Reduced Reference Watermark-Based Image Transmission Quality Metric

Mohammad Hosain Kayvanrad
Embedded Video (EV) Laboratory
Department of ECE
National University of Singapore
Block E4, Level 8, Room 27, 4
Engineering Drive, Singapore
117576
mohammad@nus.edu.sg

Shabnam Sodagari
Alireza Nasiri Avanaki
Control and Intelligent Processing
Center of Excellence
School of Electrical and Computer
Engineering, University of Tehran
shabnam@ieee.org
avanaki@ut.ac.ir

Hossein Ahmadi-Noubari
Department of ECE
University of British Columbia
University of Tehran
noubari@ece.ubc.ca

Abstract— **A novel reduced reference (RR) method of channel estimation for image transmission through lossy channels is presented. Estimation is based on the amount of distortion the low frequency content of the image or its histogram suffers during transmission. This distortion is computed by sending the low frequency content of the original image or the corresponding histogram to the receiver by means of a robust watermark. Simulations for different types of images and different channel losses yield large amounts of correlation between the estimated Mean Square Error (MSE) and the structural similarity (SSIM) index.**

Keywords- image transmission; reduced reference quality assessment; watermarking; two dimensional wavelet transform; histogram

I. INTRODUCTION

The problem of estimating channel loss effects on image transmission is important in order to maintain image quality. It is desirable to have a quantitative measure of the distortion being introduced to the image as close to perceptual judgments as possible. This measure might be used by service providers, for instance, in order to charge the customers based on the quality of image or video they receive. It is ideal to have full access to the original image or full reference quality assessment, but this is not the case in almost all practical applications. Therefore, reduced reference quality metric has been investigated in the literature [1]-[9] where the receiver has access to part of the original image or some specific features of it.

Obviously, some side information has to be sent to the receiver in addition to the original image. Nevertheless, in practice there might be no provision of a side channel. Therefore, watermarking can be used in order to embed some features of the transmitted image within itself [10]. This way, the receiver can extract the embedded features and use them as a reference for quality assessment.

In this work, we propose a channel estimation method in which assessment is based on the distortion introduced to low frequency content of the image or its histogram during transmission. Low frequency approximation coefficients of the image are computed using 2DWT and are sent to the receiver using a robust wavelet-based watermark. While in traditional watermarked-based approaches a known watermark is embedded on the image, the watermark here is some feature of the original image. At the receiver, 2DWT is again applied to the received image and the watermark is extracted. Then the Mean Square Error (MSE) between the original and received approximation coefficients or the relevant histograms is computed to quantify channel distortion. In order to estimate the efficiency of the proposed method, the estimated error (MSE) is compared to the SSIM index computed between the original image and its distorted version through the same channel. Simulations show that for various channel loss models this method results in high correlation with structural similarity (SSIM) [3].

The organization of the paper is as follows: In section II a review of previous work is presented. Our method is presented in section III. Section IV describes various simulation results and section V concludes the paper.

II. REVIEW OF PREVIOUS WORK

The aim of image quality assessment is to quantify the quality in a way that closely represents the judgment of a human observer. Three main categories of quality assessment methods exist i.e., Full Reference (FR), Reduced Reference (RR) and No Reference (NR). In the full reference category the original and the distorted image are compared in order to compute the quality. Reduced reference methods compare specific features of the original and the distorted image and no reference or blind quality assessment methods do not require any information from the original image. A performance analysis of various image quality metrics has been presented in [1].

We are interested in the reduced reference methods since we consider image transmission through lossy channels, where it is possible to send partial description of the image to the receiver. The performance of the purposed RR method is, however, evaluated by comparing the results to Structural

978-1-4244-1688-2/08/$25.00 ©2008 IEEE

Similarity (SSIM) index which is an FR method of assessment.

Several approaches have been adopted to develop reduced reference image quality metrics: in [2] a method to evaluate the quality of distorted images that only needs to know the invariant representation of the original image is presented. The authors in [3] have developed a reduced reference perceptual image quality metric which can be applied to in-service quality monitoring and link adaptation. Carnec et al. [4] discuss the choice of description features in reduced reference image quality assessment and show that using structural information and combining various feature types yield better results. Based on the analysis of local harmonic strength, which is computed from the gradient of images, a reduced reference quality metric can be developed [5]. The authors in [6] describe an RR image quality assessment metric based on multi-resolution analysis of the image followed by multi-scale edge presentation.

The use of structural similarity for image quality assessment is proposed in [11] and a structural similarity index has been developed. The structural similarity image quality paradigm is based on the assumption that the human visual system is highly adapted for extracting structural information from the scene, and therefore a measure of structural similarity can provide a good approximation to perceived image quality [12].

Watermarking can be used to acquire knowledge about channel imperfections through embedding and then extracting the watermark within the image and thereby detecting and compensating for channel errors [13]-[15].

In [16], as a blind measure for the quality of the communication link, a fragile watermark is embedded into the video stream. The MSE between the estimated and actual watermarks is then evaluated.

In the next section, we describe our method of computing the channel error in terms of MSE.

III. THE PROPOSED CHANNEL ERROR ESTIMATION SCHEME

Since most of perceptible information of images lies in their low-frequency content, in an attempt to achieve a quality measure close to perceptual assessments, we base our transfer quality metric on the amount of error the low frequency content of the image tolerates during transmission. To estimate this error the low frequency content is sent beside the original image through the same channel but in a *robust* way. It is then used, at the receiver, to compute the amount of error introduced by the channel during transmission.

The low frequency content can be sent to the receiver by means of a robust watermark. Use of wavelet transform to extract the low frequency content and sending it to the receiver via a wavelet-based watermark let us reduce the computational cost to a great deal. In this manner watermark samples, i.e. the approximation coefficients of the wavelet transform of the original image, are embedded as soon as they are computed without any need of another change of domain, which is inevitable in most watermarking schemes.

Clearly we need a blind wavelet-based watermarking scheme with as small perceptual effects on the original image as possible which is as well robust to losses introduced by the channel. The proposed scheme is described below:

Consider an image of dimensions $M \times N$. Applying l-level wavelet transform to this image we get approximation and detail coefficients, which we call $a(n)$ and $d(n)$ respectively.

$$d = [lh_1, hl_1, hh_1, lh_2, hl_2, hh_2, ..., lh_l, hl_l, hh_l] \quad (1)$$

where lh_i, hl_i, and hh_i are horizontal, vertical, and diagonal coefficients of the ith level of decomposition respectively, and all are of length $\frac{N \times M}{2^{2i}}$, for $i = 1,2,...,l$. The approximation coefficients, a, is of length $\frac{N \times M}{2^{2l}}$.

In this stage assume hh_1 coefficients have smooth variations, especially those on the lower and upper indices. We will see later that this assumption holds to a great extent for decomposition by certain wavelets and to certain levels. For $l \geq 2$ it is possible to redundantly embed a into the diagonal coefficients of the first level of decomposition, hh_1, in the following way:

$$hh_1^{embedded}(2i-1) = hh_1(2i-1) + c(a,d)a(i) \quad (2)$$

$$hh_1^{embedded}(2i) = hh_1(2i) - c(a,d)a(i) \quad (3)$$

$$hh_1^{embedded}(\frac{NM}{4} - 2i + 1) = hh_1(\frac{NM}{4} - 2i + 1) + c(a,d)a(i)$$

(4)

$$hh_1^{embedded}(\frac{NM}{4} - 2i) = hh_1(\frac{NM}{4} - 2i) - c(a,d)a(i)$$

(5)

In which $C(a,d)$ is a normalizing coefficient.

`Since the approximation coefficients, a , are of very high energy compared to other components, direct embedding of them into any detail component leads to severe distortions in the image. To overcome this problem we need a normalizing coefficient before embedding. Observations prove that a multiple of the ratio of the maximum of the detail coefficients to that of the approximation coefficients is an appropriate choice. Furthermore, the choice of the multiplicative constant determines the compromise between visible effects and the robustness of the watermark.

At the receiver, wavelet and scaling coefficients of the received image are again computed and the watermark is extracted:

$$\hat{a}(i) = [\frac{\widetilde{hh}_1^{embedded}(2i-1) - \widetilde{hh}_1^{embedded}(2i)}{2c(\widetilde{a}^{embedded}, \widetilde{d}^{embedded})} + \frac{\widetilde{hh}_1^{embedded}(\frac{NM}{4} - 2i - 1) - \widetilde{hh}_1^{embedded}(\frac{NM}{4} - 2i)}{2c(\widetilde{a}^{embedded}, \widetilde{d}^{embedded})}]/2 \tag{6}$$

$$\widetilde{hh}_1(2i-1) = \widetilde{hh}_1^{embedded}(2i-1) - c(\widetilde{a}^{embedded}, \widetilde{d}^{embedded})\hat{a}(i) \tag{7}$$

$$\widetilde{hh}_1(2i) = \widetilde{hh}_1^{embedded}(2i) + c(\widetilde{a}^{embedded}, \widetilde{d}^{embedded})\hat{a}(i) \tag{8}$$

$$\widetilde{hh}_1(\frac{NM}{4} - 2i + 1) = \widetilde{hh}_1^{embedded}(\frac{NM}{4} - 2i + 1) - c(\widetilde{a}^{embedded}, \widetilde{d}^{embedded})\hat{a}(i) \tag{9}$$

$$\widetilde{hh}_1(\frac{NM}{4} - 2i) = \widetilde{hh}_1^{embedded}(\frac{NM}{4} - 2i) + c(\widetilde{a}^{embedded}, \widetilde{d}^{embedded})\hat{a}(i) \tag{10}$$

where $\widetilde{a}^{embedded}$, $\widetilde{d}^{embedded}$, $\widetilde{hh}_1^{embedded}$ are respectively the approximation, details, and first level diagonal coefficients of the received image before watermark extraction, and $\widetilde{a}$, $\widetilde{d}$, $\widetilde{hh}_1$ are those after watermark removal. $\hat{a}$ is the extracted watermark which is nothing but the approximation coefficients of the original image. $\widetilde{a}$, $\widetilde{d}$ are then used to reconstruct the image at the receiver.

Finally, the assessment is made by computing the MSE between $\hat{a}$ and $\widetilde{a}$:

$$MSE = \frac{2^{2l}}{NM} \sum_{i=1}^{\frac{NM}{2^{2l}}} (\hat{a}(i) - \widetilde{a}(i))^2 \tag{11}$$

$$PSNR = 10\log(255^2 / MSE) \tag{12}$$

The block diagram of figure 1 shows the procedure.

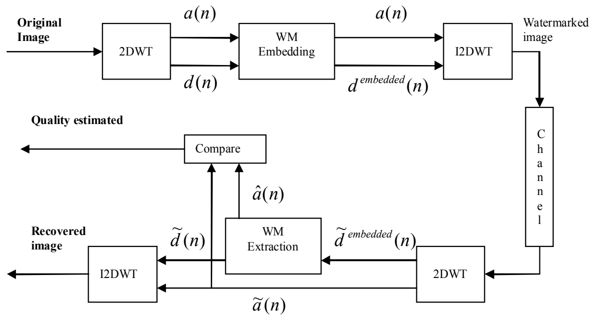

Figure 1. The Block Diagram of the Image Quality Assessment Method

Choice of wavelet as well as the level of decomposition is an important issue. In theory any level of decomposition greater than or equal to two is acceptable. Larger levels of decomposition imply more computational load though. So we have to choose a level as small as possible. However, in order for our first assumption to be true, we should find a specific resolution in which the coefficients have smooth variations for a wide variety of images. Several observations have been made and it has finally been found out that decomposition of the images using bior1.1 leads to acceptable figures in first level diagonal coefficients for decomposition to 3 or more levels. Furthermore, the higher the level of decomposition, the smaller the perceptual effects, since a smaller portion of the coefficients at that level are changed. Lager levels of decomposition deserve higher computational cost, though. So here another trade-off rises. The diagonal coefficients of level 4 seem to be proper candidates for watermark embedding. All simulations in this paper are based on decomposition using bior1.1 to 4 levels.

Although very small, the watermark itself introduces some distortions to the image before it is removed. The distortion is however so small that one can hardly notice. (See figure 2) The SSIM of the original image and its watermarked version can be used as a measure of this distortion. Table 7 shows this measure for the standard test images used in this paper. As the figures suggest, the similarity of the two images is so high that the distortion seems to be quite negligible.

However, in order to further decrease this distortion, we may think of even more compact data to be embedded. To do so we watermark the histogram of the approximation coefficients instead of the whole approximation coefficients.

The watermarking scheme is the same as before and the assessment is again made based on the MSE between the received histogram and that computed at the receiver. This method, however, works well only for losses considerably affecting the histogram of the image, as expected.

IV. SIMULATION RESULTS

We have developed a benchmark for testing our proposed algorithms which simulates channels introducing salt and pepper loss with various loss probabilities, JPEG compression with different qualities, and additive white Gaussian noise with various levels of PSNR. In order to evaluate the performance of the purposed methods, computed quality measure (MSE) is compared to SSIM index which is a widely accepted FR method for measuring the similarity between two images.

Tables 1-6 show the results for cameraman and Lena test images (shown in figure 2) obtained through transmitting the whole approximation coefficients. The results are also plotted in figures 3-5. As seen in these figures, our first purposed method has an almost linear one-to-one correlation with SSIM.

Figure 6 depicts the results of embedding the histogram of approximation coefficients instead of the coefficients themselves for a channel introducing JPEG loss to the image. As seen in the figure, computed MSE has still good correlation with SSIM. This method, however, fails to give satisfactory results for losses which do not change the histogram considerably, as mentioned before.

(a) Test images

(b) Their watermarked version

Figure 2- The standard images on which simulations have been done and their watermarked version.

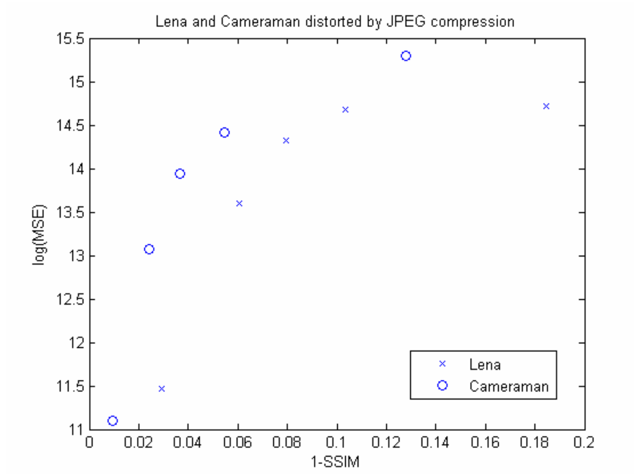

Figure 3- Simulation results for JPEG channel model obtained through transmitting the whole approximation coefficients

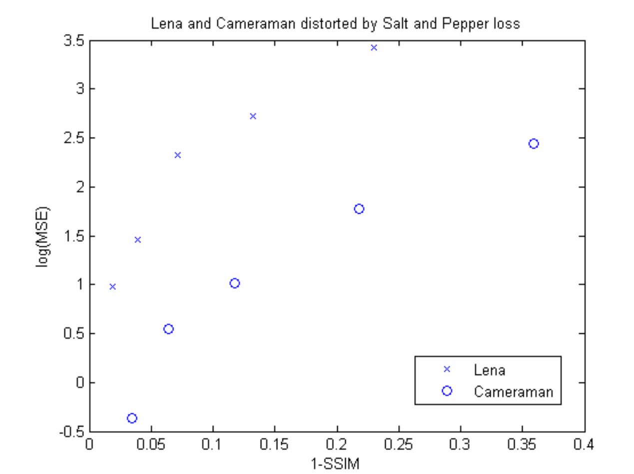

Figure 4- Simulation results for salt and pepper loss obtained through transmitting the whole approximation coefficients

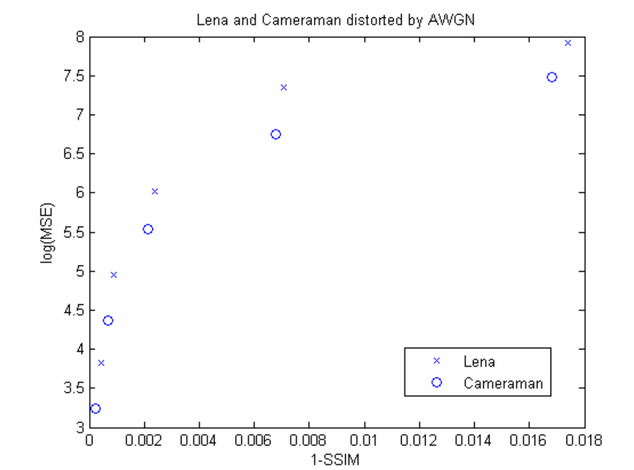

Figure 5- Simulation results for additive white Gaussian noise obtained through transmitting the whole approximation coefficients

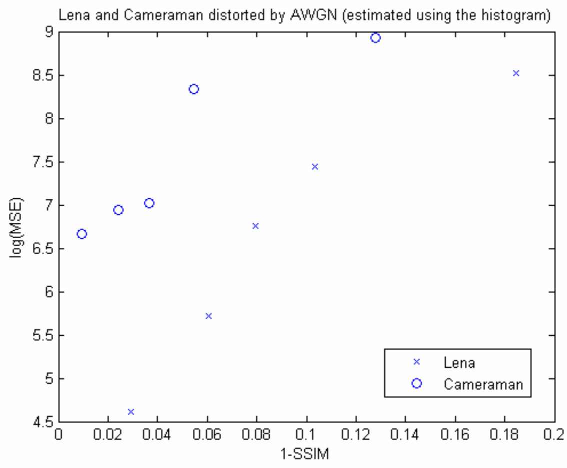

Figure 6- Simulation results for JPEG channel model obtained through embedding the histogram of approximation coefficients instead of the coefficients themselves.

Loss probability	ln(MSE)	PSNR	1-SSIM
2%	3.4294	33.236	0.23026
1%	2.7210	36.313	0.13234
0.5%	2.3251	38.032	0.071452
0.25%	1.4571	41.802	0.039062
0.125%	0.98104	43.870	0.018774

Table 1- Simulation results for Lena test image and salt and pepper loss

Loss probablity	ln(MSE)	PSNR	1-SSIM
2%	2.4364	37.549	0.35857
1%	1.7787	40.405	0.21831
0.5%	1.0147	43.723	0.11780
0.25%	0.54490	45.764	0.06423
0.125%	-0.36092	49.698	0.034553

Table 2- Simulation results for cameraman test image and salt and pepper loss

Compression Quality	ln(MSE)	PSNR	1-SSIM
10/100	14.721	-15.806	0.18456
30/100	14.671	-15.587	0.10355
50/100	14.317	-14.048	0.079536
70/100	13.603	-10.949	0.060449
90/100	11.466	-1.6674	0.029317

Table 3- Simulation results for Lena test image and JPEG compression

Compression Quality	ln(MSE)	PSNR	1-SSIM
10/100	15.295	-18.296	0.12803
30/100	14.417	-14.484	0.054443
50/100	13.941	-12.417	0.036651
70/100	13.075	-8.6541	0.024301
90/100	11.098	-0.071320	0.0096398

Table 4- Simulation results for cameraman test image and JPEG compression

SNR[dB]	ln(MSE)	PSNR	1-SSIM
20	3.8280	31.505	0.00043897
15	4.9481	26.641	0.00090868
10	6.0172	21.998	0.0023942
5	7.3559	16.184	0.0070867
1	7.9142	13.759	0.017400

Table 5- Simulation results for Lena test image and AWGN

SNR[dB]	ln(MSE)	PSNR	1-SSIM
20	3.2471	34.028	0.00021394
15	4.3718	29.144	0.00068280
10	5.5305	24.112	0.0021371
5	6.7508	18.812	0.0067970
1	7.4858	15.620	0.016804

Table 6- Simulation results for cameraman test image and AWGN

	Lena	Cameraman
SSIM	0.999894	0.999998

Table 7- Computed SSIM between the original and watermarked test images.

V. Conclusion and Future Work

We demonstrated a method for estimating image transmission quality over error prone channels and showed the estimated quality has a high correlation with structural similarity index for various images transferred over channels introducing different kinds of loss. The proposed methods are based on extraction of the low frequency content of the image or its histogram at the sender, which is then sent to the receiver by means of a wavelet-based robust watermarking method. The transmission quality is quantified in terms of MSE between the original approximation coefficients or their histogram received as a watermark and those computed at the receiver.

Finally it is worth mentioning that this method can simultaneously be used for error correction or concealment as well. Having the original low frequency content of the image sent to the receiver, it is possible to make use of this side information to correct or conceal at least some of the transmission errors.

REFERENCES

[1] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, Apr. 2004, pp. 600-612,

[2] R. Dosselmann and X. D. Yang, "Existing and Emerging Image Quality Metrics plastics," in *Proc. IEEE Canadian Conference on Electrical and Computer Engineering,* May 2005, pp. 1906-1913

[3] M.Carnec, P. Le Callet and D. Barba, "An Image Quality Assessment Method Based on Perception of Structural Information," in *Proc. Int. Conf. Image Processing*, Sept. 2005, vol. 2, pp. III-185-8

[4] T. M. Kusuma, H. –J. Zepernick, and M. Caldera, "On the Development of a Reduced Reference Perceptual Image Quality Metric," in *Proc. Systems Communications*, Aug. 2005, pp. 178 - 184

[5] M. Carnec, P. Le Callet and D. Barba, "Visual Features for Image Quality Assessment with Reduced Reference," in *Proc. Int. Conf. Image Processing*, Sept. 2005, vol. 1, pp. I-421-4

[6] I. P. Gunawan and M. Ghanbari, "Image Quality Assessment Based on Harmonic Gain/Loss Information," In *Proc. Int. Conf. Image Processing*, Sept. 2005, vol. 1, pp. I-429-32

[7] G. Zhai, W. Zhang, X. Wang, and Y. Xu, "Image Quality Assessment Metrics Based on Multi-scale Edge Presentation," In *Proc. IEEE Workshop Signal Processing Sysmtes Design and Implementation*, Nov. 2005, pp.331-336

[8] M. Carnec, P. Le Callet and D. Barba, "Full Reference and Reduced Reference Metrics for Image Quality Assessment," in *Proc. Seventh International Symposium Signal Processing and Its Applications*, Jul. 2003, vol. 1, pp. 477-480

[9] T. M. Kusuma and H. -J. Zepernick, "A Reduced-Reference Perceptual Quality Metric for In-Service Image Quality Assessment," in *Proc. Joint First Workshop on Mobile Future and Symposium on Trends in Communications*, SympoTIC, Oct. 2003, pp. 71-74

[10] M. Sendashonga and F. Labeau, "Low Complexity Image Quality Assessment Using Frequency Domain Transforms," in Proc. *Int. Conf. Image Processing*, Oct. 2006, pp. 385-388

[11] Z. Wang, G. Wu, H. R. Sheikh, E. P. Simoncelli, E. -H. Yang, and A. C. Bovik, "Quality Aware Images," *IEEE Trans. Image Process.,* vol. 15, no. 6, Jun. 2006, pp. 1680 – 1689

[12] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale Structural Similarity for Image Quality Assessment," in *Proc. Thirty-Seventh Asilomar Conf. Signals, Systems and Computers*, Nov. 2003, vol. 2. pp. 1398-1402

[13] C. B. Adsumilli, M. C. Q. Farias, S. K. Mitra, and M. Carli, "A Robust Error Concealment Technique Using Data Hiding for Image and Video Transmission Over Lossy Channels," *IEEE Trans. Circuits Syst. Video Technol.* vol. 15, no. 11, Nov. 2005, pp. 1394-1406

[14] T. Brandão, and P. Queluz, "Towards Objective Metrics for Blind Assessment of Images Quality," in *Proc. Int. Conf. Image Processing*, , Oct. 2006, pp. 2933-2936

[15] M. Holliman and M.M. Yeung, "Watermarking for Automatic Quality Monitoring," in Proc. *SPIE Security and Watermarking of Multimedia Contents*, vol. 4675, pp. 458-469

[16] P. Campisi, M. Carli, G. Giunta, and A. Neri, "Blind Quality Assessment System for Multimedia Communications Using Tracing Watermarking," *IEEE Trans. Signal Processing*, vol. 51, no. 4, Apr. 2003, pp. 996-1002