

VARIABILITY OF PERFORMANCE IN VIDEO CODING

Don Pearson
Department of Electronic Systems Engineering
University of Essex
Colchester CO4 3SQ
UK
dep@essex.ac.uk

ABSTRACT

Modern video compression techniques exhibit variability of performance as a function of time. Studies are reported of viewers' reactions to this variability, which indicate a sensitivity to particular features. Some interesting conclusions emerge for future work in video coding.

1. CONCERNS ABOUT THE INFLUENCE OF COMPRESSION ON PICTURE QUALITY

Some of us have been in the field of picture coding long enough to remember an attitude towards video compression which was quite different from that prevailing today. Currently, techniques such as MPEG-2 are widely accepted as making engineering and commercial good sense for the efficient delivery of broadcast television programmes. But this was not always so. An early assessment of three bandwidth-reduction schemes by Newell and Geddes of the BBC Research Department [1] concluded that "The quality of reproduction was in each case noticeably inferior to that which can be achieved with the conventional system", and "One result of the investigation is that the authors have come to believe that for broadcast entertainment, bandwidth compression methods of the types investigated are unlikely to prove worthwhile." When this paper was read at the Institution of Electrical Engineers in London on 10 January 1962, Professor Colin Cherry of Imperial College said, in the ensuing discussion, "I should have preferred the authors to come to conclusions on what future research ought to be done, rather than to the conclusion that it cannot be done". The authors, in reply, agreed that "our work does not justify the conclusion that bandwidth compression is impossible", but they did not take the opportunity of expressing optimism about its prospects!

In the following thirty years, systems for data reduction of video signals were developed largely outside the broadcasting domain, with

applications such as the videophone and videoconferencing in mind. As techniques such as variable-length coding, conditional replenishment and motion compensation were introduced, it became increasingly evident that coding performance was variable as a function of picture content; the higher the compression attempted, the more variable the quality. During this period broadcasters stuck firmly to the general philosophy put forward by Newell and Geddes, namely that "almost all television development has been towards the attainment of a higher degree of realism", and insisted that most compression techniques could not be used in broadcasting, because a high standard of performance could not be guaranteed.

In the 1990s, in an era of increasing competition and pressing requirements to cut costs, broadcasters have finally buried their suspicions of compression technology and have more or less wholeheartedly embraced it. But what is interesting about this little piece of history is that the issues have not changed; they remain the same in 1997 as they were thirty-five years earlier. Video compression *does* introduce degradation - not always, but sporadically, and with a degree of visibility that depends on the bit rate. What we need to do is to understand its causes and extent in order to control it *to the satisfaction of viewers*. This is the subject of my paper.

2. CAUSES OF VARIABILITY

There is still widespread ignorance, often among more theoretically-inclined research workers, about the limitations of using a single coding technique to compress video. Nowadays there are a wonderful variety of methods to choose from, among them interframe and intraframe prediction, transform coding, fractals, vector quantisation, model-based, morphological, and others. Researchers often become mentally fixated and enthused about just one of these, or some new addition to this list. They then present a series of

papers at conferences and in journals, using short sequences of video to prove their point. Invariably these video presentations involve just a few seconds of picture material. All too frequently in journals the results are expressed in terms of PSNR, sometimes without any description at all as to the video content. PSNR is an unreliable measure of video quality, but there is an additional reason for avoiding it: this is that it can induce a state of mind that is inimical to the understanding of coding performance. Performance variability stems from the interaction between coding method and scene content [2]; simple stochastic video source descriptions fail to recognise this.

The fact is that no compression technique, analogue or digital, works well with all types of scene. One of the earliest, line interlacing, produces patterning on certain types of moving object. Colour compression schemes such as NTSC or PAL exhibit severe cross-colour effects in some high spatial-frequency areas. Block-based transform coding has been known since the earliest days to have difficulty with diagonal lines traversing a block. Fractal coding may work spectacularly with certain types of iterated structure but not so well with others. Facsimile codes are much more

efficient for some types of document layout. Model-based coding does not work very well if new objects keep entering the scene.

Coding is really a high-level language for the description of pictorial material. The language saves bits by characterising groups rather than individual pixels. But a language that is efficient for one type of scene e.g. run-length coding for black-and-white documents, may not be efficient for others such as detailed continuous-tone images.

3. VARIABILITY OF CODING PERFORMANCE USING HYBRID INTERFRAME CODING

3.1 VBR mode

With hybrid interframe techniques such as those used in H261/3 and MPEG-1/2, a considerable fluctuation of performance is observed at low bit rates. In VBR mode, coded broadcast material has been found have a peak-to-mean ratio of 4.7 [3]. Lodge [4], [5] documented the entropy of 27,369 short video sequences; the results again show a wide variations as a function of sequence content. Head-and-shoulders (videoconference-type) sequences were found to be among the easiest to code.

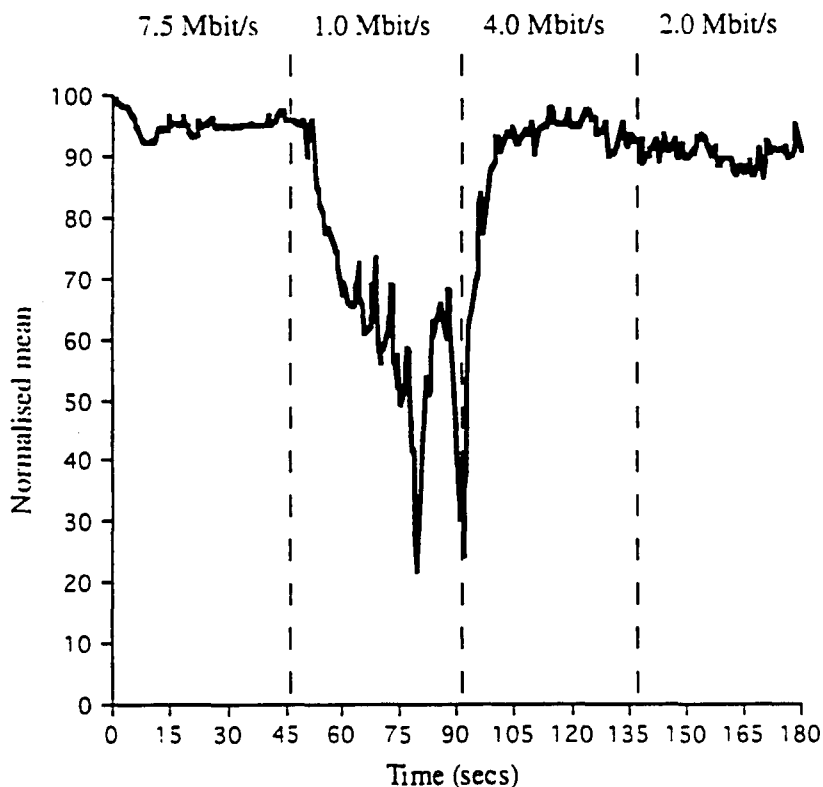


Fig. 1 SSCQE recordings of a sequence MPEG-2 coded at different bit rates [8].

3.2 CBR mode

It is interesting to compare the earlier VBR observations with recent recordings (Figs. 1 & 2) of quality variations in CBR mode, recorded by the SSCQE method, developed recently in RACE project MOSAIC [6],[7]. If MPEG-2 video is coded at a high bit rate (small compression ratio), then the picture quality as a function of time is relatively constant. But as the compression ratio is progressively increased (and bit rate decreased), so there comes a point when the interaction of the coding method with image texture and movement gives rise to sporadic visible distortions. The frequency and magnitude of these distortions both increase as bit rate decreases, until eventually a trace like that in Fig. 2 is obtained, where the distortion is visible almost all the time.

3.3 Causes of extremes in variability

Evidence suggests that at least some of the negative peaks of quality in CBR transmission are due to the movement of large textured objects, which hybrid interframe block-based methods are not efficient at coding [10]. An alternative method such as model-based coding or global motion compensation is needed to deal with such objects [10], [11].

4. EFFECT OF QUALITY VARIABILITY ON HUMAN VIEWERS

How do viewers react to quality variations? In

studying this question experimentally, we first learnt that they are disproportionately influenced by what they see in the last 10s or so of a test sequence [12]. This recency effect can alter the overall or end judgement that a subject makes of a presentation by half a scale point or more on a five-point quality scale.

Modelling of these variations by our Dutch colleagues [13] in project MOSAIC suggests that the negative peaks play an important role in determining a subject's overall response to the sequence. More recent work at Essex [14] implies that it is the magnitude rather than the duration of the negative peaks that is significant. This fits well with work by psychologists that show that magnitude rather than duration is important ("duration neglect") in human reaction to other experiences such as pain [15].

5. IMPLICATIONS FOR SWITCHED CODERS

The results of this work suggest that in current work on switched or layered coders (Fig. 3) [16], attention needs to be given in the code selection strategy to the occasions when the quality suffers badly. These have a disproportionate effect on the viewer. Overall, it confirms the approach whereby video quality is likely to be maximised by switching appropriately between several different coding methods or models.

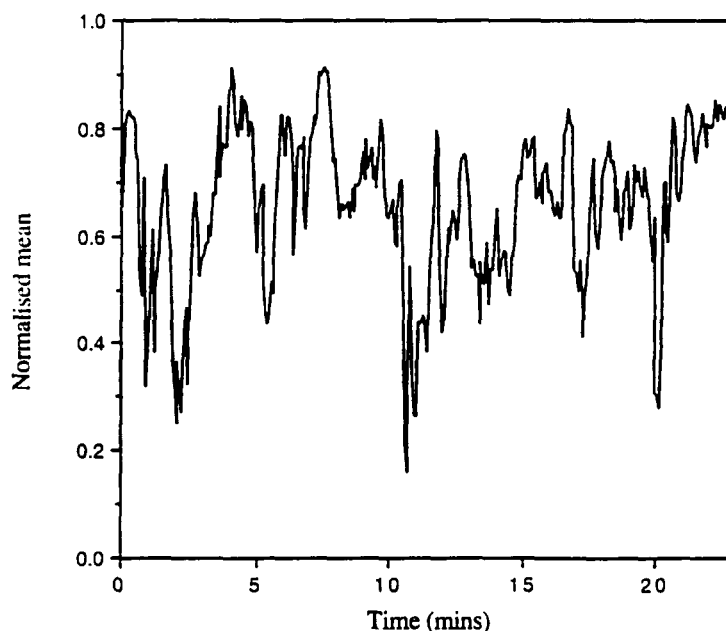


Fig. 2 Quality variations recorded using the SSCQE method [9] for a 20-minute television sequence encoded at 2.5 Mbit/s using hybrid interframe coding.

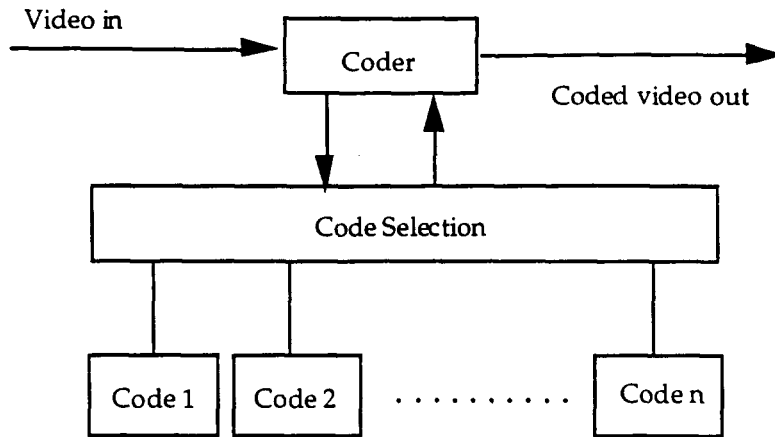


Fig. 3 Code switching to reduce quality variations

6. CONCLUSIONS

Variability in coding efficiency as a function of time and content is a characteristic of video compression techniques. With a recently-developed method of measuring the moment-by-moment variation in picture quality that results in CBR mode, we have begun to document the relationship between variability and the degree of compression. Studies of viewer reaction to variable-quality video have identified (i) the end-section and (ii) the depth of the negative peaks as being particularly influential. The latter has implications for code-selection strategies in switched coders; in particular, it points to strategies which avoid large decrements in coding performance.

7. ACKNOWLEDGEMENT

The author acknowledges with thanks the support of the Independent Television Commission and the contributions of European partners in respect of the MOSAIC project work reported in this paper.

8. REFERENCES

- [1] G.F. Newell and W.K.E. Geddes, "Tests of Three Systems of Bandwidth Compression of Television Signals", *Proc. IEE*, vol. 109, part B, pp 311-324, 1962.
- [2] M. Ghanbari and D.E. Pearson, "Components of bit-rate variation in videoconference signals", *Electronics Letters*, vol. 25, no. 4, pp. 285-286, 16 February 1989.
- [3] M. Ghanbari and D.E. Pearson, "Variable bit-rate transmission of television signals", *Electronics Letters*, vol. 24, no. 7, pp. 392-393, 1988.
- [4] N. Lodge, "A method for the characterisation of low bit-rate picture quality" PhD thesis, Heriot-Watt University, 1993.
- [5] N. Lodge and D. Wood, "Subjectively Optimising Low Bit-rate Television" *Proc. International Broadcasting Convention IBC 94*, Amsterdam, September 1994.
- [6] R. Aldridge, J. Davidoff, M. Ghanbari and D. Hands and D. Pearson, "Subjective Assessment of Time-Varying Coding Distortions", *Proc. International Picture Coding Symposium*, pp. 269-274, Melbourne, Australia, 13-15 March 1996.
- [7] Draft revision of Recommendation ITU-R BT.500-7, "Introduction of a New Method for Single Stimulus Continuous Quality Evaluation (SSCQE)", ITU-R Document 11E/9, 4 June 1996.
- [8] R. Aldridge, J. Davidoff, M. Ghanbari and D. Hands and D. Pearson, "Evaluating Digital Television Picture Quality" submitted for publication.
- [9] R. Aldridge, PhD thesis, University of Essex, 1996.
- [10] D. Pearson, "Selective Use of Model-based Coding for Large Moving Objects", *Proc. VIII European Signal Processing Conference EUSIPCO-96*, paper SS.2.7, Trieste, Italy, 10-13 September 1996.
- [11] H. Jozawa et al., "Video Coding using Adaptive Global MC and Local Affine MC", *Proc. VIII European Signal Processing Conference EUSIPCO-96*, paper SS.2.4, Trieste, Italy, 10-13 September 1996.
- [12] R. Aldridge, J. Davidoff, M. Ghanbari, D. Hands and D.E. Pearson, "Measurement of scene-dependent quality variations in digitally-coded television pictures", *IEE Proceedings on Vision, Image and Signal Processing*, vol. 142, no. 3, pp. 149-154, 1995.
- [13] H. de Ridder, "Current issues and New Techniques in Visual Quality Assessment", *Proc. IEEE Conference on Image Processing ICIP-96*, Vol. 1, pp. 869-872, 1996.
- [14] D. Hands, PhD thesis, University of Essex, in preparation.
- [15] B.L. Frederickson and D. Kahneman, "Duration neglect in retrospective evaluations of affective episodes", *J. Personality and Social Psychology*, vol. 65, pp. 45-55, 1993.
- [16] D.E. Pearson, "Developments in Model-based Video Coding" *Proc. IEEE*, vol. 83, no. 6, pp. 892-906, 1995.