EUROPEAN TRANSACTIONS ON TELECOMMUNICATIONS
Eur. Trans. Telecomms. (in press)
Published online in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/ett.1168

Letter

Communication Networks

E-model based comparison of multiple description coding and layered coding in packet networks

Yugang Zhou* and Wai-Yip Chan

Department of Electrical and Computer Engineering, Queen's University, Kingston, Ontario, K7L 3N6, Canada

SUMMARY

We examine the performance of multiple description coding (MDC) with and without the use of automatic repeat request (ARQ) protocols for packet network communication, in comparison with layered coding (LC). The rate-distortion lower bound of MDC and LC are incorporated into an E-model based performance measure, which accounts for the additional costs of excess rates and delay incurred from using ARQ. The results show that the relative merits of the schemes depend on the values of the packet loss rates and round-trip-time (RTT). LC is superior for small RTT and unaided MDC is superior for large RTT. For moderate RTT, LC is preferred for small packet loss rates and MDC aided by ARQ is preferred for large packet loss rates. Copyright © 2007 John Wiley & Sons, Ltd.

1. INTRODUCTION

Layered coding (LC) and multiple description coding (MDC) have been proposed as source-coding schemes that can offer robust communication over packet loss channels. In contrast to traditional coding schemes that generate a single bitstream, LC and MDC generate two or more bitstreams. In LC, one bitstream is sent as a base layer and the other bitstreams as progressive enhancement layers. The base layer has higher priority than the enhancement layers, as the loss of the base layer makes the information received from the enhancement layers useless. An enhancement layer can be decoded only if the base layer and all previous enhancement layers have been correctly received. Automatic repeat request (ARQ) or forward error correction (FEC) is generally employed for error control in the base layer. In the rest of this paper, unless otherwise noted, the base layer of LC is transmitted using ARQ (and requisite error detection coding). The multiple bitstreams (descriptions) of MDC, however, are equally important. Any subset of the bitstreams is decodable and reconstruction quality improves with the size of the subset. One means of realising MDC is using FEC codes [1].

LC and MDC have been compared mainly in the area of video communications [2–5]. A general view of MDC and LC, according to these comparisons, is that MDC has advantages over LC in error-prone channels with high loss probabilities or applications with strict delay constraints. However, these conclusions were drawn for specific MDC and LC coders in specific communication systems. Which scheme performs better depends on actual coder implementations and channel characteristics. For this reason, observations in the literature are not consistent. For example, Reibman *et al.* [2] compared MD transform-coded video [6] with H.263-coded video over an EGPRS wireless network. They observed that LC performs no worse than MDC. Singh *et al.* [5] compared a MD polyphase

* Correspondence to: Yugang Zhou, Department of Electrical and Computer Engineering Queen's University, Kingston, Ontario, K7L 3N6, Canada.
E-mail: zhouy@ee.queensu.ca

Received 24 October 2005
Revised 20 March 2006
Accepted 13 June 2006

transform [7] coder with a progressive JPEG coder. Based on their simulations, they concluded that MDC performs better than LC for a wide range of network scenarios. Wang *et al.* [4] compared MD motion compensation-coded [8] video with signal-to-noise ratio (SNR) scalable H.263+ video over multi-hop wireless networks. They concluded that MDC outperforms LC only for high channel loss conditions. Lee *et al.* [3] compared MD scalar quantized [9] MPEG video with SNR-scalable MPEG video. A 3-D set partitioning in hierarchical trees (3D-SPIHT [10]) based MDC video scheme was also compared with a SPIHT LC scheme. They concluded that ARQ-based or FEC-based LC outperforms MDC for low to medium channel loss rates.

In this paper, instead of using actual MDC and LC coders, we examine the performance of two-description MDC and two-layer LC based on rate-distortion (RD) lower bounds. In Reference [11], Reibman *et al.* observed that sophisticated MD source coding cannot achieve good performance without appropriate channel coding for memoryless channels. This suggests that MDC might have better performance when aided by other forms of error resilient measure, such as ARQ. Therefore, we also examine the performance of MDC+ARQ, where ARQ is applied when all descriptions are lost during transmission. MDC without using ARQ is considered in this paper as a special case of MDC+ARQ performed with only one transmission attempt. The use of ARQ increases the likelihood of successful transmissions, while increasing the total transmission bit rate and delay. We incorporate these additional costs into our performance evaluation. For a fair comparison, the excess bit rates incurred by using ARQ are compensated by allowing MDC to use the excess amount of bits for source coding. The excess delay is mapped to an impairment factor and included in the evaluation. The overall quality is evaluated using a performance model combining the effect of both rate and delay. By mapping RD performance to a rating factor that is combined with the impairment factor of delay, schemes without using ARQ are compared objectively with those using ARQ. It is worth noting that while in the above LC and MDC are regarded as two competitive schemes, they can also complement one another; a hybrid layered MDC scheme can be advantageous for multicasting [12]. A related work on comparing MDC with FEC over Gilbert channels can be found in Reference [13], where MDC is shown to be better in rate-distortion performance. Preliminary results from this work, including results for two variant ARQ schemes, were presented in Reference [14].

2. RATE-DISTORTION PERFORMANCE

Two-layer LC is considered. A source X is encoded to a base layer and an enhancement layer with bit rate R_1 and R_2, respectively. When information from both layers are available, $\hat{X}_1 + \hat{X}_2$ is reconstructed. When information from only the base layer is available, $\hat{X}_1$ is reconstructed. Denote the achievable RD region as the quadruples (R_1, R_2, d_l^1 and d_l^0), where d_l^1 is the reconstruction distortion with information from the base layer only and d_l^0 is the distortion with information from both the base layer and enhancement layer. Given the achievable region results by Tuncel *et al.* in Reference [15], the achievable quadruple (R_1, R_2, d_l^1 and d_l^0) for a zero-mean unit-variance i.i.d. Gaussian source can be shown to satisfy

$$d_l^1 \geqslant 2^{-2R_1}, \quad d_l^0 \geqslant 2^{-2(R_1+R_2)} \tag{1}$$

Correlated sources such as speech and video can be approximated as i.i.d. Gaussian after decorrelation processing such as in transform coding.

The achievable RD region for two-description MDC of a zero-mean unit-variance i.i.d. Gaussian source using square-error distortion measure was studied by Ozarow in Reference [16]. Finding an operating point to tradeoff between central and side distortion becomes a fundamental problem of MDC. Denote d_m^1 and d_m^0 as the side and central distortion of balanced MDC. For a total bit rate R_m for the two descriptions, d_m^0 satisfies [16]

$$d_m^0 \geqslant \frac{2^{-2R_m}}{1 - \left(1 - d_m^1 - \sqrt{{d_m^1}^2 - 2^{-2R_m}}\right)^2}$$
$$\text{for} \quad 2^{-R_m} \leqslant d_m^1 \leqslant \frac{1}{2}(1 + 2^{-2R_m}) \tag{2}$$

Note that $d_m^0 \geqslant \beta 2^{-2R_m}$, where $\beta \geqslant 1$ is a penalty factor the RD performance of MDC suffers, in comparison with Equation (1) for $R_m = R_1 + R_2$. The achievable side-central distortion region for LC with $R_1 = R_2 = 1$ and for MDC with $R_m = 2$ bits per sample is shown in Figure 1. Clearly, LC outperforms MDC in the lower left region. However, Figure 1 is drawn without accounting for the extra costs required to ensure the base layer transmission in LC. We are interested in fairly comparing LC with MDC in packet networks by accounting for the extra costs required to operate LC.

Consider the scenario in which each layer of LC is transmitted through a channel with packet loss rate p. ARQ is applied to the base layer and the maximum allowed

DOI: 10.1002/ett

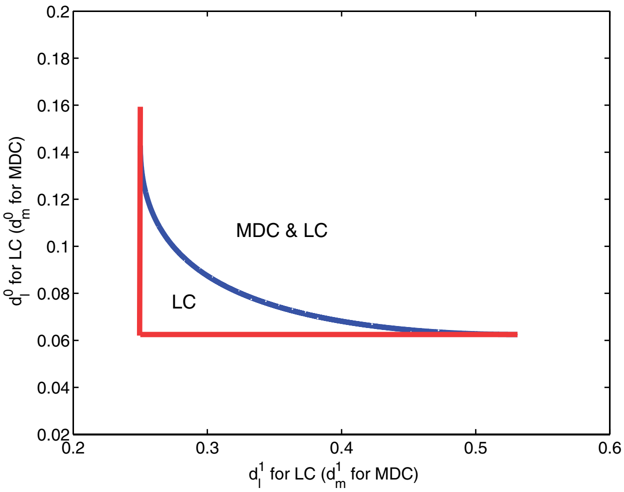

Figure 1. Achievable side-central-distortion regions for LC with $R_1 = R_2 = 1$ bit per sample and for MDC with $R_m = 2$ bits per sample.

number of transmission attempts is M. Suppose that the base layer is successfully received, the expected distortion is $(1-p)\ 2^{-2(R_1+R_2)} + (p)\ 2^{-2R_1}$, where the two terms correspond to receiving or not receiving of the enhancement layer, respectively. Since the loss probability of the base layer is p^M, the expected distortion for LC of the source under consideration is calculated as

$$D_l = \left(1-p^M\right)\left[(1-p)2^{-2(R_1+R_2)} + (p)\ 2^{-2R_1}\right] + p^M \tag{3}$$

For MDC+ARQ, ARQ is applied when both descriptions are lost. The expected distortion can be evaluated as

$$\begin{aligned}
\bar{D}_{ma} &= \underbrace{(1-p)^2 d_m^0 + 2p(1-p)\, d_m^1 + p^2}_{\text{first transmission}} \\
&\quad \times [\underbrace{(1-p)^2 d_m^0 + 2p(1-p)\, d_m^1 + p^2}_{\text{second transmission}} \\
&\quad \times [\cdots \underbrace{[(1-p)^2 d_m^0 + 2p(1-p)\, d_m^1 + p^2]}_{\text{Mth transmission}}]\cdots] \\
&= \left[1 + p^2 + \cdots + p^{2(M-1)}\right](1-p)^2 d_m^0 \\
&\quad + \left[1 + p^2 + \cdots + p^{2(M-1)}\right] 2p(1-p)\, d_m^1 + p^{2M} \\
&= \frac{1-p^{2M}}{1+p}\left[(1-p)\, d_m^0 + 2p\, d_m^1\right] + p^{2M}
\end{aligned} \tag{4}$$

where d_m^0 satisfies the inequality in Equation (2). The best tradeoff between the central distortion and the side distortion determines the MDC+ARQ operating point. The optimal operating point can be found by first inserting the lower bound in Equation (2) to Equation (4), and then setting the derivative to zero. That is, the lower RD bound is obtained by minimising the expected distortion in Equation (4) over d_m^1,

$$D_{ma} = \min_{d_m^1}\ \bar{D}_{ma}(d_m^1) \tag{5}$$

MDC without using ARQ is regarded as a special case of MDC+ARQ with $M = 1$. The RD lower bound for MDC can be obtained using Equations (4) and (5) with $M = 1$.

The above model only considers packet erasures. We assume that bit errors, if present, can be handled using FEC codes. Packets with residual errors are erased. In this manner, our model covers also the scenario with bit errors.

3. COST OF ARQ IN PACKET NETWORKS

ARQ is a widely used error-recovery approach for data communication over noisy channels. ARQ reduces packet loss rates at the expense of increased overall transmission rates and delay. In this paper, we incorporate the additional costs due to using the standard stop-and-wait ARQ in the performance comparison between LC and MDC. The results for other types of ARQ, such as Go-Back-N and selective retransmission, can be found in Reference [19].

3.1. Excess rate

For LC with the base layer protected by ARQ, the probability of attempting to transmit the base layer information for the i-th time is $p^{i-1}, i \in \{1, \ldots, M\}$. Let the source-coding rate be R_1 for the base layer and R_2 for the enhancement layer. The overall transmission rate accounting for retransmissions can be calculated as

$$\begin{aligned}
R_l^{\text{tot}} &= \left(1 + p + \cdots + p^{M-1}\right) R_1 + R_2 \\
&= \frac{1-p^M}{1-p} R_1 + R_2
\end{aligned} \tag{6}$$

The excess rate of LC, $R_l^{\text{tot}} - R_1 - R_2 = \frac{p-p^M}{1-p} R_1$, becomes larger as p increases.

For MDC+ARQ with retransmission when both descriptions are lost, the probability of attempting to transmit the two descriptions for the i-th time is $p^{2(i-1)}, i \in \{1, \ldots, M\}$. Let the source-coding rate of MDC+ARQ

DOI: 10.1002/ett

and MDC be R_{ma} and R_m, respectively. The overall transmission rate for MDC+ARQ is

$$R_{ma}^{\mathrm{tot}} = \left[1 + p^2 + \cdots + p^{2(M-1)}\right] R_{ma} = \frac{1 - p^{2M}}{1 - p^2} R_{ma} \tag{7}$$

and the overall transmission rate for MDC is identical to its source-coding rate,

$$R_m^{\mathrm{tot}} = R_m \tag{8}$$

When $p = 0$, the transmission rates for all three schemes are identical to their source-coding rates.

For fair comparison, all three schemes, LC, MDC+ARQ and MDC, use the same total average transmission rate which includes any excess rate due to retransmissions. To this end, we set the overall transmission rates in Equations (6)–(8) to be equal. The source-coding rate for MDC+ARQ is therefore given by

$$R_{ma} = \frac{1 - p^2}{1 - p^{2M}} \left(\frac{1 - p^M}{1 - p} R_1 + R_2 \right) \tag{9}$$

and the source-coding rate for MDC is given by $R_m = R_l^{\mathrm{tot}}$. Excess rate is not incurred in MDC so that more bits are available for source-coding.

3.2. Excess delay

Consider point-to-point communication between two users. The overall transmission time for MDC+ARQ and LC is larger than that for MDC. Let the round-trip-time (RTT) of the network be T_0. Similar to the excess rate analysis in Section 3.1, the overall transmission time for LC can be calculated as

$$T_l = \frac{1 - p^M}{1 - p} T_0 \tag{10}$$

and the overall transmission time for MDC+ARQ is

$$T_{ma} = \frac{1 - p^{2M}}{1 - p^2} T_0 \tag{11}$$

For MDC without using ARQ, no acknowledgement message is required so that the overall transmission time is

$$T_m = \frac{T_0}{2} \tag{12}$$

4. PERFORMANCE MODEL

Both the RD performance and delay affect overall quality. Low RD performance or large delay degrades user perceived quality. However, delay and RD performance are expressed on different scales. A single measure of quality is desired. We are inspired by the ITU-T G.107 E-model [17, 18], which is used to evaluate end-user opinion of narrowband telephone service quality. The E-model assesses the combined effect of various transmission parameters. It is based on the assumption that transmission impairments can be represented as psychological factors on a scale such that the factors are additive. The combined effect is represented with a rating reflecting the user opinion of service quality.

The prime criterion for assessing the quality of audio and video communications is subjective quality. However, subjective evaluation involving human subjects is typically time intensive and costly. Objective quality models that offer accurate prediction of user perceived quality are convenient and economical. ITU-T G.107 Annex B provides a relationship (Table 1) between the E-model quality rating and mean opinion scale (MOS) for conversational situations. A rating of 80 or more is considered good quality.

We adopt the E-model framework for estimating end-user opinion of the quality of real-time audio and audio-visual communication services. User opinion is assessed by combining the effect of distortion and delay as prescribed by the E-model. While the conventional measure of mean squared error may not be directly related to subjective quality, there exists high correlation between them. For example, the widely used peak signal-to-noise ratio (PSNR) for image and video coding can be converted to MOS [19] using the mapping in Table 2. For speech coding, a correlation of 0.77–0.78 exists between segmental SNR and subjective voice quality [20]. Following the E-model, we define a quality measure

$$W_E = W_0 - \lambda I_{\mathrm{d}} \tag{13}$$

Table 1. Relation between the E-model rating and MOS for conversational situations [17].

E-model rating (lower limit)	MOS (lower limit)	User satisfaction
90	4.34	Very satisfied
80	4.03	Satisfied
70	3.60	Some users dissatisfied
60	3.10	Many users dissatisfied
50	2.58	Nearly all users dissatisfied

DOI: 10.1002/ett

Table 2. Possible PSNR to MOS conversion [19].

PSNR (dB)	MOS
> 37	5 (excellent)
31–37	4 (good)
25–31	3 (fair)
20–25	2 (poor)
< 20	1 (bad)

where the rating factor W_0 is associated with the RD performance and I_d is the delay impairment factor. The rating W_E reflects the overall quality accounting for both the RD performance and delay. The model parameter $\lambda \geqslant 0$ can be chosen to reflect the impact of delay for different applications.

Two mapping functions are required for the objective quality measure in Equation (13): one maps SNR to the corresponding rating factor W_0 and the other maps delay to the impairment factor I_d. For the first mapping, we notice that the subjective MOS value varies slightly for very high or low SNR, for example PSNR >37 dB or PSNR<20 dB. For moderate SNR, the MOS value increases approximately linearly as SNR increases. A similar relationship can be found between signal-to-correlated-noise ratio (see ITU-T P.810 [21] for modulated noise reference unit, MNRU) and subjective MOS value for speech coding. In addition, we assume that one can achieve best subjective quality using a source-coding rate of $r_{\max}$. In practice, the value of $r_{\max}$ may vary for different applications. For example, a selectable mode vocoder (SMV) can score a high of 4.1 MOS at a rate of 8.5 kbps, that is, 1.1 bits per source sample, with clean input speech. A H.264 video coder [22] can achieve PSNR of 38 dB at a rate of 1800 kbps, that is, 0.6 bits per source sample, with CIF sequences at 30 Hz frame rate. Therefore, we construct the following mapping to convert from SNR to W_0

$$W_0 = \frac{16.6\gamma}{r_{\max}} + \frac{0.95\gamma}{r_{\max}^3}(\gamma - 0.5\gamma_0)(\gamma_0 - \gamma) \qquad (14)$$

where $\gamma = -10\log_{10} D$ is the SNR in dB and D is the distortion obtained from Section 2. The parameter $\gamma_0 = -10\log_{10} 2^{-2r_{\max}}$ is the SNR theoretically attainable with the source-coding rate $r_{\max}$. The mapping for $r_{\max} = 2$ bits per source sample is depicted in Figure 2.

Consider the mapping function converting delay to the delay impairment factor I_d for multimedia communication. The one-way delay for voice conversation should be less than 400 ms. 'Toll quality' is achieved with delay less than 100 ms [23]. In a video conference call, the one-way delay should be less than 100 ms for the participants to interact naturally [24]. The video stream should be synchronised with the audio stream for audio-visual communication. In this paper, we use the delay impairment mapping function in ITU-T G.107 [17] to calculate I_d:

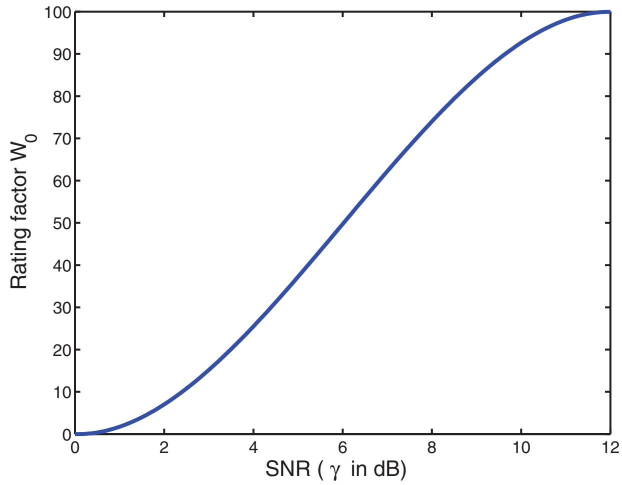

Figure 2. A mapping for converting the SNR to the rating factor W_0 for $r_{\max} = 2$ bits per source sample.

$$I_\mathrm{d} = \begin{cases} 25\left\{[1+Y^6]^{1/6} - 3[1+(Y/3)^6]^{1/6} + 2\right\} \\ \qquad\qquad \text{for} \quad T > 100\,\mathrm{ms} \\ 0, \qquad\quad \text{for} \quad T \leqslant 100\,\mathrm{ms} \end{cases} \qquad (15)$$

where $Y = \log_2(T/100)$ and T is given in Equations (10)–(12). Note that while the above mapping in Equation (15), for speech communication is used in this paper, the parameter λ in Equation (13) can be chosen to adapt the delay impairment factor to other applications. The larger is the value of λ, the more sensitive is the application to delay impairment.

5. NUMERICAL RESULTS

Numerical results are provided for i.i.d. Gaussian source with zero mean and unit variance. Both the excess rate and excess delay are accounted for when LC and MDC+ARQ are compared with MDC. The excess delay depends on the value of RTT, which varies for different types of data networks. The RTT for local networks can be negligible. The RTT for Internet is in the range of 30–300 ms depending on the number of hops between the two users. The RTT for satellite networks can be greater than 600 ms [25].

DOI: 10.1002/ett

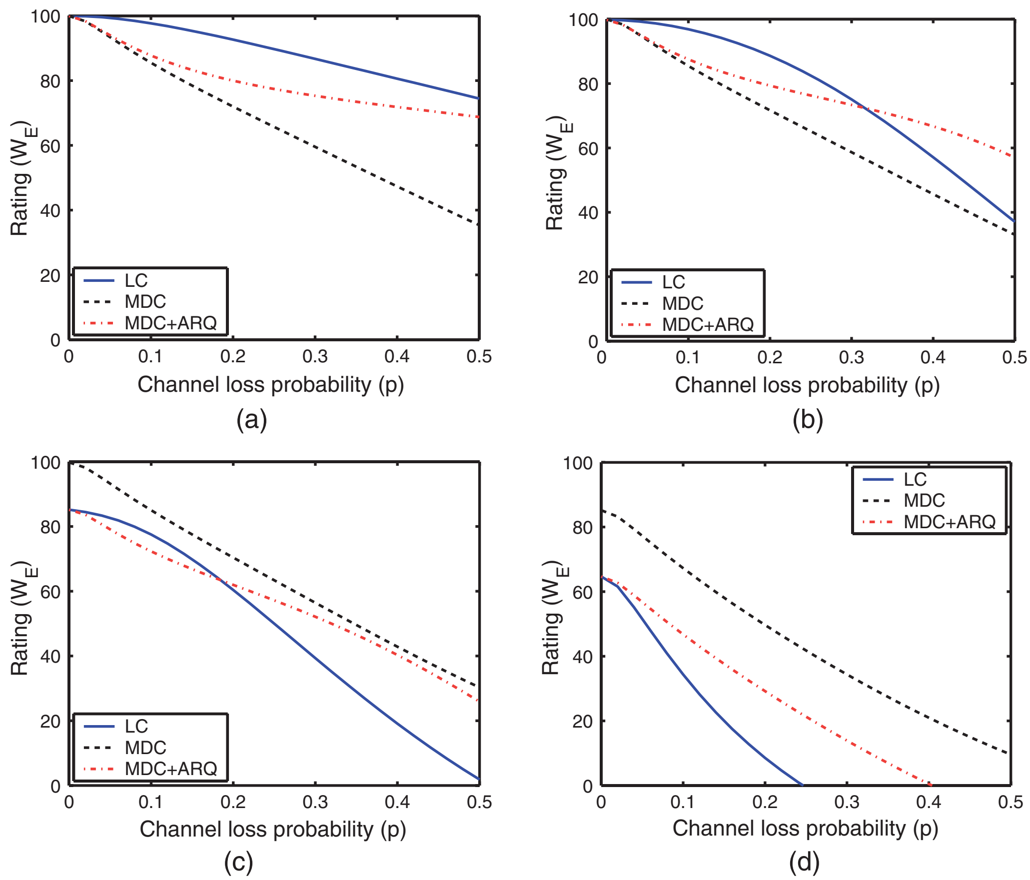

Figure 3. Performance of LC, MDC+ARQ and MDC for various T_0 with $\lambda = 1$, $R_1 = R_2 = 1$ and $r_{\max} = 2$ bits per sample; (a) $T_0 = 40$ ms and $M = 10$; (b) $T_0 = 150$ ms and $M = 3$; (c) $T_0 = 300$ ms and $M = 2$; (d) $T_0 = 600$ ms and $M = 1$.

We set $r_{\max} = 2$ and choose the source-coding rate of LC as $R_1 = R_2 = 1$ bit per sample. Figure 3 shows the overall rating of the three schemes for various types of networks. For local networks or Internet with small RTT (40 ms), the maximum allowed transmission attempts could be large, for example $M = 10$ for a delay constraint of 400 ms. In such case, LC performs better than MDC+ARQ and MDC, as shown in Figure 3(a). As RTT increases, the excess delay of LC and MDC+ARQ increases so that their performance curves drop. The larger is the packet loss rate p, the larger is the delay impairment. MDC incurs no excess delay, hence, the performance curve of MDC changes little until the RTT becomes large. For Internet communication with a moderate RTT of 150 ms, Figure 3(b) shows that MDC+ARQ outperforms LC for $p > 30\%$, as the performance of LC drops faster than MDC+ARQ due to the larger excess delay and rate. For a large RTT of 300 ms in Figure 3(c), the delay impairment from using ARQ is so large that MDC outperforms LC and MDC+ARQ for all p. For satellite networks with very large RTT (600 ms) in Figure 3(d), no retransmission is allowed so that LC and MDC are operated with $M = 1$. In such case, MDC performs the best.

The proposed objective quality measure can be applied to different applications by adjusting the model parameter values, for example λ in Equation (13) and $r_{\max}$ in Equation (14). The value of λ determines the weight of the delay impairment factor. A larger value of λ is associated with delay sensitive applications. The performance curves of LC and MDC+ARQ would drop faster when λ is increased. The value of $r_{\max}$ can also be adjusted to suit different types of applications, for example 0.5–1 bits per sample for video communication. We performed simulations for different values of λ and $r_{\max}$. Similar performance trends were obtained and thus omitted.

DOI: 10.1002/ett

While the above observations are obtained based on specific mapping functions and model parameters, it is highly possible that similar trends would be observed for other reasonable mappings and parameters. Indeed, when RTT is small, the overall transmission delay in Equations (10) and (11) is also small. The associated delay impairment is thus negligible for any reasonable choice of the delay impairment mapping function. In such case, the overall rating is mainly determined by the RD performance. Recall that the gap between LC and MDC in Figure 1 indicates that LC can offer better RD performance than MDC. The gap provides room for LC to outperform MDC when the extra costs from using ARQ are small. Hence, for small p where the excess rate is small, LC still performs better than MDC+ARQ and MDC for small to moderate RTT (Figure 3(a) and (b)). As p increases, the excess rate becomes larger and LC is outperformed by MDC+ARQ. When RTT is large, the delay impairment dominates the performance of all ARQ-aided schemes. MDC performs better than the other two ARQ-aided schemes in that case. With reasonable mapping functions and model parameters, the relative ranking among the three schemes is expected to follow the above trend. Nevertheless, the specific values of RTT and p beyond which one scheme performs better than the others vary, depending on the choice of the mapping functions and model parameters. Future work including experimental verification and refinement of our performance model is desirable.

6. Conclusions

We have compared the performance of LC and MDC, with and without retransmissions, in packet networks. Performance is evaluated using RD lower bounds combined with the effect of excess rate and delay incurred from retransmissions. We summarise based on simulation results for an i.i.d. Gaussian source as following:

- When retransmission is not permitted or retransmission is allowed but the RTT is very large, MDC performs better than LC and MDC+ARQ for all packet loss rates.
- When retransmission is permitted and the RTT is small, LC performs better than MDC and MDC+ARQ for all channel loss rates. For moderate RTT, LC performs the best for low packet loss rates and MDC+ARQ performs the best for high packet loss rates.

In the above analysis of the three schemes, the transmitter is assumed to have knowledge of p. If p is varying sufficiently slowly, the receiver can estimate p and request the transmitter to use the best transmission scheme for the specific value of p. A scheme employing erasure codes to form multiple-description packets from an embedded bit stream is described by Mohr *et al.* in Reference [26]. With this scheme, one source coder capable of generating an embedded bit stream serves both LC and MDC communication modes, so that mode switching can be effected at a lower layer of the protocol stack.

REFERENCES

1. Albanese A, Blömer J, Edmonds J, Luby M, Sudan M. Priority encoding transmission. *IEEE Transaction on Information Theory* November 1996; **42**:1737–1744.
2. Reibman AR, Wang Y, Qiu X, Jiang Z, Chawla K. Transmission of multiple description and layered video over an EGPRS wireless network. In *Proceedings of the IEEE, International Conference Image Processing (ICIP)*, September 2000; **2**:136–139.
3. Lee YC, Kim J, Altunbasak Y, Mersereau RM. Layered coded vs. multiple description coded video over error-prone networks. *Signal Processing-Image Communication* May 2003; **18**: 337–356.
4. Wang Y, Panwar S, Lin S, Mao S. Wireless video transport using path diversity: multiple description vs. layered coding. In *Proceedings of the IEEE International Conference Image Processing (ICIP)*, September 2002; **1**:21–24.
5. Singh R, Ortega A, Perret L, Jiang W. Comparison of multiple description coding and layered coding based on network simulations. In *Proceedings of SPIE Conference Visual Communication Image Processing*, April 2000; PP. 929–939.
6. Reibman AR, Jafarkhani H, Wang Y, Orchard MT, Puri R. Multiple-description video coding using motion-compensated temporal prediction. *IEEE Transaction Circuits Systems of Video Technology* March 2002; **12**:193–204.
7. Jiang W, Ortega A. Multiple description coding via polyphase transform and selective quantization. In *Proceedings of Visual Communications and Image Processing (VCIP)*, January 1999; pp. 998–1008.
8. Wang Y, Lin S. Error-resilient video coding using multiple description motion compensation. *IEEE Transaction Circuits Systems Video Technology* June 2002; **12**(6):438–452.
9. Vaishampayan VA. Design of multiple description scalar quantizers. *IEEE Transactions Information Theory* May 1993; **39**:821–834.
10. Kim J, Mersereau RM, Altunbasak Y. Error-resilient image and video transmission over the Internet using unequal error protection. *IEEE Transactions on Image Processing* February 2003; **12**(2):121–131.
11. Reibman AR, Jafarkhani H, Orchard MT, Wang Y. Performance of multiple description coders on a real channel. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 1999; **5**:2415–2418.
12. Chou PA, Wang HJ, Padmanabhan VN. Layered multiple description coding. In *Proceedings of Packet Video Workshop*, France, 2003; 7 pages.
13. Kim MY, Kleijn WB. Rate-Distortion comparisons between FEC and MDC based on Gilbert channel model. In *Proceedings of IEEE International Conference on Networks (ICON)*, September 2003; PP. 495–500.
14. Zhou Y, Chan W-Y. Performance comparison of layered coding and multiple description coding in packet networks. In *Proceedings of IEEE Global Telecommunication Conferences (GLOBECOM)*, November 2005; 5 pages, St. Louis.

15. Tuncel E, Rose K. Additive successive refinement. *IEEE Transactions on Information Theory* August 2003; **49**(8):1983–1991.
16. Ozarow L. On a source-coding problem with two channels and three receivers. *Bell System Technology Journal* December 1980; **59**(10):1909–1921.
17. ITU-T Recommendation G.107, The E-model, a computational model for use in transmission Planning, March 2005.
18. Takahashi A. Opinion model for estimating conversational quality of VOIP. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2004; **5**:1072–1075.
19. Ohm J. *Bildsignalverarbeitung fuer multimedia-systeme*, Skript, 1999.
20. Quackenbush S, Barnwell T III, Itoh K. *Objective Measures of Speech Quality*. Prentice Hall: New York, 1988.
21. ITU-T Recommendation P.810, Modulated noise reference unit (MNRU), 1996.
22. Wiegand T, Sullivan G, Bjøntegaard G, Luthra A. Overview of the H.264/AVC Video coding standard. *IEEE Transaction Circuit Systems for Video Technology* July 2003; **13**:560–576.
23. ITU-T Recommendation G.114, One-way transmission time, 1996.
24. Baldi M, Ofek Y. End-to-end delay analysis of video-conferencing over packet-switched networks. *IEEE/ACM Transactions Networking* August 2000; **8**:479–492.
25. Dillon D, Butalia GS, Joshi PK. Secure Web Browsing over Long-Delay Broadband Networks—Recommendations for Web Browsers. In *Proceedings of the International Conference E-business and Telecom. Networks (ICETE)* 2004; PP. 160–168. Setúbal, Portugal.
26. Mohr AE, Riskin EA, Ladner RE. Unequal loss protection: graceful degradation of image quality over packet erasure channels through forward error correction. *IEEE Journal on Selected Areas in Communication* June 2000; **18**:819-829.

DOI: 10.1002/ett