INVITED PAPER *Special Section on Multimedia QoS Evaluation and Management Technologies*

Subjective Quality Assessment of the H.264/AVC In-Loop De-Blocking Filter

Matthew D. BROTHERTON[†a)], **Damien BAYART**[†], *and* **David S. HANDS**[†], ***Nonmembers***

SUMMARY Next generation codecs, benchmarked by the H.264/AVC standard, are providing substantial compression efficiency for the coding and transmission of video. Coupled with technologies offering larger transmission bandwidths over DSL, wireless and satellite networks, the capability of delivering high quality video services to the home is now a reality. The perceptual quality of the content delivered over communications networks will be crucial in ensuring a first-class customer experience. It is therefore important to assess the advantages and disadvantages of the optional features offered by next generation codecs. This paper describes a subjective assessment that was carried out to investigate the perceptual effects of switching the in loop de-blocking filter within the H.264/ AVC CODEC on or off. Although the filter is believed to substantially improve the perceptual quality of video, it has been suggested that in some cases negative perceptual effects can be produced. The H.264/AVC architecture allows de-blocking to be switched off in cases where there are limited processing resources or it is considered a negative perceptual effect may be introduced. This paper describes a study that examined the perceptual effects of de-blocking by employing a standardised subjective assessment methodology. The Absolute Category Rating (ACR) method was used to capture Difference Mean Opinion Scores (DMOS) for a range of video. Content was selected to span a wide and representative range of coding complexity. This content was then encoded at a variety of bit-rates to represent high, medium and low qualities. Results were used to examine the end-user perception of video quality when the de-blocking filter is switched on or off. The experimental design allowed the overall effects of the de-blocking filter to be examined and additionally the relationship between content and quality on the filter performance. The experiment found that the performance of the de-blocking filter was content-dependent. Results were used to discuss the advantages and disadvantages of in-loop de-blocking and there is an examination of content properties (e.g. spatial and temporal complexity) that influence the performance of de-blocking.
key words: *subjective assessment, video quality, H.264/AVC, de-blocking*

1. Introduction

H.264/AVC [1] is the most recent video coding standard and emerged from a joint project between the ITU-T video coding experts group (VCEG) and the motion picture experts group (MPEG). The new standard has set the benchmark for next generation codecs. Empirical studies have shown H.264/AVC can achieve 50% coding gain over MPEG-2, 47% coding gain over H.263 baseline, and 24% coding gain over H.263 high profile encoders [2]. This improved coding performance has created a path for high quality video services to be delivered over DSL, wireless and satellite networks. The possibility of triple play (voice, data, and video) down a single access line is now a service that can be offered by telecommunications companies. There is scope for broadcasters to reduce the bandwidth requirements per video channel in order to increase the number of channel offerings or support HDTV content. Packaged media providers can also offer an increased quantity of content as higher compression efficiency creates more free space per disk. More importantly, satisfying customer requirements for higher perceptual quality can be achieved. When considering customer requirements for quality it is essential to understand the techniques and features offered by new video coding standards and assess not only their advantages but also their potential disadvantages.

H.264/AVC is known to be considerably more complex to implement than previous coding standards. The H.264/AVC baseline decoder has been shown to be 2.5 times more complex than the H.263 baseline decoder [3]. This complexity introduces a high processing power requirement making it unsuitable for real time implementations on some current digital signal processing (DSP) devices [4]. Profiles are available in baseline, main and extended and are split into two distinct layers, the video coding layer (VCL) and the network adaptation layer (NAL). The VCL is inherently similar to previous video coding standards (MPEG-2 and MPEG-4). Translational block-based motion compensation, transform based residual coding, scalar quantization, adjustable quantization, ziz-zag scanning and run length variable length coding of quantized transform coefficients are all included in baseline, main and extended profiles. The improved coding performance is achieved by some significant changes in the inherent features plus the addition of advanced coding functions. Primarily a much more flexible model for motion compensation supporting the use of multiple reference frames and several block search sizes is implemented. Additionally motion vectors can be specified with 1/8 pixel accuracy. New coding features include integer DCT replacing floating point DCT. This offers reduced blocking, ringing and encoder-decoder mismatches in the reverse transform [2]. The standard also features a more complex context-based arithmetic coding (CABAC) for entropy coding.

H.264/AVC also specifies the use of an improved deblocking filter within the motion compensation loop. The de-blocking filter can be switched on or off depending on coding requirements to reduce blocking artefacts. Blocking is one of the most common artefacts produced by video compression and is a common characteristic of block-based

Manuscript received May 2, 2005.
Manuscript revised July 22, 2005.
[†]The authors are with British Telecom Broadband Applications Research Center, UK.
a) E-mail: matthew.brotherton@bt.com
DOI: 10.1093/ietcom/e89–b.2.273

transform codecs. Two methods of de-blocking are currently used and are classified by their implementation inside or outside the coding loop. Filters within the coding loop are appropriately referred to as 'in loop filters' while filters outside the coding loop are known as 'post filters.' Loop filters have been empirically shown to improve both the objective and subjective quality compared to filters outside the coding loop [5]. In this work both objective and subjective measurements were made. However the objective measurements were based on the PSNR which is known to correlate poorly with the actual perceived quality of video [6]. The subjective measurements have been by 'golden-eye' judgments from expert users which are very far from being representative of the average home user. The typical home user, typically referred to as 'Novice users' in the literature, are known to be less sensitive to distortions, and to be less critical of perceptual quality, than an expert eye [7], [8]. Further, 'golden eye' tests are often performed by individuals who are closely associated with a technology, product or service and are thus open to the change of bias. Nevertheless, it has been suggested that in specific cases using de-blocking can actually introduce degradations in perceptual quality [4]. Unfortunately, this study provides no empirical evidence to indicate when any depression in quality may be introduced by the de-blocking process.

Because the in loop de-blocking filters are applied within the motion compensation loop of the encoder the filter must also be applied at the decoder. This requirement introduces several issues in the coding process and implementation. For generic loop filters the encoder and decoder must use an identical filter. This introduces a limitation for interoperability unless standardised filters are used, like in the H.264/AVC standard. In-loop filters more importantly increase the complexity of the codec. There is a subsequent increase in requirement for bus bandwidth [3].

A recent study [9] examined the perceptual quality of H.264/AVC de-blocking when applied to CIF content, employing a novel subjective method known as the User Feedback Quality Measurement Method (UFQ). This study found variable results for the performance of de-blocking and suggested that for some content users fall into two groups. This bi-modal distribution in subjective ratings is inconsistent with established subjective methods that tend follow normal distribution. The results of this study are of interest and question the requirements for de-blocking when there is need for performance optimisation in H264/AVC CODECS. Predominantly, there has been little published research to investigate the end-user perception, measured by a standardised subjective assessment procedure, of standard resolution video content with an in loop de-blocking filter applied. This paper describes a subjective quality study that was designed to examine the perceptual effects of switching the de-blocking filter on or off within the H.264/AVC codec. A selection of video material spanning a wide range of coding complexity was encoded at various bit-rates to represent high, medium and low quality. The aim of this study was to assess the performance of the de-blocking on a variety of content at a range of quality levels. Two basic research questions were addressed. Firstly, to what extent does de-blocking improve the perceptual quality of H.264 encoded video, as indicated by subjective measurements? Secondly, if there are instances where de-blocking introduces negative perceptual effects, how can the causes be characterised?

2. Method

2.1 Participants

Subjects were recruited from a pool of employees working for British Telecom at Adastral Park, Ipswich, UK. A total of 15 subjects took part in the test, 4 females and 11 males. All subjects worked in the telecommunications industry but had no experience of video coding or issues related to video quality. None of the subjects had taken part in a subjective test for at least 6 months. Subjects were screened for the following:

- Normal (20/20) visual acuity (Snellen Test)
- Normal colour vision (Ishihara Test)
- Sufficient familiarity with the language to comprehend instructions and to provide valid responses

2.2 Apparatus

Each subject completed the test in a sound insulated cabinet. Test scenes were presented to subjects on 21" Sony professional grade monitor (BVM-20F1E). Rating scales were visually presented after each clip on the monitor. Subjects provided quality ratings using a Bluetooth number pad. The Bluetooth number pad was connected to a PC outside the cabinet where the results were stored. This PC was also used to store and play the video content.

2.3 Stimuli

The subjective study consisted of 9 reference scenes chosen to span the range of coding complexity. Each scene was full resolution (720 × 576 pixels) and full frame rate (25 fps). All of the scenes except one were interlaced. Each reference scene was encoded to represent three quality levels, high/medium/low, using the H.264 AVC codec. The experimenter was responsible for selecting the quality levels. Additionally each scene was encoded twice, once with the de-blocking filter switched on and once with the filter switched off. There were a total of 63 scenes presented in the test; 9 reference scenes and 54 test scenes. The 54 test scenes contained each of the 9 scenes encoded at 3 quality levels with the filter switched on or off. To obtain the three different quality levels across the test scenes, bit-rates ranging from 200 kbps to 4000 kbps were used. For the low quality scenes a bit rate of 200 k was used. For medium quality scenes bit rates were either 500 k or 1000 k. For the high quality scenes bit rates between 1500 k and 4000 k were used. The test scenes were stored on a ClipStationPro digital video player

Table 1 Reference scenes (720 × 576 pixels, 25 fps, PAL).

Scene ID #	Characteristics	Duration (seconds)	Interlaced/ Progressive
1#	High Detail, High Motion, Panning	8	**Interlaced**
2#	High Detail, High Motion, Fast Panning, 2 Angles.	8	**Progressive**
3#	High Detail, High Motion, Panning, Zooming	8	**Interlaced**
4#	High Motion, Panning, 2 Camera Angles	8	**Interlaced**
5#	High Detail, High Motion.	8	**Interlaced**
6#	High Detail	8	**Interlaced**
7#	High Detail, Low Motion	8	**Interlaced**
8#	Low Motion	8	**Interlaced**
9#	Low Motion	8	**Interlaced**

in uncompressed 4:2:2 format. Table 1 describes the characteristics of the reference scenes used.

2.4 Design

The experiment was a repeated measures three way design (9 (content) × 2 (filter) × 3 (quality)). Subjects provided data in every condition. The three factors were Content, Filter and Quality. There were nine levels in the Content factor (1#, 2#, 3#, 4#, 5#, 6#, 7#, 8#, and 9#). There were two levels in the Filter factor (ON or OFF). There were three levels in the Quality factor (High, Medium, and Low). The dependent variable in the experiment was the difference mean opinion score (DMOS).

2.5 Procedure

The single stimulus 5 grade quality scale (ACR) was selected as the subjective testing methodology. This single stimulus methodology has been defined by experts working in the field of video quality [10], [11]. The ACR method is characterised by a series of short test sequences presented one at a time. Each test sequence is then rated independently on a discrete five grade category scale. The method specifies that after each presentation the subjects are asked to evaluate the overall quality of the scene shown using the descriptors: Excellent, Good, Fair, Poor, and Bad. The category scale used to select the quality of each scene is shown in Fig. 1. For analysis purposes each of the descriptors was coded with the values 1 (poor) through to 5 (excellent). These numerical codes were included on the voting dialog as they indicated which button the subject was required to press on the Bluetooth number pad. The time pattern for the stimulus presentation is illustrated by Fig. 2. In this study there was no limit on the time available for subjects to vote.

Subjects were seated 6 picture heights away from the monitor which is in line with ITU-R Rec. BT.500-11 [12]. Subjects were given written instructions prior to starting the test. The test session for this assessment lasted approximately 20 minutes, including training. The training session consisted of four practice scenes. The four scenes used within the practice session spanned the quality and content that was used within in the test. The content used in the training session was not replicated in the main assessment. Results from the practice session were collected but not used in the results analysis. Presentation of test scenes randomised across subjects according to a pseudo-Latin square design. The function of this randomisation was to reduce any contextual effects that might be introduced by the order of scene presentation [13]. Reference scenes were included in the test but their presence was not made explicit to the subject.

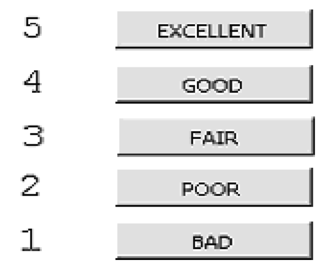

Fig. 1 Voting dialog for ACR methodology.

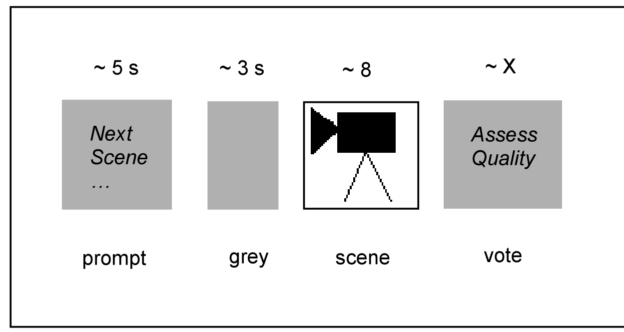

Fig. 2 Time schedule for ACR methodology.

3. Results

For each test scene a DMOS was calculated by subtracting the test scene MOS from the corresponding reference scene MOS. This procedure is know as hidden reference removal (HRR) and was proposed to be used within ACR testing by the Video Quality Experts Group [11]. A previous study using single stimulus continuous quality evaluation utilised HRR to improve the stability of subjective quality assessments [14]. In the case of the ACR methodology, HRR is applied to provide improved stability in the subjective data. This is achieved by calculating subjective quality ratings relative to a benchmark or standard quality provided by the reference sequence. It is worth noting that the proposed advantages of HRR have yet to be established empirically.

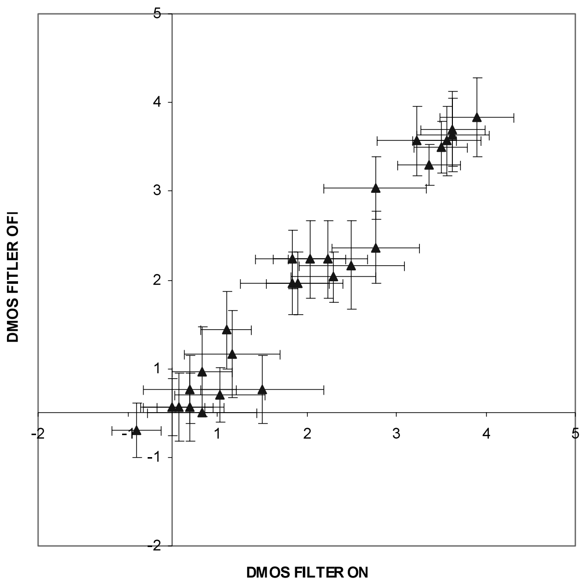

Fig. 3 Comparison of DMOS for the de-blocking filter ON or OFF. Vertical and horizontal bars indicate the 95% confidence intervals for each test scene for the filter on (horizontal bars) or off (vertical bars).

The use of HRR here is in accordance with the experienced judgement of the Video Quality Experts Group. When interpreting DMOS, the reader should be aware that smaller values indicate higher opinions of video quality. Therefore a DMOS of 0 represents no subjective difference between the reference scene and test scene. A DMOS of 100 indicates maximum difference in quality between the reference and test video. In this subjective test, some DMOS were negative due to the fact the highest quality test scene was rated higher than the reference scene. Negative DMOS were marginal and are known to occur at high fidelities in subjective assessment methodologies [14]. As the negative DMOS values were marginal, they were included in the data analysis. Figure 3 shows the DMOS for all test scenes when the de-blocking filter was switched on or off.

Figure 3 illustrates that there is clear clustering of DMOS in three areas. These three areas represent a distinct definition of content encoded to represent high, medium and low bit-rates. The initial distribution of DMOS in Fig. 3 shows that in some test conditions the DMOS is lower (better perceptual quality) when the filter was switched on; for other conditions, DMOS is lower when the filter was switched off. As a quality indicator of the subjective data, the distributions of the 95% confidence intervals were calculated and are shown in Fig. 3. The average confidence interval was ± 0.403 DMOS units.

Figure 4 displays DMOS scores for each individual scene. Initially Fig. 4 reveals there are some clear differences in DMOS for particular clips. Recall that higher DMOS scores represent a lower subjective opinion of quality.

For content encoded at low bit rates, scenes 4#, 7#, and 9# were rated perceptually better quality with the deblocking was switched on. For all other scenes encoded at low bit rates there was no perceptual difference for deblocking switched on or off.

For content encoded at medium quality bit-rates scenes 1#, 5# and 9# were rated with a higher perceptual quality for de-blocking switched on. However, for scenes 2#, 6#, and 8# there was a higher perceptual quality when de-blocking was switched off. For other scenes encoded at medium bitrate there was no perceptual difference indicated by DMOS ratings.

For the majority of scenes encoded at high bit rates there were no differences when de-blocking was switched on or off. However two scenes, 5# and 8#, showed considerably lower perceptual quality when de-blocking was on.

To investigate the results more conclusively a three way repeated measures ANOVA was performed. Table 2 displays the F- and p- values for all of the main effects and interactions in this subjective study. There were main effects for two of the factors, Content and Quality. These effects were expected as content was selected to span a wide range of coding difficulty and quality levels were selected to represent high, medium and low qualities. There was no main effect for the filter factor. There was an expected interaction effect between Content and Quality. There was however no overall interaction effect between Content, Filter and Quality.

Additionally there were was no interaction effect between Filter and Quality. However there was a significant interaction effect between Content and Filter. Figure 5 displays the interaction plot for Content and Filter.

The scenes within Fig. 5 have been listed along the abscissa according to the largest positive difference in DMOS between switching the filter on or off. Scenes listed closest to the ordinate have better perceptual quality when deblocking was switched off. Scenes farthest away from the ordinate are those that have better perceptual quality when the filter was switched on. Ordering scenes this way produces a mid-point in Fig. 5 (Scene 3#) where it appears there is a threshold in which one direction de-blocking has a negative effect and in the other it has a positive effect. Identifying the characteristics of content either side of this threshold would create a useful mechanism to flag when de-blocking is appropriate. This mechanism would require the development of a metric that detects characteristics of content that produce negative perceptual effects when de-blocking is applied.

4. Discussion

The three-way repeated measures ANOVA described in the results section show a significant interaction between scene content and filter. The results found that turning the deblocking filter on produced both positive and negative effects on perceptual quality ratings. These results show that the performance of the H.264/AVC de-blocking filter is content-dependent. Previous research has suggested that it is possible for de-blocking to produce negative Perceptual effects [4].

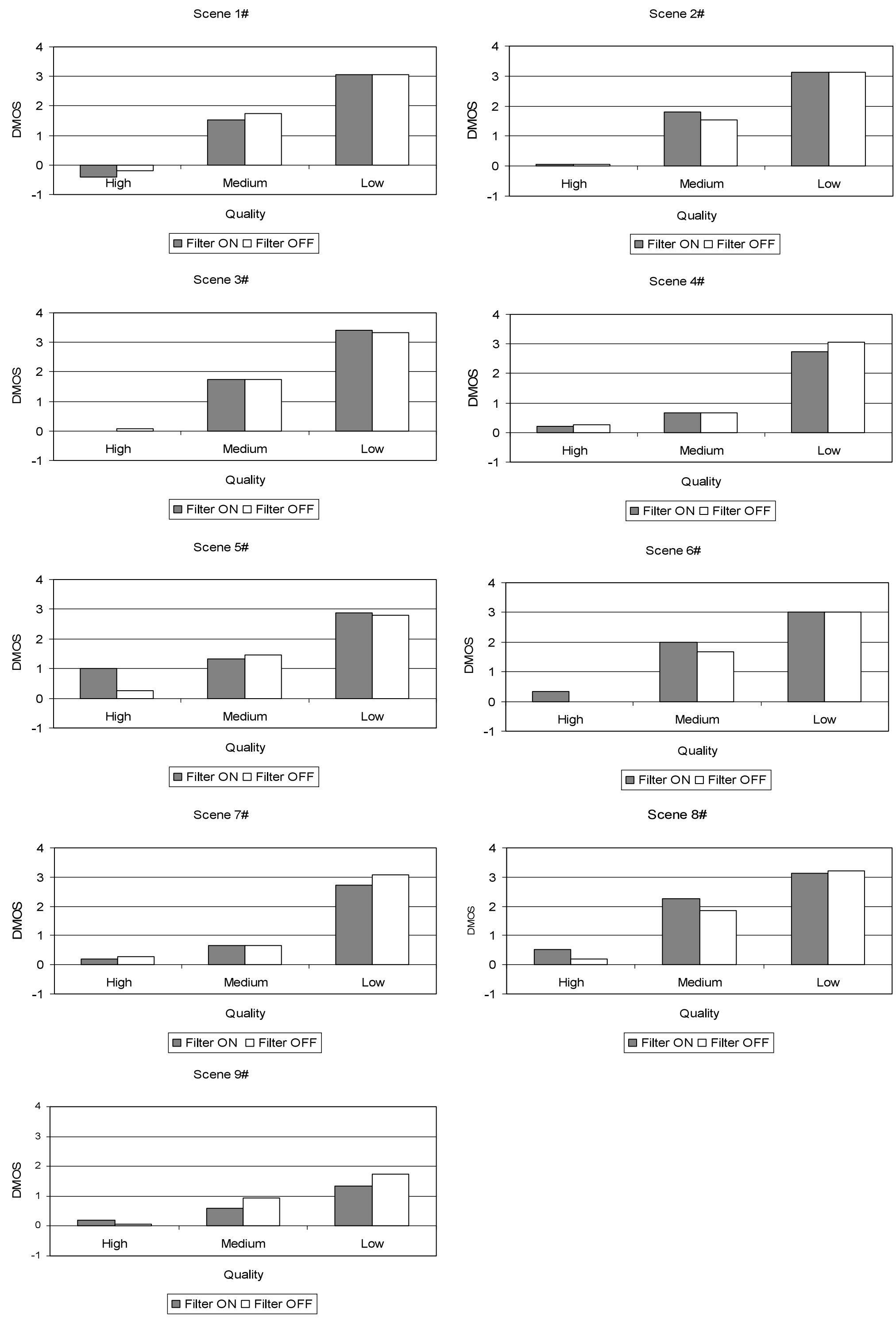

Fig. 4 DMOS for individual scenes at high, medium and low qualities with the de-blocking filter switched on or off.

A recent study [9] suggested that for identical content with de-blocking on or off, subjects fell into two groups. For some subjects there was a preference for de-blocking on, however a second group of subjects indicated a preference for de-blocking off. This bimodal distribution of subjective ratings was suggested for both low and medium motion content. For high motion content there was found to be no clear user preference for applying de-blocking or not. This study employed a non-standardised subjective testing method. This method required to subjects to state a preference when presented with a series of identical tests scenes with variable de-blocking filter settings. It should be noted

Table 2 F- and p- values for main effects and interaction effects.

	df	F-Value	p-Value
Content	8	5.279	<.0001
Filter	1	0.62	ns
Quality	2	446.852	<.0001
Content * Filter	8	2.037	<.05
Content * Quality	16	13.486	<.0001
Filter * Quality	2	2.405	ns
Content * Filter * Quality	16	1.648	ns

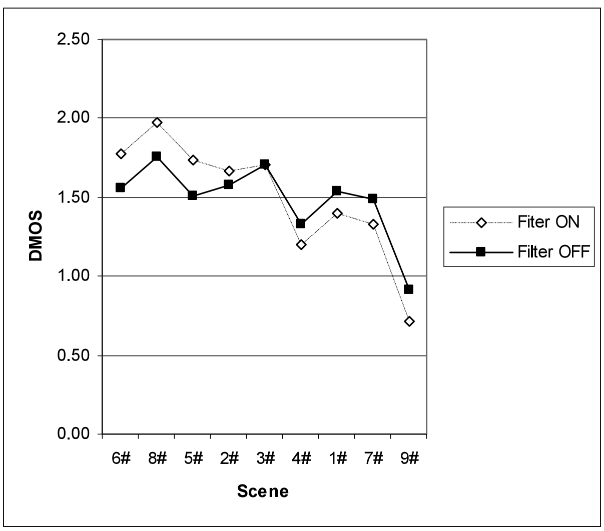

Fig. 5 Interaction plot of Content × Filter for DMOS. DMOS shown have been averaged across all three levels of Quality. Scenes ordered along abscissa according to (Filter ON DMOS—Filter OFF DMOS).

that the unusual distribution of subjective ratings may have been a consequence of the methodology applied. Although the conclusions derived by the authors pose interesting questions to the requirements and suitability of de-blocking the data from this study fails to provide clear evidence on the performance of H.264/AVC de-blocking. The work reported here provides a more extensive evaluation of the perceptual effects of H.264/AVC de-blocking. The present work examined a wider range of content and considered factors beyond solely motion in effecting de-blocking performance. Additionally, the present work used a standardised subjective test method requiring subjects to provide ratings of video quality for all test scenes. Unfortunately no direct comparisons could be made between the results from [9] and this study due to the methodologies used and lack of statistical data presented.

Given the results of this present study, a useful feature of the H.264/AVC codec would be to automatically decide whether de-blocking should be applied. Such a decision could be based on a flag within the header of unprocessed video which is then read by the encoder. Useful deployment of this flag is dependent on whether video characteristics can be identified that improve or worsen perceptual quality when de-blocking is applied. Identifying these video characteristics may provide a set of criteria that can serve as guidelines when deciding to switch de-blocking on or off. The properties of the test scenes used in this study were examined in an attempt to identify characteristics indicative of whether de-blocking is beneficial or detrimental to perceptual quality

For test scenes 2#, 5#, 6# and 8# it was apparent that perceptual quality was worse when de-blocking was applied. Degradations in these scenes were characterized by an over smoothing of detailed regions resulting in a blurring effect and general loss of detail. Figures 6 and 7 illustrate two synchronous frames from scene 8# encoded at 200 k with de-blocking on or off. In Fig. 6 the scene was encoded with de-blocking applied and in Fig. 7 no de-blocking was applied. It is clear to see that the frame in Fig. 6 has considerably less detail than Fig. 7 due to an over smoothing effect. The subjective results of this study would indicate that for some content an over smoothing effect and loss of fine detail is less preferable than blocky images that retain fine detail. However it should be noted that examining single frames in an unrealistic method of evaluating trends in subjective ratings.

As an initial basis of characterising these scenes, the spatial and temporal complexities were calculated for each of the reference scenes used in this subjective study and are outlined in Table 3. Spatial and temporal measures were derived from an objective video model designed by British Telecom [16]. There was no obvious association between complexity and the perceptual effect of de-blocking. Scenes that are negatively affected by the filter share similar characteristics of high spatial complexity and low temporal complexity. However filtered scenes with large positive perceptual effects also share similar characteristics of high spatial and low temporal complexity. The only difference between these scenes, displaying negative and positive effects, is the presence of zooming and panning within the scene. The scenes with the largest negative perceptual effects all included prominent zooming effects and, to a lesser degree, panning. A possible reason for the negative perceptual effects therefore may be caused by characteristics of low spatial and temporal complexity coupled with zooming within the scene.

The performance of the in-loop de-blocking filter may be adversely affected by zooming due to the method employed to reduce blocking artifacts. The de-blocking algorithm smoothes the edges of macroblocks to attenuate the perception of artificial boundaries. During a zooming scene, most of the macroblocks are assigned motion vectors pointing in the same direction for a particular amount of time. If in-loop filtering is activated on these blocks, their edges are subsequently smoothed followed by a movement in the direction of their motion vector. In the next frame, pixels defining the macroblock edges are going to be located inside or outside the macroblock. If the in-loop filtering is still applied, smoothing will be applied to the new pixels defining the macroblock edge. This could ultimately produce an

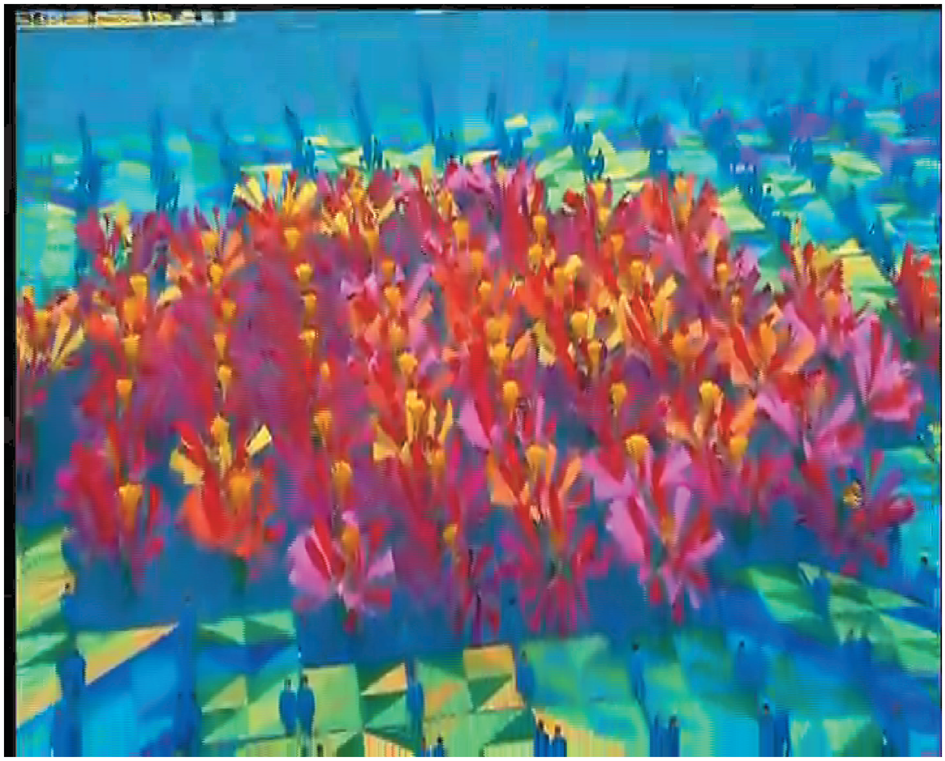

Fig. 6 Scene 8# with de-blocking applied.

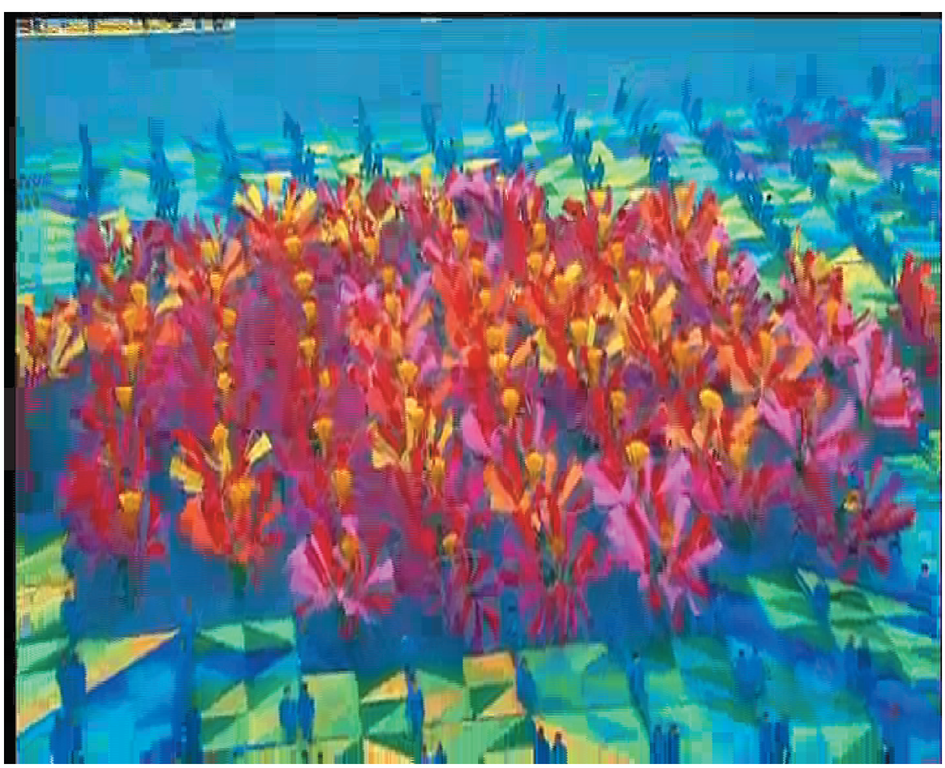

Fig. 7 Scene 8# without de-blocking applied.

effect of blurred pixels propagating along a motion vector. The same effect might be noticeable on panning scenes, although in a less visible way because simple panning creates less residual pixels to code than in a zooming scene. Therefore during a zoom scene de-blocking should be applied less strongly to reduce blurring effects.

Table 3 Spatial and temporal complexity of each reference scene. Scenes ranked according to filter ON DMOS—Filter OFF DMOS.

Scene ID #	Perceptual Effect of Filter	Spatial Complexity	Temporal Complexity
6#	-	43.00	4.71
8#	-	38.89	5.94
5#	-	20.55	8.60
2#	-	30.46	7.96
3#	=	19.16	11.98
4#	+	27.85	11.53
1#	+	18.39	15.30
7#	+	36.02	7.02
9#	+	36.04	3.53

5. Conclusion

This study identified both negative and positive effects of H.264 in-loop filtering on subjective perceptual quality ratings. The tests, using standard resolution video, found a significant interaction between the de-blocking filter and scene content. From the results of this study it is concluded that the performance of the H.264/AVC in-loop de-blocking filter is content-dependent, and that application of the filter can on occasion introduce perceptible degradations to video sequences.

Degradations introduced by the de-blocking filter were characterized by an over smoothing of detailed regions resulting in a blurring effect and general loss of detail. We suggest this could be introduced when de-blocking is applied to scenes with zooming and low temporal complexities. De-blocking could cause blurring when applied to scenes with zooming because the macroblock boundaries constantly change as de-blocking is applied. This could ultimately produce an effect of blurred pixels propagating along a motion vector. Further study is necessary to fully validate this claim and to define much more precisely the image characteristics that affect the performance of the de-blocking filter. Furthermore a metric to quantify panning and zooming would be needed to determine when to flag the application of the de-blocking filter to video. Such a tool would be a valuable option for H.264 encoders so that de-blocking is only applied to appropriate video content.

References

[1] ITU-T Rec. H.264, "Advanced video coding for generic audiovisual services," Geneva, March 2005.

[2] N. Kamaci and Y. Altunbasak, "Performance comparison of the emerging H.264 video coding standard with the existing standards," IEEE Int. Conf. Multimedia and Expo, AIVP-P2.3, pp.I-345–348, Baltimore, MD, July 2003.

[3] M. Horowitz, A. Joch, F. Kossentini, and A. Hallapuro, "H.264/AVC baseline profile decoder complexity analysis," IEEE Trans. Circuits Syst. Video Technol., vol.13, no.7, pp.704–716, July 2003.

[4] A. Dueñas, J.C. Pujol, A. Martín, C. Peláez, F. Díaz, G. Gomez, and F. Martin, "H.264/MPEG-4 AVC real time implementation on DSP," IBC Conference, Amsterdam, Sept. 2003.

[5] J. Lainema and M. Karczewicz, "TML 8.4 loop filter analysis," ITU-TSG16 Doc. VCEG-N29, 2001.

[6] S. Wolf, M. Pinson, A. Webster, G. Cermak, and E.P. Tweedy, "Objective and subjective measures of MPEG video quality," ANSI T1A1 Contribution Number T1A1.5/96-121, Oct. 1996.

[7] J. Allnatt, Transmitted-picture Assessment, John Wiley & Sons, Chichester, 1983.

[8] D.S. Hands, M.D. Brotherton, A. Bourret, and D. Bayart, "Subjective quality assessment for objective quality model development," Electron. Lett., vol.41, no.7, pp.408–409, March 2005.

[9] Y. Zhong, I. Ricardson, A. Miller, and Y. Zhao, "Perceptual quality of H.264/AVC de-blocking filter," IEE VIE 2005, pp.379–384, Glasgow, April 2005.

[10] ITU-T Rec. P.910, "Subjective video quality assessment methods for multimedia applications," Geneva, 1999.

[11] Video Quality Experts Group, www.vqeg.org

[12] ITU-R Rec. BT.500-11, "Methodology for the subjective assessment of the quality of television pictures," Geneva, 2002.

[13] P. Corriveau, C. Gojmerac, B. Hughes, and L. Stelmach, "All subjective scales are not created equal: The effects of context on different scales," Signal Process., vol.77, pp.1–9, 1999.

[14] M. Pinson and S. Wolf, "Comparing subjective video quality testing methodologies," SPIE Video Communications and Image Processing Conference, pp.573–582, Lugano, Switzerland, July 2003.

[15] VQEG, "Final report from the video quality experts group on the validation of objective models of video quality assessment," http://www.vqeg.org/, March 2000.

[16] R. Voelcker, D. Hands, A. Davis, A. Bourret, and D. Bayart, "Video quality measurement—Predicting subjective quality using perceptual models," J. Communications Network, vol.1, no.3, pp.40–45, 2003.

Matthew D. Brotherton received a B.Sc. in Cognitive Science in 2004 from Keele University. Whilst studying for his degree he completed two internships conducting subjective tests for development of world class objective measurements tools at British Telecom. He now works in the Perceptual Engineering Group within BT's Broadband Applications Research Centre. He currently works on issues regarding subjective quality assessment for video, evaluation of emerging video technologies and verification testing for objective models.

Damien Bayart has been working on issues regarding perceptual video quality assessment since joining BT in 2001. Prior to joining BT, Damien graduated with a degree in Telecommunication Technologies from ENIC in France. He is currently working in the Perceptual Engineering Group within BT's Broadband Applications Research Centre. Damien has worked on the development of the Full Reference and Reduced/No Reference perceptual quality models. He is now working on the evaluation and the optimization of implementations of emerging video technologies such as MPEG-4 AVC.

David S. Hands has been working in the field of objective and subjective video quality assessment for the past ten years. He was involved in the RACE II MOSAIC project that defined the internationally standardised continuous quality subjective assessment method and is leading BT's development of objective video quality tools. He is co-chair of the VQEG multimedia group and is an active member of ITU-T SG9 and ITU-R WP6Q.