RETRANSMISSION STRATEGIES FOR OPTIMIZED JPEG 2000 IMAGE TRANSMISSION IN WIRELESS ENVIRONMENT

Marco Grangetto, Enrico Magli, Gabriella Olmo

CERCOM - Center for Multimedia Radio Communications
Dept. of Electronics, Politecnico di Torino
Corso Duca degli Abruzzi 24
10129 Torino, Italy
marco.grangetto@polito.it
enrico.magli@polito.it
gabriella.olmo@polito.it

ABSTRACT

Transmission of images and video over lossy packet networks is known to pose severe technical challenges; FEC and ARQ are often used to make the transmission more robust towards potential data losses. In this paper we investigate ARQ strategies aimed at optimizing the transmission of progressive codestreams over wireless environments characterized by packet losses. The analysis is specifically tailored to the JPEG 2000 family of standards, and considers Part 1 and 3 for image and video compression, and Part 11 for additional error protection tools. We show that ad-hoc retransmission strategies significantly outperform basic ARQ schemes, providing improved end-to-end image and video quality.

1. INTRODUCTION

JPEG 2000 [1] is the most recent lossy and lossless image compression technology developed by the ISO/IEC SC 29 WG 1. With respect to previous standards, and to other compression schemes, JPEG 2000 provides a number of remarkable features, including improved coding efficiency, fine-grain scalability, support for region of interest coding, random codestream access, and error resilience. Thanks to its advanced features, JPEG 2000 has plenty of possible applications, ranging from digital cameras to medical imaging, remote sensing, compression and storage of compound documents, as well as delivery of multimedia images and video.

Among all possible target applications for JPEG 2000, a key role is played by the wireless imaging applications. Wireless imaging is known to pose severe technical challenges, due to the fact that the compressed data must be transmitted from the source to the destination through a wireless transmission medium, which can potentially cause data losses. Hence, wireless image transmission requires a provision for codestream error robustness, in a such a way that unrecovered bit errors or packet losses do not dramatically reduce the quality of the received data. For these reasons, Part 1, 2, and 3 of the JPEG 2000 standard embody error-resilience tools, which make a compressed codestream relative robust toward errors.

Nevertheless, these tools have not been designed to provide the very high degree of robustness required by some wireless applications. Therefore, when the transmission channel is very harsh, additional error protection must often be provided, e.g. in terms of additional forward error correction (FEC) coding, automatic repeat request (ARQ), or any other tool that can guarantee a given degree of quality even in presence of errors. The wireless imaging problem is insofar important, that the JPEG 2000 committee is developing a new ad-hoc part of the standard, namely Part 11, which addresses the issues relevant to the wireless applications.

In the past, there has been a lot of research work related to error-resilient image transmission over channels prone to bit errors; examples can be found in [3, 4, 5]. However, less work has been done in the scenario of transmission over packet networks potentially subject to packet losses. In [6] FEC is used to generate multiple equi-important descriptions of an image; the decoder exhibits robustness toward packet erasures. However, FEC usually requires a relatively high redundancy, so that, except for multicast applications, ARQ is usually preferred. Typical ARQ schemes are completely unaware of the characteristics of the data that are being transmitted. As a consequence, all packets are treated in the same way. However, since multimedia data are typically highly structured, classical ARQ schemes are often suboptimal, in that they do not exploit this structure. As an example, images are often compressed and stored in a progressive fashion, meaning that the most important part of the data (i.e., those contributing the most to the final quality) are placed at the beginning of the codestream. Intuitively, better performance should be expected if packets containing those data are transmitted more reliably (e.g., by allowing a larger number of retransmission attempts) than the other packets.

This paper deals with the investigation of optimized retransmission strategies for reliable image transmission over lossy packet networks; 3G mobile communication systems and 802.11 wireless LANs are the most widespread (though not the only) examples of possible applications. The analysis is specifically tailored to the JPEG 2000 standard, which supports both image [1] and video [2] compression. We exploit a new tool currently being developed within Part 11 of the standard, i.e. the Error Sensitivity Descriptor (ESD) marker segment [7, 8]. This marker segment allows an encoder to describe the error sensitivity of different parts of the codestream, and to embed this information into the codestream itself. The error sensitivity description can be exploited for several tasks, including: i) optimizing the codestream robustness via unequal error protection; ii) performing intelligent ARQ, by dedicating a higher number of retransmissions to

This work was partially supported by the Italian Ministry of Education and Research under the CERCOM (Center for Multimedia Radio Communications) and PRIMO (Reconfigurable Platforms for Wideband Wireless Communications) grants.

the most important packets; iii) carrying out optimized video streaming, by prefetching the most important packets into the packet scheduler; iv) optimizing the operation of a rate transceiver, which can look up the error sensitivity information to understand which quality would be delivered at a given rate. The main contribution of this paper consists in the performance evaluation of new and existing ARQ schemes for JPEG 2000 in the context of 3G communication systems, possibly exploiting the error sensitivity information, in order to optimize the end-to-end quality of the received images. Although ARQ has already been studied for other encoders and video systems, to the authors' best knowledge an analysis targeted to JPEG 2000 has not been reported yet.

2. NETWORK MODEL AND VIDEO APPLICATION

As stated, we consider image and video communications in 3G mobile cellular systems using JPEG 2000 and Motion JPEG 2000 as target encoders. In particular, the network model is depicted in Fig. 1. The video data are conveyed from a wired network to a base station, which forwards them to the end-user via a wireless link. The base station may contain a proxy server as in [9], which parses the incoming packets in a smart way, so as to optimally handle retransmissions to the end-user. The video data can be generated e.g. by a streaming video server, or by another mobile terminal in case of conversational application. In this paper we consider the downlink between base station and terminal. This corresponds to the whole wireless link in the streaming video case, whereas it only represents the base station to terminal air interface in case of conversational applications; in this latter case, the proxy server should also ask for retransmissions from the video source terminal.

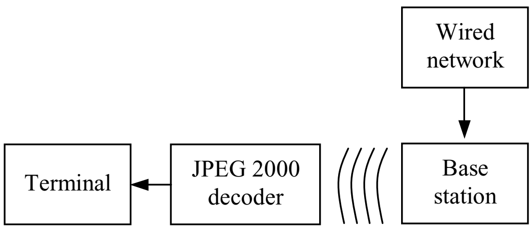

Figure 1: Wireless network model.

We consider intra-frame video compression using [2]. Grayscale images in CIF (Common Intermediate File Format) and QCIF (Quarter CIF) format have been considered (i.e., 352x288 and 176x144 pixels respectively). Channel data rates of 64, 128, 384 and 512 kbit/s have been taken into account; these rates can be easily accommodated by 2.5G/3G mobile communication system, and can provide reasonably good video quality. Packets of 80 bytes each have been considered at the RLC (radio link control) layer, and the RLC acknowledged mode has been used to carry out ARQ at the link layer. This setting is typical of image and video transmission over a UMTS radio link for conversational applications [10]. The video frame-rate has been selected so as to provide a reasonably good quality to the end user (i.e. temporally smooth playback with sufficiently high PSNR) with respect to the available bandwidth and the JPEG 2000 compression capability. This has led to the selection of 10 and 15 frames/s for QCIF pictures, and 5 and 10 frames/s for higher resolution CIF pictures; these corresponds to source bit-rates from 0.2 to 1.5 bit-per-pixel (bpp). The data rate, packet size and frame rate constrain the maximum number N of packet transmission attempts available for each picture (including possible retransmissions).

3. ARQ STRATEGIES

We have considered the following ARQ policy that exploits the error sensitivity information contained in the ESD data structure. We assume that a proxy server at the base station intelligently manages retransmissions towards the end-user terminal. The intelligence lies in the capability of the proxy server to parse a JPEG 2000 or Motion JPEG 2000 codestream and i) identify the picture number and codestream start and end bytes, ii) identify the main and tile-part headers, which are vital for correct decoding, iii) extract the ESD data structure. The ESD can be employed to adapt the retransmission strategy according to the error sensitivity of the different packets, and/or to carry out rate adaptation with exact knowledge of the delivered quality for each possible rate by means e.g. of the absolute sensitivity information. Notice that, if a codestream employs the layer-based progression order to optimize the expected PSNR at the receiver, a good packet retransmission order is one that follows increasing packet numbering, so that the most important packets are requested first. This approach is not guaranteed to be optimal in terms of quality, because it does not take into account either decoder error concealment by means of the JPEG 2000 error resilience tools, or rate-distortion information possibly contained in the ESD marker segment. However, in our experiments it has turned out to provide excellent results. As a consequence, if the proxy server must only carry out retransmissions, accurate rate-distortion information for different portions of the codestream is not necessary. Nevertheless, it must be noticed that this information is crucial to enable accurate quality control for rate adaptation, or hybrid ARQ and FEC [6, 9].

Based on these remarks, we have considered an ARQ system which processes packets sequentially; in particular, each unacknowledged packet is assigned a maximum number M of retransmissions before the following packet is processed. Transmission is stopped upon reaching N transmission attempts, corresponding to the available bandwidth for the current picture. This is a customary approach when the proxy or base station is not aware of the content of each packet, and attempts to provide the same delay for each packet. As a consequence, some packets may be dropped, and the JPEG 2000 decoder must employ its error resilience capability to conceal the effect of losses. The case for $M = 0$ corresponds to no error protection via ARQ at all; consequently, all the available bandwidth is used for coded source data.

Moreover, we have investigated the behavior of an ARQ system which only enforces a maximum number of retransmissions per picture N, but does not constrain the number of retransmissions for each packet; this is equivalent to the previous ARQ strategy in the limit of $M = \infty$, constrained to the maximum number N. Such system is expected to provide improved performance, since it recognizes that the delay constraint can be enforced on a per-picture basis, and hence employs the maximum number of leftover retransmission for each packet, thus guaranteeing that the first and most important packets in the codestream are assigned a higher number of attempts.

4. EXPERIMENTAL RESULTS

In the following we report experimental results on the transmission of JPEG 2000 coded pictures according to the scenario described above. The MOTHER & DAUGHTER (QCIF) and MISS AMERICA (CIF) pictures have been used to evaluate the performance of the considered ARQ schemes. The JPEG 2000 encoder provides error resilience by using the SOP and EPH markers [1], resetting the contexts and terminating the arithmetic coder at each coding pass, and inserting the segmentation symbol. The layer-progressive order is employed using 50 layers. It is assumed that the headers are received error-free by the terminal. This can be obtained by either using FEC codes to protect the headers as in [7], or by always devoting $M \times N$ retransmissions to the packets that contain the main and tile-part headers. The former strategy has been used in our experiments. Expected PSNR (computed by linearly averaging the MSE) has been used as quality metric, averaging the results of a Monte Carlo simulation over 1000 picture transmissions for each test condition. Random packet loss rates (PLR) ranging from 5% to 20 % have been selected as representative of several wireless scenarios including mobile communications.

In Tab. 1 the results for the MOTHER & DAUGHTER picture at 10 frames/s are reported for several PLRs and data rates. As can be seen, if no retransmissions at all are used (M=0), a poor average quality is obtained also at PLRs as low as 5%, even though the headers are received error-free.

Table 1: Average PSNR for MOTHER & DAUGHTER (QCIF format, 10 fps).

		PSNR (dB)			
PLR	M	64 kbit/s	128 kbit/s	384 kbit/s	512 kbit/s
0.05	0	26.74	27.37	27.64	27.64
0.05	1	29.03	32.07	38.26	39.71
0.05	2	29.10	32.30	40.40	43.57
0.05	∞	29.10	32.30	40.56	43.93
0.10	0	25.60	25.81	25.85	25.85
0.10	1	28.40	30.68	33.42	33.72
0.10	2	28.78	31.87	38.68	40.68
0.10	∞	28.82	32.01	39.97	43.18
0.15	0	24.76	24.84	24.86	24.86
0.15	1	27.58	29.13	30.36	30.45
0.15	2	28.42	31.26	36.33	37.42
0.15	∞	28.56	31.72	39.36	42.49
0.20	0	24.11	24.15	24.15	24.15
0.20	1	26.90	27.93	28.50	28.53
0.20	2	27.94	30.36	33.56	34.02
0.20	∞	28.29	31.42	38.76	41.68

This points out that some form of error control is required in order to provide the user with an acceptable quality level. In fact, with M=0 the received quality is dominated by the position of the first lost packet, and this justifies the fact that PSNR is little dependent on the data rate.

Employing one or two retransmissions yields a large PSNR gain, showing the suitability of the ARQ approach for this application. It is worth noticing that the $M = \infty$ approach always provides the best results, highlighting the validity of the proposed approach. However, at PLRs as low as 5%, two retransmissions yield quality very close or equal to the maximum values. On the other hand, for very harsh conditions, the gain of the $M = \infty$ approach becomes significant, pointing out that limiting the number of retransmissions per packet is a suboptimal strategy. This is due to the fact that the probability of losing one of the first and most critical packets is high, thus vanishing the quality improvements conveyed by the next received packets.

The performance improvement of the $M = \infty$ approach with respect to $M = 2$ increases with higher channel data rates, since this allows the proxy server to better manage the retransmission of the most critical packets.

Tab. 2 reports the results of the same experiment, encoding the picture at 15 frames/s. Apart from the PSNR decrease due to the increased frame rate (and hence smoother playback), similar comments can be made as for the quality achieved by the considered ARQ strategies at various PLRs and data rates.

Table 2: Average PSNR for MOTHER & DAUGHTER (QCIF format, 15 fps).

		PSNR (dB)			
PLR	M	64 kbit/s	128 kbit/s	384 kbit/s	512 kbit/s
0.05	0	25.88	27.10	27.61	27.64
0.05	1	27.07	30.29	35.94	37.52
0.05	2	27.10	30.40	36.88	39.15
0.05	∞	27.10	30.40	36.93	39.26
0.10	0	25.15	25.74	25.85	25.85
0.10	1	26.76	29.38	32.66	32.22
0.10	2	26.91	30.00	35.84	37.76
0.10	∞	26.93	30.07	36.37	38.72
0.15	0	24.51	24.82	24.86	24.86
0.15	1	26.30	28.27	30.08	30.29
0.15	2	26.71	29.53	34.46	35.79
0.15	∞	26.78	29.76	35.92	38.22
0.20	0	23.97	24.13	24.15	24.15
0.20	1	25.91	27.38	28.40	28.48
0.20	2	26.44	28.90	32.58	33.32
0.20	∞	26.60	29.44	35.51	37.76

Tab. 3 and 4 show the results of the same experiment for a CIF image encoded at 5 and 10 frames/s respectively. It can be observed that the same considerations as before still hold for this higher resolution image format. As a consequence, it can be seen that the proposed $M = \infty$ strategy consistently provides the best results for this kind of application.

5. DISCUSSION AND CONCLUSIONS

As has been shown, error sensitivity information can be exploited by a proxy server at the base station to optimize the retransmission management. In particular, it has been found that the $M = \infty$ strategy, which does not constrain the number of retransmission attempts for each packet, but rather for each frame, provides improved performance. This is mainly due to the fact that this strategy privileges the retransmission of the first packets of the codestream, achieving a good trade-off between channel rate and delivered quality. If the codestream has been encoded in layer-progressive order, the default mode for error sensitivity description turns out to be adequate.

Table 3: Average PSNR for MISS AMERICA (CIF format, 5 fps).

		PSNR (dB)			
PLR	M	64 kbit/s	128 kbit/s	384 kbit/s	512 kbit/s
0.05	0	21.46	21.47	21.48	21.48
0.05	1	33.00	34.40	35.56	35.75
0.05	2	35.23	38.08	42.16	43.36
0.05	∞	35.23	38.10	42.29	43.55
0.10	0	19.58	19.58	19.58	19.58
0.10	1	28.66	29.07	29.28	29.30
0.10	2	33.75	35.67	37.60	37.95
0.10	∞	34.97	37.86	42.11	43.27
0.15	0	17.94	17.94	17.94	17.94
0.15	1	24.68	24.80	24.85	24.85
0.15	2	31.26	32.24	32.99	33.09
0.15	∞	34.73	37.60	41.88	43.04
0.20	0	16.85	16.85	16.85	16.85
0.20	1	22.78	22.83	22.84	22.85
0.20	2	28.04	28.45	28.70	28.73
0.20	∞	34.47	37.36	41.59	42.75

Table 4: Average PSNR for MISS AMERICA (CIF format, 10 fps).

		PSNR (dB)			
PLR	M	64 kbit/s	128 kbit/s	384 kbit/s	512 kbit/s
0.05	0	21.38	21.46	21.48	21.48
0.05	1	31.06	33.00	34.94	35.21
0.05	2	32.30	35.23	39.63	40.58
0.05	∞	32.30	35.23	39.69	40.66
0.10	0	19.55	19.58	19.58	19.58
0.10	1	27.89	28.66	29.19	29.23
0.10	2	31.39	33.75	36.56	36.98
0.10	∞	32.03	34.97	39.53	40.46
0.15	0	17.93	17.94	17.94	17.94
0.15	1	24.39	24.68	24.83	24.84
0.15	2	29.75	31.26	32.64	32.79
0.15	∞	31.81	34.73	39.35	40.24
0.20	0	16.85	16.85	16.85	16.85
0.20	1	22.62	22.78	22.84	22.84
0.20	2	27.26	28.04	28.59	28.64
0.20	∞	31.57	34.47	39.11	40.02

On the other hand, it is worth noticing that, if other progression orders are employed, requesting packets sequentially may not be the best approach. As an example, consider the transmission of a color RGB picture in componentwise progression order, i.e. the codestream contains first all data related to the R, then to the G and B components. A typical rate-distortion curve of such a codestream is shown in Fig. 2 for the BOATS image (PSNR is referred to each component separately). It can be clearly seen that, in this case, the codestream portions that mostly contribute to the received quality are not placed at the beginning of the codestream, but at the beginning of each component. As a consequence, the $M = \infty$ approach could be used by retransmitting packets not in sequential order, but according to the rate-distortion information in the ESD data structure.

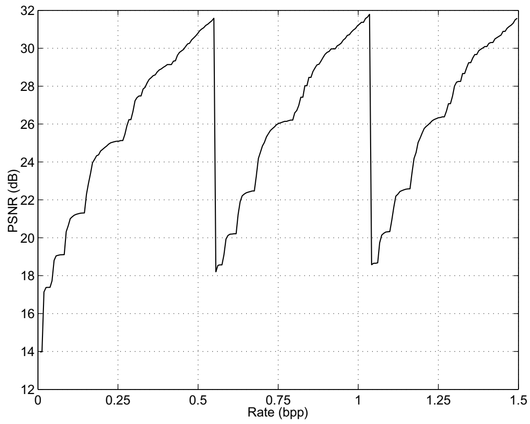

Figure 2: Typical rate-distortion curve of an RGB picture in component-progressive order (the plot refers to the BOATS image).

REFERENCES

[1] ISO/IEC 15444-1, "JPEG 2000 part 1 - core coding system," 2000.

[2] ISO/IEC 15444-3 Amd 2, "Motion JPEG 2000", Oct. 2002.

[3] V. Chande and N. Farvardin, "Progressive transmission of images over memoryless noisy channels," *IEEE J. on Sel. Areas in Comm.*, vol. 18, no. 6, pp. 850–860, June 2000.

[4] B.A. Banister, B. Belzer, and T.R. Fischer, "Robust image transmission using jpeg2000 and turbo-codes," *IEEE J. Sel. Areas Commun.*, vol. 9, no. 4, pp. 117–119, Apr. 2002.

[5] O. G. Sherwood and K. Zeger, "Progressive image coding for noisy channels," *IEEE Sig. Proc. Lett.*, vol. 4, pp. 189–191, July 1997.

[6] A.E. Mohr, R.E. Ladner, and E.A. Riskin, "Unequal loss protection: graceful degradation of image quality over packet erasure channels through forward error correction," *IEEE J. on Sel. Areas in Comm.*, vol. 18, no. 6, pp. 819–828, June 2000.

[7] Document WG1N3111, "JPEG 2000 image coding system - Part 11: Wireless JPEG 2000 - Working Draft version 2.0"

[8] M. Grangetto, E. Magli, G. Olmo, "Error sensitivity data structures and retransmission strategies for robust JPEG 2000 wireless imaging", *IEEE Trans. Cons. Elec.*, special issue on "JPEG 2000: implementations and applications", to appear Nov. 2003.

[9] A. Majumdar, D. Grobe Sachs, I.V. Kosintsev, K. Ramchandran, and M.M. Yeung, "Multicast and unicast real-time video streaming over wireless LANs," *IEEE Trans. on Cir. Sys. Video Tech.*, vol. 12, no. 6, pp. 524–534, June 2002.

[10] ITU-T VCEG-N37, *Common Test Conditions for RTP/IP over 3GPP/3GPP2 - Software and Amendments, VCEG (SG16/Q6).*