Copyright © IEEE.
Citation for the published paper:

This material is posted here with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of BTH's products or services. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by sending a blank email message to pubs-permissions@ieee.org.

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

Quality Assessment of an Adaptive Filter for Artifact Reduction in Mobile Video Sequences

Ulrich Engelke*, Andreas Rossholm[†]*, Hans-Jürgen Zepernick*, and Benny Lövsström*

*Blekinge Institute of Technology

PO Box 520, SE-372 25 Ronneby, Sweden

E-mail: {ulrich.engelke, hans-jurgen.zepernick, benny.lovsstrom}@bth.se

[†]Ericsson Mobile Platforms AB

Nya Vattentorget, SE-221 83 Lund, Sweden

E-Mail: andreas.rossholm@ericsson.com

Abstract—In this paper, we examine an adaptive deblocking deringing filter for mobile video sequences in H.263 format. The considered filter has been designed with reference to the constraints of computational complexity and working memory of mobile terminals. The post filter suggested by the International Telecommunications Union (ITU) in Recommendation H.263 App. III is also included as a reference. Given that fidelity metrics such as the peak signal-to-noise ratio (PSNR) do not necessarily correlate well with video quality as experienced by the user, we consider in this paper objective quality metrics that can incorporate knowledge about the user's perception into the quality assessment. Guidelines for choosing filter parameters in relation to user-perceived video quality are obtained from the numerical results.

I. INTRODUCTION

Modern mobile multimedia terminals differ substantially from the devices used in previous generations of mobile radio systems, which were mainly designed for speech services. Especially, the tremendous advances in integrated circuit technologies potentially allow for the implementation of powerful processing algorithms for encoding and decoding of video sequences on a compact mobile terminal. In view of constraints typically associated with mobile terminals such as hardware limitations, power consumption, and scarce bandwidth resources, the design of computationally efficient video compression algorithms becomes a challenging task.

The most used video codecs, such as those based on the H.263 format recommended by the International Telecommunications Union (ITU) and the Motion Pictures Expert Group (MPEG) standard MPEG-4, provide efficient data compression as needed by mobile video applications. However, a common shortcoming of these codecs is the introduction of blocking and ringing artifacts. These impairments to the original video sequence are due to quantization of the discrete cosine transform (DCT) coefficients and motion estimation associated with block-based codecs. In more recent codecs such as H.264 this problem is reduced by an in-loop deblocking filter. However, for other codecs a post filter is needed to reduce blocking and ringing effects. In order to support a justified selection of favorable filters and filter settings, quality assessment of the filters becomes an interesting objective.

In this paper, we examine the adaptive deblocking deringing

filter presented in [1], which has been designed considering the constraints of computational complexity and working memory of mobile terminals. As in [1], the post filter recommended in H.263 App. III [2] is also included as a reference.

Given that fidelity metrics such as the peak signal-to-noise ratio (PSNR) do not necessarily correlate well with video quality as experienced by the user, we consider in this paper objective quality metrics that can incorporate knowledge about the user's perception into the quality assessment. In particular, metrics that are based on structural video properties such as features or artifacts are beneficial as these metrics relate to the characteristics of the human visual system [3].

This paper is organized as follows. In Section II, the main principles behind the examined adaptive filter for artifact reduction due to video compression algorithms are presented. The concepts of the perceptual-based video quality assessment approach are described in Section III. The considered metrics used to support the quality assessment of the examined adaptive filter are also provided. Numerical examples for a variety of filter settings are reported and discussed in Section IV. Finally, conclusions are drawn in Section V.

II. ADAPTIVE FILTER FOR ARTIFACT REDUCTION

The adaptive deblocking deringing filter that is examined in this paper has been presented in [1]. This adaptive filter operates as a post-processing step after a hybrid differential pulse code modulation (DPCM) transform codec that uses 8×8 blocks for spatial decorrelation as with H.263 and MPEG-4. The filter is adapted on the level of compression where higher compression results in increased filtering. It can be adjusted in two ways; the amount or level of filtering of the frame being processed and the filter strength. The level of filtering refers to the amount of pixels of each block that are being processed. In this paper, we apply filtering to the luminance data in three levels as follows (see also Fig. 1) while chrominance data is not considered in this work:

- L1: Only the first tier of border pixels in every 8×8 block,
- L2: The first and second tier of pixels in every 8×8 block,
- L3: The entire 8×8 pixels in every block are filtered.

The strength of the filter refers to how much low-pass filtering a certain compression level results in. The compres-

Fig. 1. Filter levels with reference to 8×8 blocks of pixels.

sion level is related to the quantization parameter (QP). The different filter strengths are achieved by adding an offset to the QP value at the input of the filter weight generator. In the sequel, we consider the following cases:

- S0: Deblocking filtering is not performed,
- S1: Nominal strength,
- S2: An offset of 4 is added,
- S3: An offset of 6 is added.

III. PERCEPTUAL-BASED VIDEO QUALITY ASSESSMENT

For comparison purposes, we take the fidelity metric PSNR into account and the measure of structural information (SSIM) proposed in [3]. Then, we briefly review the ideas behind the hybrid image quality metric (HIQM) that has been proposed in [4] and then modify this approach to suit both the particulars of quality assessment of mobile video sequences and filter classification. This will result in the metric referred to as normalized HIQM (NHIQM) that we propose for use in perceptual-based video quality assessment.

A. Peak Signal-to-Noise Ratio

Video fidelity is an indication about the similarity between a video frame and a distorted video frame and measures pixel-by-pixel closeness between those pairs. The most commonly used fidelity metric is PSNR, which shall here be used with reference to video frames. Let us first calculate the average of the square of the errors or pixel differences between an original video frame and the related filtered video frame. This is called the mean squared error (MSE) and is defined as

$$MSE = \frac{1}{UV} \sum_u \sum_v [I_o(u, v) - I_f(u, v)]^2 \quad (1)$$

where $I_o(u, v)$ denotes the intensity value at pixel location (u, v) in the original video frame, $I_f(u, v)$ denotes the intensity value at pixel location (u, v) in the filtered video frame, U is the number of rows in a video frame, and V is the number of columns in a video frame. Errors are computed on the luminance signal only, so the pixel values $I(u, v)$ range between 0 (black) and 255 (white). Given the MSE value $MSE(n)$ for the n^{th} pair of original and filtered video frame, then the related $PSNR(n)$ value is calculated as

$$PSNR(n) = 10 \cdot \log \frac{m^2}{MSE(n)} \quad (2)$$

and the PSNR for an entire video sequence of length N can be written as

$$PSNR = \frac{1}{N} \sum_{n=1}^N PSNR(n) \quad (3)$$

where m denotes the maximum intensity value. For the luminance component of a video sequence with 8 bits per pixel, m is 255. It is also noted that PSNR often does not correlate well with what is actually perceived by the human eye [5]. In other words, a high PSNR value indicating high fidelity may actually be perceived as low in quality by viewers.

B. Measure of Structural Similarity

This metric is based on the degradation of structural information in the viewing area. Its outcome quantifies the structural similarity between the reference and the distorted image, which in our case relates to an original and a filtered video frame, as [3]

$$SSIM(n) = \frac{[2\mu_x(n)\mu_y(n) + C_1][2\sigma_{xy}(n) + C_2]}{[\mu_x^2(n) + \mu_y^2(n) + C_1][\sigma_x^2(n) + \sigma_y^2(n) + C_2]} \quad (4)$$

where $\mu_x(n)$, $\mu_y(n)$ and $\sigma_x(n)$, $\sigma_y(n)$ denote the mean intensity and contrast of the n^{th} original video frame x and filtered video frame y , respectively. The constants C_1 and C_2 are used to avoid instabilities in the structural similarity comparison that may occur for certain mean intensity and contrast combinations such as

$$\mu_x^2(n) + \mu_y^2(n) = 0 \quad \text{or} \quad \sigma_x^2(n) + \sigma_y^2(n) = 0 \quad (5)$$

Similar as with PSNR, the SSIM value for an entire video sequence of length N may be calculated as

$$SSIM = \frac{1}{N} \sum_{n=1}^N SSIM(n) \quad (6)$$

C. Hybrid Image Quality Metric

This metric uses standard feature extraction algorithms to quantify artifacts [6]–[9]. A total of five features relating to blocking f_1^* , blur f_2^* , edge-based activity f_3^* , gradient-based activity f_4^* , and intensity masking f_5^* , respectively, are extracted and combined using the corresponding relevance weights. In particular, HIQM is formulated as [4]

$$HIQM = \sum_{i=1}^5 w_i f_i^* \quad (7)$$

where w_i denotes the relevance weights associated with the feature f_i^* . HIQM has been shown to correlate well with the quality as perceived by human observers with the Pearson correlation coefficient [10] reaching nearly 90%. The interested reader is referred to [11] for more details on the development, characteristics, and applications of HIQM. Although this metric can be used for in-service quality assessment as it stands, application to solve classification problems may be difficult as several features with different range of feature values are combined.

TABLE I

QUALITY ASSESSMENT OF VIDEO SEQUENCES FOR DIFFERENT ADAPTIVE FILTER SETTINGS AND H.263 APP. III DEBLOCKING FILTER

		$QP^{*)}$	Bit rate (kb/s)	No Deblocking S0	Deblocking Filter									H.263 App. III
					L1			L2			L3			Deblocking Filter
					S1	S2	S3	S1	S2	S3	S1	S2	S3	
PSNR (dB)	Cart	16	96	29.00	29.08	29.08	29.08	29.10	29.10	29.08	29.12	29.11	29.10	29.25
		29	48	26.16	26.25	26.24	26.24	26.27	26.25	26.24	26.28	26.25	26.23	26.42
	Foreman	10	96	32.04	32.10	32.08	32.06	32.10	32.02	31.99	32.10	32.02	31.98	32.07
		16	48	29.55	29.60	29.58	29.57	29.59	29.54	29.51	29.60	29.53	29.50	29.64
	Mobil	19	96	24.50	24.50	24.49	24.49	24.49	24.47	24.46	24.49	24.46	24.45	24.53
		27	48	22.77	22.76	22.75	22.74	22.73	22.69	22.68	22.70	22.65	22.64	22.80
PSNR (MOS)	Cart	16	96	55.00	55.17	55.18	55.18	55.24	55.22	55.20	55.27	55.25	55.22	55.57
		29	48	49.14	49.31	49.31	49.30	49.35	49.31	49.29	49.38	49.32	49.29	49.66
	Foreman	10	96	62.00	62.14	62.09	62.05	62.13	61.96	61.86	62.13	61.96	61.86	62.06
		16	48	56.14	56.25	56.22	56.20	56.24	56.12	56.06	56.26	56.12	56.05	56.36
	Mobil	19	96	45.92	45.93	45.91	45.90	45.92	45.87	45.85	45.90	45.86	45.83	45.98
		27	48	42.88	42.87	42.84	42.83	42.81	42.75	42.73	42.76	42.68	42.65	42.92
SSIM (MOS)	Cart	16	96	59.49	60.15	60.23	60.24	60.42	60.46	60.43	60.57	60.62	60.59	60.74
		29	48	51.84	52.58	52.59	52.58	52.79	52.72	52.68	52.90	52.80	52.74	52.93
	Foreman	10	96	65.48	65.95	65.96	65.93	66.11	65.99	65.90	66.19	66.08	65.99	66.40
		16	48	59.99	60.50	60.53	60.51	60.67	60.59	60.52	60.78	60.68	60.60	61.09
	Mobil	19	96	54.89	54.94	54.90	54.87	54.91	54.78	54.71	54.87	54.72	54.64	54.63
		27	48	47.57	47.47	47.37	47.33	47.18	46.93	46.84	46.84	46.53	46.41	46.99
NHIQM (MOS)	Cart	16	96	51.82	61.91	66.71	68.14	61.58	65.96	67.51	58.33	61.64	63.02	72.80
		29	48	41.28	58.69	62.81	64.41	59.03	62.30	63.85	50.36	52.81	53.99	62.02
	Foreman	10	96	63.14	74.91	80.73	82.64	73.94	79.23	81.05	70.81	75.16	76.94	80.96
		16	48	50.31	61.72	68.25	70.57	61.24	68.03	70.41	57.61	63.57	65.55	70.96
	Mobil	19	96	66.68	71.89	74.23	75.06	71.50	74.24	75.30	70.14	72.73	73.76	94.45
		27	48	56.13	65.94	69.32	70.37	66.39	69.98	71.31	62.22	65.29	66.63	88.78

*) QP : Quantization Parameter

D. Extreme Value Normalized Hybrid Image Quality Metric

In view of the intended assessment of the impact of different deblocking deringing filter settings on video quality, it is more appropriate to apply a feature value normalization such that the range of values is the same for each of the features. In our context, we normalize feature values f_i^* as follows [12]:

$$f_i = \frac{f_i^* - \min_{j=1, \dots, J} (f_{i,j}^*)}{c_i}, \quad i = 1, \dots, 5 \quad (8)$$

where the feature values $f_{i,j}^*$, $j = 1, \dots, J$ are taken from a training set $\mathcal{J}$ of size J and

$$c_i = \max_{j=1, \dots, J} (f_{i,j}^*) - \min_{j=1, \dots, J} (f_{i,j}^*) \quad (9)$$

The NHIQM value, $NHIQM(n)$, of the n^{th} frame in the video sequence is then obtained as a relevance weighted combination of the five normalized feature values $f_i(n)$, $i = 1, \dots, 5$ for that frame, that is

$$NHIQM(n) = \sum_{i=1}^5 w_i f_i(n) \quad (10)$$

where $w_i \in [0, 1]$, $i = 1, \dots, 5$ are the relevance weights. The similarity between an original and the related filtered video frame may then be measured by the absolute difference value

$$\Delta_{NHIQM}(n) = |NHIQM_o(n) - NHIQM_f(n)| \quad (11)$$

where $NHIQM_o(n)$ and $NHIQM_f(n)$ denote the NHIQM value of the n^{th} original and filtered video frame. In order to provide a quality assessment of an entire video sequence, an average absolute difference value can be calculated as

$$\Delta_{NHIQM} = \frac{1}{N} \sum_{n=1}^N \Delta_{NHIQM}(n) \quad (12)$$

where integer $n \in [1, N]$ denotes the discrete index of the considered frames in the video sequence of length N .

E. Prediction of Subjective Quality

The derivation of prediction functions for subjective quality is usually based on subjective ratings that have been obtained from experiments. These subjective ratings from a group of non-experienced viewers are typically averaged to produce a mean opinion score (MOS) [13] of a particular image or in our case translates to a video frame. An exponential prediction

function has been suggested in [11], to translate difference HIQM values to predicted mean opinion scores (MOS). The rationale behind this choice is the fact that quality metrics are considered here that relate to degradation of structural information and video frame distortion. Highly impaired video frames would produce low MOS while those frames with little distortions would result in high MOS. In addition, an exponential prediction function accounts for saturation mechanisms in the human visual system for severely distorted viewing experiences. In the context of this paper, we therefore adopt this approach and use the following prediction function for mapping the considered quality metrics (QM) to predicted MOS:

$$MOS_{QM} = a e^{-b QM} \quad (13)$$

where a and b are parameters that need to be obtained from curve fitting the relationship between MOS values from subjective tests to the corresponding values of the quality metric $QM \in \{PSNR, SSIM, \Delta_{NHIQM}\}$.

IV. NUMERICAL RESULTS

The particulars of the scenarios that are used here for assessing the quality of the considered deblocking deringing filter for mobile video sequences are as follows. Three H.263 video sequences with fixed quantization are examined, namely, 'Cart', 'Foreman', and 'Mobile' each presented at the two different bit rates of 48 kb/s and 96 kb/s. The video sequences comprised of 150 frames and were given in 176×144 pixels quarter common intermediate format (QCIF) at 15 frames-per-second (fps).

A. Prediction Functions

The quality assessment for these video sequences has been performed with reference to PSNR, SSIM, and NHIQM. As far as NHIQM is concerned, the relevance weights that are required in (10) were deduced from our data base of subjective tests for images. These tests were conducted at the Department of Signal Processing at the Blekinge Institute of Technology and have been reported earlier in detail, e.g. [11], [14]. Given that we process the video sequence on a frame-by-frame basis, the MOS values from the aforementioned subjective test can be used to relate the extreme value normalized features to user-perceived quality. Using curve fitting procedures, the relevance weights were obtained as

$$\begin{aligned} w_1 &= 0.77 \\ w_2 &= 0.35 \\ w_3 &= 0.61 \\ w_4 &= 0.16 \\ w_5 &= 0.35 \end{aligned}$$

It can be seen from these values that the blocking artifact receives the highest relevance weighting of $w_1 = 0.77$, which clearly supports the effort of designing efficient deblocking filters. It should also be mentioned that one would theoretically expect less blocking with more filtering at the expense of

Fig. 2. Frame samples of video sequence 'Cart' (left column) of bit rate 48 kb/s and their zoomed versions (right column) for filter strength S1 and different filter levels. Top to bottom: No filtering, L1, L2, L3, H.263 App. III.

increased blur. In view of the relatively low value of the relevance weight for blur of $w_2 = 0.35$, this increase may not be as perceptually significant as compared to blocking. Finally, it is noted that the obtained relevance weights for the extreme value normalized features are equal to the weights that were reported in [11], [14] for the non-normalized features. This is understood to be attributed to the fact that the imposed normalization does not effect the statistical relationship between video frame impairments and subjective scores. In other words, the correlation between an individual feature and the MOS values from subjective test is not affected by this

normalization.

The exponential prediction functions that relate the three considered metrics to predicted MOS are obtained from curve fitting as

$$\begin{aligned} MOS_{PSNR} &= 17.36 \cdot e^{+0.039 PSNR} \\ MOS_{SSIM} &= 14.93 \cdot e^{+1.662 SSIM} \\ MOS_{NHIQM} &= 95.20 \cdot e^{-2.782 \Delta_{NHIQM}} \end{aligned}$$

The accuracy of the predicted MOS with respect to the actual MOS in our data base from the subjective tests can be quantified by the Pearson correlation coefficient. The related correlation coefficients r of these predictions are obtained as

$$\begin{aligned} r_{PSNR} &= 0.778 \\ r_{SSIM} &= 0.600 \\ r_{NHIQM} &= 0.894 \end{aligned}$$

Clearly, NHIQM achieves by far the highest correlation to the MOS from the subjective tests followed by PSNR and SSIM. This finding is also interesting in view of both PSNR and SSIM being full-reference metrics that rely on the availability of the original video frames while NHIQM as a reduced-reference metric only requires knowledge about features.

B. Quality of Video Sequences

Table I shows the results from the quality assessment of the examined three video sequences for different adaptive filter settings and each of the considered metrics. The numerical values for each filter level and strength setting of the considered deblocking filter are given for PSNR in dB as well as in predicted MOS and for both SSIM and NHIQM in predicted MOS. The results obtained for the H.263 App. III deblocking filter are also presented for comparison.

It can be seen from the results presented in Table I that PSNR does not support a distinct differentiation among the analyzed filter settings. This applies to both when PSNR is measured in dB and when it is mapped to predicted MOS. The same behavior can be observed for SSIM, which provides only little differentiation between the filter settings for a given QP value and bit rate.

On the other hand, NHIQM is in fact able to distinguish among different quality levels for the different adaptive filter settings. It is observed that for a given filter level, the quality in terms of predicted MOS is increased by increasing the filter strength. As far as the filter level is concerned, it appears to be sufficient to use level L1 while an increase to level L3 would actually reduce the quality. This is because the higher filter levels do not further decrease the blocking but increase the blur.

C. Quality of Video Frame Samples

Figure 2 shows samples of the 72nd frame of video sequence ‘Cart’ with quantization parameter $QP = 29$, bit rate of 48 kb/s, filter strength S1 and different filter levels. The zoomed version is also given for each of the frame samples. It should be mentioned that zoomed frames have been produced

using the pixel replication technique which is a special case of nearest neighbor interpolation [15]. Distortions due to this zoom operation have not been observed for the considered frame samples. In addition, it is noted that video frame samples are used as a means of visualizing while the conclusions given in the sequel are more pronounced when comparing the actual streaming videos. It can be seen that an increase in filter level from no filtering to level L1 gives the most perceptual improvement. This is especially feasible in the areas around the left back wheel of the cart. The blocking artifact can be clearly seen in the non-filtered sample in the top row of the figure. Imposing the filter level L1, the samples in the second row are produced much smoother with the blocking largely reduced. An additional increase in the filter level does not appear to improve the perceptual quality for those samples. In fact, the filter level L3 starts to introduce some blur to the sample as can be seen from the third row of the figure. Finally, it can be observed that with the H.263 App. III deblocking filter some degree of blocking still remains but overall provides comparable performance as the examined adaptive filter suggested in [1].

D. Feature Values, Predicted MOS, and Filter Settings

Similar conclusions can be drawn from Fig. 3, which shows the progression of the actual feature values over the four filtering strengths. The most significant reduction in the blocking feature f_1 is obtained when changing from no deblocking filtering performed S0 to nominal filter strength S1. Further increase in filter strength produces only minor blocking decrease. Although the blur feature f_2 seems to increase with stronger filtering as expected, the introduced impairment is only minor. As far as an increase of filter level is concerned, it can be observed from the figure (e.g. Fig. 3a-c) that although the blocking reduces for all three levels with increase of filter strength, the absolute feature value of blocking increases with increase of the filter level from a value below 0.3 in Fig. 3a to a value above 0.3 in Fig. 3c. This effect is due to the algorithm deployed here for the extraction of the blocking artifact [6]. This algorithm not only accounts for blocking but also indirectly considers some degree of blur. As such, an increase of blur with the increase of filter level in turn may cause the feature value f_1 to settle at higher values. As far as the blur feature f_2 itself is concerned, it also increases with an increase of filter level (e.g. Fig. 3a-c).

Figure 4 visualizes the comprehensive results of Table I for the three video sequences in terms of predicted MOS for different bit rates, filter levels, and filter strengths when using NHIQM for quality assessment. It can be seen that the video quality increase with the bit rate from 48 kb/s to 96 kb/s and again with the increase of filter strength from S0 to S4. It can also be observed from these results that increasing the filter level beyond L2 would not improve quality but may reduce it for level L3 due to the introduction of blur with this level. Also, as can be seen from Fig. 4a and Fig. 4c, the increase in quality between the different filter strengths can be higher for the lower bit rate of 48 kb/s.

Fig. 3. Extreme value normalized feature values for the considered video sequences of bit rate 48 kb/s for the three different filter levels L1, L2, and L3, and the four different filter strength S_0 , S_1 , S_2 , and S_3 : (a) ‘Cart’ for filter level L1, (b) ‘Cart’ for filter level L2, (c) ‘Cart’ for filter level L3; (d) ‘Foreman’ for filter level L1, (e) ‘Foreman’ for filter level L2, (f) ‘Foreman’ for filter level L3; (g) ‘Mobile’ for filter level L1, (h) ‘Mobile’ for filter level L2, (i) ‘Mobile’ for filter level L3 (Considered features: Blocking f_1 , blur f_2 , edge-based activity f_3 , gradient-based activity f_4 , intensity masking f_5).

E. Filter Design Guidelines

In the sequel, we summarize some of the main findings as supported by the results shown in Table I and the Figs. 2-4:

- PSNR, the conventional fidelity metric, is not able to differentiate among the artifact reductions related to the different filter settings as only minor changes in their values are observed (see Table I).
- SSIM varies also only very little with the different filter settings for a given video sequences and hence does not support selection of favorable filter parameters.
- NHIQM, which is a perceptual-based video quality metric, clearly distinguishes between the video quality related

to the considered filter settings (see Table I and Fig. 1). It also provides insights into structural information with respect to the individual features.

- The examined filter can largely reduce compression induced blocking with only minor increase of other artifacts (see Fig. 2 and Fig. 3).
- The different filter levels appear not to have a large impact on the quality of the considered video sequences. In view of mobile video applications, the examined adaptive filter may only deploy the least complex level L1 (see Fig. 3).
- The NHIQM-based approach can be used to select parameters of the considered adaptive filter to suit mobile video applications and to support given quality constraints.

V. CONCLUSIONS

In this paper, we have examined an adaptive deblocking deringing filter for mobile video sequences in H.263 format. The considered filter supports different parameter settings in terms of filter levels and filter strengths depending on complexity and memory constraints. In order to guide the selection of favorable filter parameters for mobile video applications, a quality assessment has been performed using the conventional fidelity metric PSNR and the perceptual-based metrics SSIM and NHIQM. The quality assessment with NHIQM as proposed in this paper turns out to be suitable to drive the filter design. On the other hand, PSNR and SSIM are not able to give a pronounced differentiation for different filter settings. The quality assessment with NHIQM reveals that the examined deblocking deringing filter can indeed largely reduce compression induced blocking. In view of deployment of the filter in mobile video applications, it may be sufficient to use only the least complex filter level. A set of comprehensive numerical results are also provided which can be used to guide the selection of filter parameters with respect to user-perceived quality.

REFERENCES

- [1] A. Rossholm and K. Andersson, "Adaptive De-blocking De-ringing Filter," IEEE Int. Conf. on Image Proc., Genoa, Italy, Sept. 2005, pp. 1042-5.
- [2] ITU-T Recommendation H.263 Appendix III, "Examples for H.263 Encoder/Decoder Implementations," June 2000.
- [3] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE Trans. on Image Proc., vol. 13, no. 4, pp. 600-612, April 2004.
- [4] T. M. Kusuma and H.-J. Zepernick, "A Reduced-reference Perceptual Quality Metric for In-service Image Quality Assessment," IEEE Symp. on Trends in Commun., Bratislava, Slovakia, Oct. 2003, pp. 71-74.
- [5] H. R. Wu and K. R. Rao, "Digital Video Image Quality and Perceptual Coding," Boca Raton: CRC Press, 2006.
- [6] Z. Wang, H. R. Sheikh, and A. C. Bovik, "No-reference Perceptual Quality Assessment of JPEG Compressed Images," IEEE Int. Conf. on Image Processing, vol. 1, Rochester, USA, Sep. 2002, pp. 477-480.
- [7] P. Marziliano, F. Dufaux, S. Winkler, T. Ebrahimi, "A No-reference Perceptual Blur Metric," IEEE Int. Conf. on Image Processing, vol. 3, Rochester, USA, Sep. 2002, pp. 57-60.
- [8] S. Saha and R. Vemuri, "An Analysis on the Effect of Image Features on Lossy Coding Performance," IEEE Signal Processing Letters, vol. 7, no. 5, pp. 104-107, May 2000.
- [9] A. R. Weeks, "Fundamentals of Electronic Image Processing," SPIE Optical Engineering Press, 1996.
- [10] S. Winkler, "Digital Video Quality - Vision Models and Metrics," Chichester: John Wiley & Sons, 2005.
- [11] T. M. Kusuma, "A Perceptual-based Objective Quality Metric for Wireless Imaging," Ph.D. thesis, Curtin University of Technology, Perth, Australia, 2005.
- [12] J.-R. Ohm, "Multimedia Communication Technology: Representation, Transmission and Identification of Multimedia Signals," Berlin: Springer, 2004.
- [13] "Methodology for the Subjective Assessment of the Quality of Television Pictures," ITU-R, Rec. BT.500-11, 2002.
- [14] U. Engelke, H.-J. Zepernick, and T. M. Kusuma, "Perceptual Evaluation of Motion JPEG2000 Quality over Wireless Channels," IEEE Symp. on Trends in Commun., Bratislava, Slovakia, Jun. 2006, pp. 92-96.
- [15] R. C. Gonzalez and R. E. Woods, "Digital Image Processing," Upper Saddle River: Prentice Hall, 2001.

Fig. 4. Predicted video quality, MOS_{NHIQM} , for different filter levels L1, L2, and L3 applied to the video sequence (a) 'Cart', (b) 'Foreman', (c) 'Mobile'; (Bit rates: 48 kb/s in red tone/foreground and 96 kb/s in blue tone/background; Parameter: Filter strengths S0, S1, S2, and S3).