | 0.8314 | 0.8212 | 0.8383 | 0.8301 | 0.8069 | 0.8322 | 0.8034 | 0.7777 |
| bit errors | VQM | 7.9616 | 8.6240 | 8.5521 | 8.0981 | 8.5534 | 8.9897 | 8.3212 | 9.0984 | 9.6205 |
| Dropped | PSNR | 20.3449 | 19.6873 | 19.5257 | 20.3702 | 19.6044 | 19.6678 | 20.2464 | 19.3571 | 19.1471 |
| Packets | SSIM | 0.8531 | 0.8337 | 0.8173 | 0.8432 | 0.8240 | 0.8173 | 0.8438 | 0.8203 | 0.8064 |
| | VQM | 7.9381 | 8.5682 | 8.6494 | 7.9961 | 8.7009 | 8.5840 | 8.0118 | 8.9021 | 8.9974 |
| Duplicated | PSNR | 20.0585 | 19.9127 | 19.5275 | 20.3817 | 16.9115 | 19.2271 | 20.0224 | 19.4605 | 18.3453 |
| Packets | SSIM | 0.8383 | 0.8399 | 0.8172 | 0.8469 | 0.8207 | 0.8115 | 0.8413 | 0.8235 | 0.7874 |
| | VQM | 8.1903 | 8.4714 | 8.6804 | 7.9784 | 8.7301 | 8.9495 | 8.2056 | 8.6302 | 9.6858 |
| Packets | PSNR | 20.0325 | 19.5379 | 17.2584 | 19.7236 | 19.2488 | 19.1554 | 20.5293 | 19.4198 | 18.8919 |
| out of | SSIM | 0.8422 | 0.8336 | 0.8296 | 0.8258 | 0.8142 | 0.8020 | 0.8549 | 0.8265 | 0.8114 |
| Order | VQM | 8.2040 | 8.6268 | 8.4388 | 8.4492 | 8.8614 | 8.9369 | 7.8635 | 8.7721 | 8.9162 |
| Random | PSNR | 19.7739 | 19.3600 | 19.1580 | 20.0240 | 19.9237 | 19.5490 | 20.5818 | 19.6512 | 19.5735 |
| Packets | SSIM | 0.8305 | 0.8217 | 0.8177 | 0.8451 | 0.8270 | 0.8243 | 0.8533 | 0.8303 | 0.8117 |
| Delayed | VQM | 8.7732 | 8.7813 | 9.0326 | 8.2128 | 8.6951 | 8.4956 | 7.8202 | 8.1486 | 8.4578 |
| Packet | PSNR | 19.6911 | 19.8307 | 19.3965 | 20.7949 | 20.0638 | 19.5568 | 20.7529 | 20.0260 | 19.3317 |
| Delay | SSIM | 0.8355 | 0.8372 | 0.8095 | 0.8663 | 0.8430 | 0.8119 | 0.8608 | 0.8261 | 0.8207 |
| w/ Jitter | VQM | 8.3702 | 8.4594 | 8.7054 | 7.3300 | 8.2303 | 8.7803 | 7.6800 | 8.9461 | 8.1418 |
| Random | PSNR | 20.2168 | 19.5581 | 18.9796 | 20.3854 | 19.8865 | 19.6843 | 20.3734 | 19.4515 | 19.5048 |
| Network | SSIM | 0.8464 | 0.8302 | 0.8044 | 0.8423 | 0.8660 | 0.8682 | 0.8441 | 0.8277 | 0.8160 |
| Disconnect | VQM | 8.2902 | 8.8092 | 8.9844 | 8.1179 | 8.3055 | 9.0128 | 8.0453 | 8.8572 | 8.5957 |

the proposed method may also be successfully applied.

REFERENCES

- [1] B.H. Kilani, E. Vorakitolan, J.P. Havlicek, M.P. Tull, and A.R. Stevenson, "Distributed ITS control and the Oklahoma virtual TMC," in *Proc. IEEE Int'l. Conf. Intel. Transportation Syst.*, St. Louis, MO, Oct. 2009, pp. 785–790.
- [2] B.H. Kilani, P.A. Campbell, J.P. Havlicek, M.P. Tull, A.R. Stevenson, and R.D. Barnes, "Peer-based communication for a distributed, virtual traffic management center," in *Proc. IEEE Int'l. Conf. Intel. Transportation Syst.*, Washington, DC, Oct. 2011, pp. 2068–2073.
- [3] Rate control and H.264. Visited October, 2013. [Online]. Available: http://pixeltools.com/rate_control_paper.html
- [4] Z.G. Li, F. Pan, K.P. Lim, X. Lin, and S. Rahardja, "Adaptive rate control for H.264," in *Proc. IEEE Int'l. Conf. Image Process.*, Singapore, Oct. 2004, pp. 745–748.
- [5] H. Lee, T. Chiang, and Y. Zhang, "Scalable rate control for MPEG-4 video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 10, pp. 878–894, Sep. 2000.
- [6] X. Chen and F. Lu, "A reformative frame layer rate control algorithm for H.264," *IEEE Trans. Consum. Electron.*, vol. 56, pp. 2806 – 2811, Nov. 2010.
- [7] J. Yang, Q. Zhao, and L. Zhang, "The study of frame complexity prediction and rate control in H.264 encoder," in *Proc. IEEE Int'l. Conf. Imag Anal. and Signal Process.*, Taizhou, Apr. 2009, pp. 187–191.
- [8] Z. Cui and X. Zhu, "SSIM-based content adaptive frame skipping for low bit rate H.264 video coding," in *Proc. IEEE Int'l. Conf. Commun. Technol.*, Nanjing, Nov. 2010, pp. 484–487.
- [9] M. Jiang, Y. Xiaoquan, and N. Ling, "Improved frame-layer rate control for H.264 using MAD ratio," in *Proc. IEEE Int'l. Conf. Circuits and Syst.*, Vancouver, Canada, May 2004, pp. 813–816.
- [10] The wide area network emulator. Visited October, 2013. [Online]. Available: <http://wanem.sourceforge.net/>
- [11] SmartCapture PRO: USB based H.264/AAC encoder device. Visited October, 2013. [Online]. Available: <http://www.fastvdo.com/SmartCapture/>
- [12] MSU video quality measurement tool. Visited October, 2013. [Online]. Available: http://compression.ru/video/quality_measure/video_measurement_tool_en.html
- [13] A.B. Watson, J. Hu, J.F. McGowan, III, and J.B. Mulligan, "Design and performance of a digital video quality metric," in *Human Vision and Electronic Imaging IV*, Proc. SPIE, vol. 3644, 1999, pp. 168–174.
- [14] B. Bochner and T. Walden, "Effectiveness of red light cameras," Texas Transportation Institute, Tech. Rep., July 2010.
- [15] H. Song, X. Le, C. Yang-Zhou, and G. Yuan-Yuan, "Weather identifying system based on vehicle video image," in *Proc. IEEE Int'l. Conf. Intel. Control and Automat.*, Taipei, Jun. 2011, pp. 172–175.