UTILISING OBJECTIVE PERCEPTUAL IMAGE QUALITY METRICS FOR IMPLICIT LINK ADAPTATION

Tubagus Maulana Kusuma, Manora Caldera, and Hans-Jürgen Zepernick

Western Australian Telecommunications Research Institute
39 Fairway, Nedlands, WA 6907, Australia
E-mail: {mkusuma, caldera, hans}@watri.org.au

ABSTRACT

An implicit link adaptation technique based on hybrid automatic repeat request (H-ARQ) using a soft-combining algorithm is considered for image transmission over wireless channels. Instead of the conventional link-level metrics such as bit error rate and frame error rate that do not correlate well with the quality as perceived by the end user, adaptation is carried out using metrics obtained from models that take the human perception into account. In this paper, a novel reduced-reference hybrid image quality metric (RR-HIQM) is proposed and used as a criterion to terminate the retransmission request in link adaptation. Numerical results show that the use of RR-HIQM in link adaptation provides robust link performance while meeting satisfactory quality constraints.

1. INTRODUCTION

In recent years, there has been an increasing demand for multimedia communications involving voice, image, and video over wireless channels. These services use compressed data in order to reduce the storage requirements and the bandwidth. Furthermore, the time-varying nature of the wireless channel caused by multipath propagation and the changing interference conditions make the channel very unreliable. Therefore, the multimedia services over wireless channels are impaired not only by lossy source compression but also by burst errors associated with the channel. Additionally, limitations are imposed by the available bandwidth. These have led to extensive research aimed at achieving reliable and spectrally efficient wireless multimedia communications.

Link adaptation in wireless communications has gained a great deal of attention due to the significant performance gains that it provides [1]–[3]. Explicitly, with link adaptation, the transmission parameters are modified to compensate for the variations in channel conditions, which may be estimated based on feedback from the receiver. Currently, link adaptation techniques for wireless communications are based on conventional measures such as signal-to-noise ratio (SNR), bit error rate (BER), frame error rate (FER) or combinations of them as the quality indicators or the metrics. However, in case of multimedia communications, it has been shown that these metrics do not necessarily correlate well with the quality as perceived by humans [4], [5].

Therefore, in this paper, we aim at link adaptation techniques for image transmission over wireless channels focusing on the use of objective perceptual quality metrics, which take into account the characteristics of human perception. This requires identification and development of novel objective perceptual quality measures along with the suitable theoretical framework, methods and algorithms for exploiting these measures for link adaptation purposes. An implicit link adaptation technique based on hybrid automatic repeat request (H-ARQ) using a soft-combining algorithm [6] is considered in this paper to provide robust link performance while meeting objective quality constraints.

This paper is organised as follows. In Section 2, link adaptation based on H-ARQ is discussed. Section 3 presents an overview of the objective perceptual image quality measures. The proposed hybrid image quality metric is detailed in Section 4. Numerical results are presented in Section 5 and Section 6 concludes the paper.

2. IMPLICIT LINK ADAPTATION

Link adaptation methods based on H-ARQ have been frequently used in wireless multimedia communications to improve the quality of service. H-ARQ is an implicit link adaptation technique, which uses an error-correcting code to detect and if possible to correct transmission errors. In this case, the link-layer acknowledgments are used for retransmission decisions. With conventional H-ARQ schemes, whenever a codeword is to be retransmitted, the initial codeword is discarded and replaced by its retransmitted copy. Thus the decoder ignores the information gathered through the failed decoding attempt leading to largely reduced throughput efficiencies over noisy channels. This is avoided in this paper using a

0-7803-8554-3/04/$20.00 ©2004 IEEE.

soft-combining algorithm, which preserves the information obtained with each decoding attempt and incorporates it with the retransmitted copies of the codeword. With this, the reliability information is included in the soft-combining algorithm using the soft values on a symbol-by-symbol basis.

Since all the transmitted blocks may not need the same number of retransmissions to achieve reliable decoding, a performance measure, which characterises the channel variations, is required before terminating the request for retransmissions. As the commonly used performance measures such as BER and FER do not necessarily correlate well with the quality as perceived by the end user, considerable attention has been paid on the development of objective perceptual quality measures.

3. OBJECTIVE PERCEPTUAL IMAGE QUALITY MEASURES

Most of the available objective perceptual image quality measures focus only on errors introduced by compression algorithms. However in wireless communications, additional errors occur due to the time-varying nature of the channel and the multipath effect, introducing further perceptual artifacts. The mean square error (MSE) and the peak signal-to-noise ratio (PSNR) are among the most widely adopted statistical metrics for measuring the picture quality. However, quality results based on both of these metrics are poorly correlated with subjective test results. For example, in some cases, even though the decoded image would provide a high PSNR result, the subjective quality could be perceived low by the human eye providing a low mean opinion score (MOS) [7]. Moreover, these measures are not suitable for real time monitoring as required in link adaptation techniques for wireless multimedia communications due to the unavailability of the reference image at the receiver. Therefore, major emphasis in recent research is given to the measurement techniques that do not require a reference, and provide high correlation with the results of subjective measurement methods [8]. In this paper, a suitable objective perceptual image quality metric for wireless channels is developed and utilised in link adaptation.

4. REDUCED-REFERENCE HYBRID IMAGE QUALITY METRIC

Objective perceptual quality measurement techniques based on the availability of the source signal can be broadly separated into three categories as full-reference, reduced-reference and no-reference. The full-reference and the reduced-reference systems require the knowledge of the original signal as an exact copy or as an extracted version of the source signal in calculating the quality metrics. In contrast to that, the no-reference system does not require a reference signal. Since the reference signal is not naturally available at the receiver, the no-reference systems are more practical in communication systems. However, the quality measures obtained using no-reference systems may not be suitable for link adaptation schemes, as the benchmark used for retransmission termination may not always be reached. Therefore, a reduced-reference hybrid image quality metric (RR-HIQM), which takes into account the artifacts related to blocking, lost block, image activities, blur, and contrast, is proposed in this paper. The algorithms to measure these individual artifacts are selected considering the computational efficiency to minimise the delay associated with them.

The blocking artifact is measured using the algorithm proposed by Wang et al. [9]. The lost blocks are also captured in the proposed RR-HIQM using the blocking metric (BM) given by

$$\mathrm{BM} = \alpha + \beta A^{\gamma 1} B^{\gamma 2} C^{\gamma 3} \tag{1}$$

Here, α, β, $\gamma 1$, $\gamma 2$ and $\gamma 3$ are the model parameters that are estimated using a subjective test. In this paper, the values of $\alpha = -245.9$, $\beta = 261.9$, $\gamma 1 = -0.024$, $\gamma 2 = 0.016$ and $\gamma 3 = 0.0064$ are used as suggested in [9]. Moreover, A, B and C are the average differences across block boundaries, the average absolute difference between in-block image samples, and zero-crossing rate, respectively [9].

The activity measure that indicates the busyness of the received image based on the structure and the approximation of the spatial gradient amplitude, are obtained using the methods in [10] and [11]. This metric has also been used in the RR-HIQM to acquire the artifacts relating to blur, ringing, and lost blocks. In this measure, two types of activities, namely, the edge activity measure (EAM) and the gradient activity measure (GAM) are employed. For an $M \times N$ binary image having M number of rows and N number of columns, the EAM and the GAM are respectively given by [10]

$$\mathrm{EAM} = \left[\frac{1}{MN} \sum_{i=1}^{M \cdot N} B(i) \right] \cdot 100 \tag{2}$$

and

$$\mathrm{GAM} = \frac{1}{MN} \left[\begin{array}{l} \sum_{i=1}^{M-1} \sum_{j=1}^{N} |I(i,j) - I(i+1,j)| \\ - \sum_{i=1}^{M} \sum_{j=1}^{N-1} |I(i,j) - I(i,j+1)| \end{array} \right] \tag{3}$$

In Eq. (2), $B(i)$ denotes the value of the detected edge at pixel location i and $|\cdot|$ and $I(i,j)$ in Eq. (3) denote the absolute value and the intensity value at pixel location (i,j), respectively.

Next, the contrast measure (CM) based on the standard deviation of the first-order image histogram, which indicates the distribution of the image data values, is obtained. With this method, many insights into the character of an image can be acquired. The histogram of an $M \times N$ image is defined as the percentage of pixels within the image at a given gray level. For a 256 gray level image, the histogram h_i at gray level i is given by

$$h_i = \frac{n_i}{MN}, \qquad \text{for} \quad 0 \leq i \leq 255 \tag{4}$$

where n_i denotes the number of pixels at gray level i. From the histogram information, the image perceived brightness can be measured using the average gray level given by

$$brightness = \sum_{i=0}^{255} i \cdot h_i \tag{5}$$

The CM can now be obtained by estimating the average gray level variation within the image given by [11]

$$\text{CM} = \sqrt{\sum_{i=0}^{255} i^2 h_i - brightness^2} \tag{6}$$

Finally, the proposed overall quality measure RR-HIQM is obtained using a weighted sum of all the aforementioned quality metrics. Here, each individual metric behaves as a visual sensor. Since human vision reacts differently to various artifacts, the perceptual weight factor allocation for these individual sensors was based on the impact of the output of the sensor on the overall perceptibility of images by human vision. The fine-tuning of the weights was done empirically and was justified by requesting opinion from a group of unbiased test persons [12].

Initially, all the sensors were given the same weight and were then adjusted based on the contribution of each sensor to the perceptibility of image by human eye [12]. For example, in joint photographic experts group (JPEG) images, the output of the blocking sensor is given a higher weight compared to other sensors. This is because blocking is the most frequently observed artifact in this particular image format and can be easily perceived by human vision.

In utilizing the proposed RR-HIQM in Hybrid-ARQ using soft-combining scheme, the baseline, which is the quality measure of the original image, is incorporated with the first transmission. The baseline, which provides the receiver an idea about the transmitted image is used at the receiver to compare the quality of the received image. The retransmission of the same image is requested until the received image quality reaches the baseline. In some cases, where quality degradation is tolerable, the comparison is done with a predefined threshold from the baseline. As the retransmission termination in the H-ARQ scheme solely depends on the baseline, it is recommended to use a strong error control coding technique to protect the baseline. If a retransmission of the same image is requested, the soft-output values of the received image are incorporated in the retransmitted copies of the image as suggested in the soft-combining algorithm [6]. This process is continued until the received image reaches the required quality or the chosen maximum number of retransmissions.

5. NUMERICAL RESULTS

The performance of the proposed implicit link adaptation scheme based on H-ARQ is obtained using computer simulations. The developed RR-HIQM is used for retransmission termination. In the simulations, a test JPEG image "House" with quality baseline of 1.32 is transmitted 50 times over an uncorrelated flat Rayleigh fading channel in the presence of additive white Gaussian noise (AWGN) with an average bit energy to noise power spectral density ratio (E_b/N_0) of 5dB. A $(31, 21)$ Bose-Chaudhuri-Hocquenghem (BCH) code is used for error protection purposes and modulated using binary phase shift keying (BPSK). The maximum number of retransmissions in the soft-combining algorithm is set to 10.

For this test scenario, the perceptual weights of the visual sensors BM, EAM, GAM, and CM were selected empirically as 1, 1, 0.5, and 0.3, respectively. The results obtained using the proposed RR-HIQM are compared with that of PSNR as it is the most commonly used quality metric for testing image and video communication systems.

Figures 1(a) and 1(b) show the average performance comparison between PSNR and RR-HIQM of the received image, obtained by averaging the individual measures considering fully openable images. The PSNR and the RR-HIQM values corresponding to the images that are not openable are defaulted at 0 and 100, respectively. It is observed from the figures, that the quality metric provided by the RR-HIQM is inversely related to the PSNR. That is, a better quality image is represented by a higher PSNR value or a lower RR-HIQM value. Furthermore, it is observed that the quality measures based on the RR-HIQM reach the quality baseline within a tight margin after the 3^{rd} retransmission. Hence, the retransmission request in the given example can be terminated after the 4^{th} retransmission. On the other hand, it is observed that the PSNR value never comes close to the reference quality. Thus, the termination will never occur and the link adaptation will continue until the maximum number of retransmissions allocated is reached, leading to a very inefficient scheme. Furthermore, the proposed RR-HIQM provides a better quality measure as perceived by the end user that can be incorporated into link adaptation as a method to more efficiently terminate the retransmission request.

Fig. 2 shows the progress of image quality improvement with each retransmission. The images presented in this figure represent the ones having the worst performance after the 1^{st} retransmission. It is observed that the H-ARQ scheme using soft-combining provides a large improvement in the image quality with each retransmission. Moreover, a further improvement in the quality is not observed after the 3^{rd} retransmission. This gives an indication to the system of when to terminate the retransmissions. Again, the proposed RR-HIQM constitutes an extra criterion for stopping the retransmissions.

6. CONCLUSIONS

A novel reduced-reference hybrid image quality metric that takes into account the human perception was proposed. This metric has been used as a criterion to terminate the retransmission request in link adaptation based on H-ARQ using a soft-combining algorithm over wireless channels. Moreover, the RR-HIQM provides a good quality estimate of the received image without needing the reference image at the receiver. Also, the numerical results show that the use of proposed objective perceptual quality measure in link adaptation over wireless channels provides more efficient use of the limited bandwidth while meeting the satisfactory perceived quality constraints.

7. REFERENCES

[1] A. J. Goldsmith and Soon-Ghee Chua, "Variable-rate Variable-power MQAM for Fading Channels," IEEE Trans. on Commun., vol. 45, no. 10, pp. 1218-1230, Oct. 1997.

[2] D. L. Goeckel, "Adaptive Coding for Time-varying Channels Using Outdated Fading Estimations," IEEE Trans. on Commun., vol. 47, no. 6, pp. 844-855, June 1999.

[3] J. M. Torrance and L. Hanzo, "Latency and Networking Aspects of Adaptive Modems over Slow Indoors Rayleigh Fading Channels," IEEE Trans. on Veh. Tech., vol. 48, no. 4, pp. 1237-1251, July 1999.

[4] S. Winkler, E. D. Gelasca, and T. Ebrahimi, "Perceptual Quality Assessment for Video Watermarking," in Proc. of Inter. Conf. on Infor. Tech., Coding and Computing, Nevada, USA, pp. 90-94, April 2002.

[5] A. W. Rix, A. Bourret, and M. P. Hollier, "Models of Human Perception," Journal of BT Technology, vol. 17, no. 1, pp. 24-34, Jan. 1999.

[6] H.-J. Zepernick, B. Rohani, and M. Caldera, "A Soft-combining Technique for LUEP Codes," IEE Electronics Letters, vol. 38, no. 5, pp. 234-235, February 2002.

[7] K. N. Ngan, C. W. Yap, and K. T. Tan, *Video Coding for Wireless Communications Systems*, Marcel Dekker, N.Y., 2001.

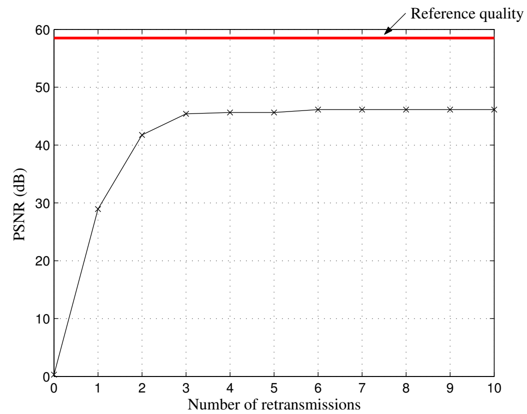
[8] Z. Yu and H. R. Wu, "Human Visual System Based Objective Digital Video Quality Metrics," in Proc. of 5th International Conf. on Signal Processing, Beijing, China, vol. 2, pp. 1088-1095, August 2000.

[9] Z. Wang, H.R. Sheikh, and A.C. Bovik, "No-Reference Perceptual Quality Assessment of JPEG Compressed Images," in Proc. of IEEE Int. Conf. on Image Proc., New York, USA, pp. 477-480, Sept. 2002.

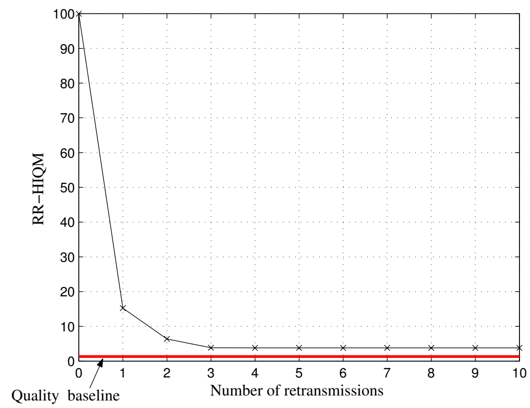
[10] S. Saha and R. Vemuri, "An Analysis on the Effect of Image Activity on Lossy Coding Performance," in Proc. of IEEE Int. Symp. on Circuits and Systems, Geneva, Switzerland, pp. 295-298, May 2000.

[11] A. R. Weeks, *Fundamentals of Electronic Image Processing*, SPIE/IEEE Series on Imaging Science and Engineering, 1998.

[12] T. M. Kusuma and H.-J. Zepernick, "In-Service Image Monitoring Using Perceptual Objective Quality Metrics," Journal of Electrical Engineering, vol. 54, no. 9-10, pp. 237-243, Dec. 2003.



(a)



(b)

Fig. 1. Average performance of link adaptation for image "House": (a) Using PSNR and (b) Using RR-HIQM.

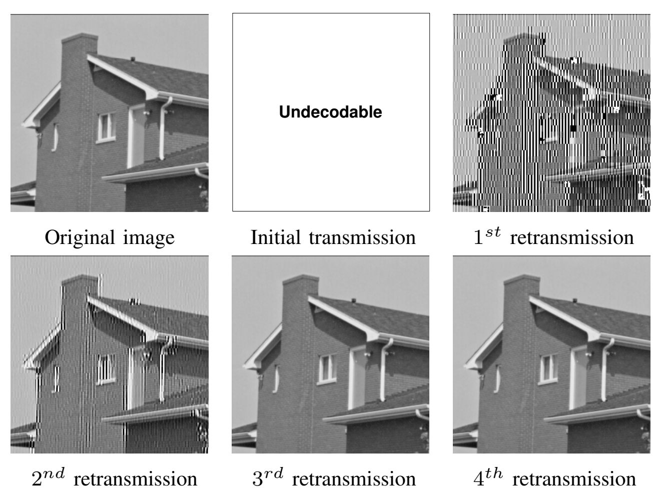


Fig. 2. Quality improvement of image "House".