A Fractal Video Communicator

J. Streit, L. Hanzo

Department of Electronics and Computer Sc., University of Southampton, UK, S09 5NH

Abstract

The image quality and compression ratio trade-offs of five different 176 × 144 pels quarter common intermediate format (QCIF) fractal image codecs are investigated by simulation. The average peak signal-to-noise ratio (PSNR) ranges from 29 dB to 37 dB, while the coding rate from 0.24 bit/pel (bpp) to 1.22 bit/pel, as seen in Table 1. Two of the candidate codecs, a 0.28 bit/pel and a 1.1 bit/pel codec, were subjected to bit-sensitivity analysis and protected by the source-sensitivity matched shortened binary BCH(122,80,6) and BCH(122,52,11) codes and transmitted using coherent pilot symbol assisted (PSAM) square-constellation 16-level quadrature amplitude modulation (16-QAM). The proposed fractal video communicators required a channel signal-to-noise ratio (SNR) and signal-to-interference ratio (SIR) of about 15 dB in order to maintain a video peak SNR (PSNR) of 31 dB and 35 dB at signaling rates of 39 kBaud and 156 kBaud, respectively, over Rayleigh-fading channels having a propagation frequency of 1800 MHz and a pedestrian speed of 2 mph.

1 Introduction

Previously proposed fractal codec designs were targeted at high-resolution images having large intra-frame domain-block pools [1], [2]. Following the approaches proposed by Barnsley [3], Jacquin [2], Monro [1], [5], Ramamurthi [6] et.al and Beamont [4] in this study we explored the range of trade-offs available using five different head-and-shoulders fractal video-phone codecs (Codecs A-E) and compared their complexity, compression ratio and image quality specifically for low-resolution, small-pool 176 × 144 pixels Quarter Common Intermediate Format (QCIF) CCITT standard videophone images.

In Section 2 on fractal coding two of the candidate codecs were then selected for further investigations when incorporated in a portable fractal video transceiver. In order to maintain high robustness against channel errors and low signaling rate, source sensitivity-matched binary Bose-Chaudhuri-Hocquenghem (BCH) coding combined with unequal protection 16-level quadrature amplitude modulation (16-QAM) is proposed in Section 3, while Section 4 protrays the performance of our transceiver, before concluding in Section 5.

2 Fractal Image Codecs

In fractal image coding the QCIF image to be encoded is typically divided into 4-by-4 or 8-by-8 pels non-overlapping range blocks (RB), which perfectly tile the original image [3]. Every RB is then represented by the contractive affine transformation [1] of a larger, typically quadruple-sized domain block (DB) taken from the same frame of the original image. In general, the larger the pool of domain blocks, the better the image quality, but the higher the computational complexity and the bit rate, requiring a compromise. For gray-scale coding of two-dimensional images a third dimension representing the brightness of the picture must be added, before affine transformation takes place. Furthermore, for the sake of reduced complexity the legitimate affine transforms are restricted to the following manipulations [2]:

1. Linear translation of the block.
2. Rotation of the block by 0, 90, 180 and 270 degrees.
3. Reflection about the diagonals, vertical and horizontal axis.
4. Luminance shift of the block.
5. Contrast scaling of the block.

In order to achieve the required compromise, the Collage Theorem [3] and contractivity requirements permit only the following ($DB \times RB$ size combinations : $(16\times16, 8\times8)$, $(16\times16, 4\times4)$, $(8\times8, 4\times4)$ and our codecs attempt to match every RB with every DB of the same frame allowing rotations by $0^o, 90^o, 180^o or 270^o$. Using the mean squared error of

$$MSE = \sqrt{\sum_{i=1}^{p}\sum_{k=1}^{p}(X_{ik} - Y_{ik})^2} \qquad (1)$$

as block-matching distortion measure, where p is the RB size, the optimum contrast scaling factor a and luminance shift b for the contracted DBs Y_{ik} and RBs X_{ik} can be derived by

0-7803-1927-3/94/$4.00

Codec	DB Size	RB Size	Classi-fication	Split	PSNR (dB)	Rate (bpp)
A	16	8	None	No	31	0.28
B	16	4	None	No	35	1.1
C	8	4	None	No	37	1.22
D	16	8/4	Twin	Yes	36	0.84
E	16	8/4	Quad	Yes	29	1.0

Table 1: Comparison of five fractal codecs

minimising the MSE defined above leading to:

$$b = \frac{\sum_{i=1}^{p}\sum_{k=1}^{p} X_{ik}Y_{ik} \sum_{i=1}^{p}\sum_{k=1}^{p} X_{ik}}{\sum_{i=1}^{p}\sum_{k=1}^{p} X_{ik}^2 - \sum_{i=1}^{p}\sum_{k=1}^{p} X_{ik}^2 p^2} - \frac{\sum_{i=1}^{p}\sum_{k=1}^{p} Y_{ik} \sum_{i=1}^{p}\sum_{k=1}^{p} X_{ik}^2}{\sum_{i=1}^{p}\sum_{k=1}^{p} X_{ik}^2 - \sum_{i=1}^{p}\sum_{k=1}^{p} X_{ik}^2 p^2} \quad (2)$$

$$a = \frac{\sum_{i=1}^{p}\sum_{k=1}^{p} Y_{ik}}{\sum_{i=1}^{p}\sum_{k=1}^{p} X_{ik}} - \frac{p^2}{\sum_{i=1}^{p}\sum_{k=1}^{p} X_{ik}} \cdot b. \quad (3)$$

The minimum achievable MSE when using the optimum coefficients a and b is given by:

$$E = \sqrt{\sum_{i=1}^{p}\sum_{k=1}^{p}(X_{ik} - (aY_{ik} + b))^2}. \quad (4)$$

In order to identify the range of design trade-offs five different codecs, Codecs A-E, were simulated and compared in Table 1. The DB indeces and four different rotations were Gray-coded, while the luminance shift b and contrast scaling a were Max-Lloyd quantised using four bits. Comparison of the three basic schemes, Codecs A-C featured in Table 1, suggested that a RB size of 4x4 used in Codecs B and C was desirable in terms of image quality, having a peak signal-to-noise ratio (PSNR) of 5-6 dB higher than Codec A. However, Codec A had an approximately four times higher compression ratio or lower coding rate expressed in bits/pels (bpp). Furthermore, had four times more DBs than Codec B, which yielded quadrupled block-matching complexity, but the resulting PSNR improvement was limited to about 1 dB and the bit rate was about 20 % higher due to the increased DB addressing. These finding were also confirmed by informal subjective assessments.

In order to find a compromise between the four times higher compression ratio of the 8×8 RBs used in Codec A and the favourable image quality of the 4×4 RBs of Codec B and C, we decided to split inhomogeneous RBs in two, three or four sub-blocks [2]. Initially the codec attempted to encode an 8×8 RB and calculated the MSE associated with the particular mapping. If the MSE was above a certain threshold, the codec split up the block into four non overlapping sub-blocks. The MSE of these sub-blocks was checked against the error threshold individually and if necessary one or two 4×4 sub-blocks were coded in addition. However, for three or four poorly matching sub-blocks, the codec stored only the transforms for the four small sub-cells. This splitting technique was used in Codecs D and E of Table 1.

Block Type	Description / Edge Angle	Frequency (%)
Shade	no significant gradient	18.03
Midrange	moderate gradient, no edge	46.34
Edge	steep gradient, edge detected	26.39 (total)
	0 deg	2.66
	45 deg	1.85
	90 deg	5.69
	135 deg	3.18
	180 deg	3.92
	225 deg	2.96
	270 deg	3.92
	315 deg	2.22
Mixed	Edge angle ambiguous	9.24

Table 2: Classified Block Types and Their Relative Frequencies in Codec E

In addition to the above splitting technique, the subjectively important edge representation of Codecs D and E was improved by a block classification algorithm [2], [6]. Accordingly, the image blocks were classified into four classes:

1. Shade blocks taken from smooth areas of an image with no significant gradient.
2. Midrange blocks having a moderate gradient but no significant, edge.
3. Edge blocks having steep gradient and containing only one edge.
4. Mixed blocks with steep gradient that contain more than one edge and hence the edge orientation is ambiguous.

Codec D used a basic twin-class algorithm, differentiating only between shade and non-shade blocks, whereas Codec E used the above quad-class categorisation. The relative frequencies of all registered sub-classes of Codec E are shown in Table 2.

In Codecs D and E after the classification of all DBs and RBs normal coding ensued, but with the advantage that the codec predetermined by what angle the DB had to be rotated and it attempted to match only blocks of the same class. Namely, if for example a RB was classified as an edge block with a certain orientation, the codec exploited this by limiting the required search to the appropriate DB pool. Furthermore, shade blocks were not fully encoded, only their mean was transmitted to the decoder, yielding a significant reduction in complexity and bit rate.

Bit Index	Parameter
1-2	Rotation
3-6	RB X coordinate
7-10	RB Y coordinate
11-14	Scaling
15-18	Offset

Table 3: Bit allocation per RB for Codecs A and B

A comparison of the five QCIF videophone codecs is presented in Table 1. Codecs A-C represent basic fractal codecs with no RB classification and splitting. When comparing Codecs A and B using RB sizes of 8×8 and 4×4, respectively, the compression ratio of Codec A is four times higher, but its PSNR is 5 dB lower at similar complexity. The 1 dB PSNR advantage of Codec C does not justify its quadruple complexity. Codecs D and E deploy twin- or quad-class block classification combined with RB splitting, if the MSE associated with a particular mapping is above a certain threshold. Interestingly, the less complex Codec D has a higher compression ratio and higher image quality. The lower performance of codec E is attributed to the limited size of the DB pool provided by our QCIF images. Table 1 provides further interesting trade-offs for system designers.

Having designed a range of fractal video codecs we shortlisted the 0.28 bits/pel codec A and the 1.1 bits/pel Codec B for further investigation in the proposed video transceiver. Both of these codecs have an identical bit allocation scheme for each RB, which is portrayed in Table 3.

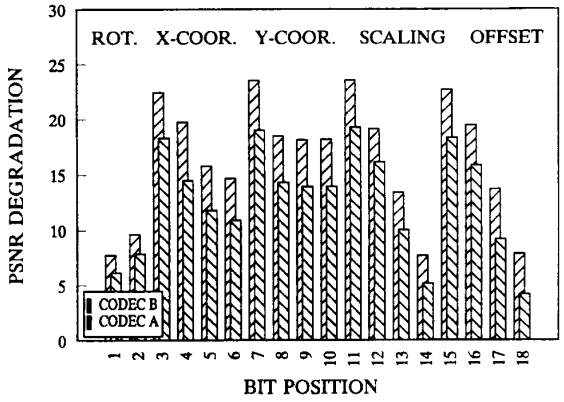

Figure 1: PSNR Degradation versus Bit Index for Codecs A and B

Bits 1-2 represent four possible rotations, bits 3-6 and 7-10 are the X and Y range-block coordinates, respectively, while bits 11-14 are the Max-Lloyd quantised scaling gains and bits 15-18 represent offset values used in the random collage algorithm. However, codec A uses $22 \times 18 = 396$ 8×8 pels RBs associated with a rate of $R = 18/64 \approx 0.28$ bits/pel, while Codec B has $44 \times 38 = 1584$ 4×4 RBs, which is associated with a quadrupled bit rate of $R = 18/16 \approx 1.1$ bits/pel. The number of bits per frame becomes $396 \cdot 18 = 7128$ bits/frame for Codec A and 28512 for Codec B, corresponding to bit rates of 71.28 kbits/sec and 285.12 kbits/sec, respectively, at a scanning rate of 10 frames/sec. The associated PSNR values are about 31 and 35 dB, respectively.

Both codec A and B were subjected to bit sensitivity analysis by consistently corrupting each bit of the 18-bit frame and evaluating the PSNR degradation inflicted. These results are shown for both codecs in Figure 1. Observe from these figures that the significance of the specific coding bits can be explicitly inferred from the PSNR degradations observed. Therefore the more sensitive bits have to be protected more strongly than their less vulnerable counterparts, an issue to be addressed in the next Section.

3 The Video Transceiver

The schematic diagram of the proposed fractal transceiver is shown in Figure 2. The fractal coded bits are mapped in two sensitivity classes and protected by the twin-class source sensitivity-matched binary Bose-Chaudhuri-Hocquenghem (BCH) encoders shown in the Figure. The BCH-coded information is then block-diagonally interleaved over an image frame in order to disperse burst errors before this information is mapped to the input of the 16-level Quadrature Amplitude Modulator (16-QAM) employed.

A variety of QAM schemes having different strengths and weaknesses have been proposed in the literature [7]. In order to maintain as low a transmitted power requirement as possible we have opted for a second-order diversity assisted coherent Pilot Symbol Assisted Modem (PSAM) using the well-known maximum minimum distance square shaped 16-QAM constellation. The performance of this modem over a Rayleigh-fading channel has been documented in reference [8] for a pedestrian speed of 4 mph, propagation frequency of 1.8 GHz and various pilot symbol spacing distances.

In order to achieve high fade-tracking efficiency in our proposed transceiver we used a pilot separation of ten symbols. Under these conditions this modem provides two independent QAM subchannels that exhibit different bit error rates (BER), which is about a factor three to four times lower for the higher integrity path referred to as Class 1 (C1) subchannel than for the C2 subchannel over Rayleigh-fading channels. This consistent BER difference is maintained over a range of channel signal to noise ratios (SNR) around 20 dB, a value realistically targeted in the benign microcellular personal communications (PCN) environment, provided that similarly favourable interference levels can be maintained.

This property can be exploited to provide un-equal source sensitivity-matched error protection for the fractal video codec [7]. If the BER ratio of these subchannels does not match the integrity requirements of the source codec, it can

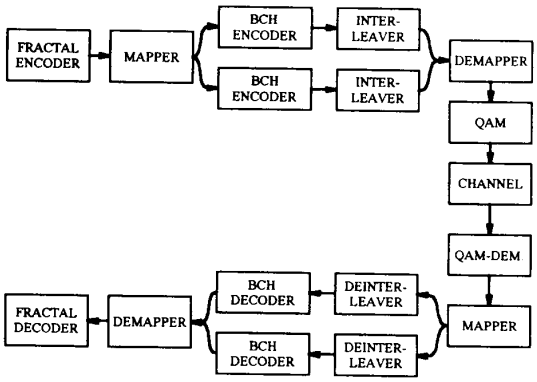

Figure 2: Fractal Video Transceiver Schematic

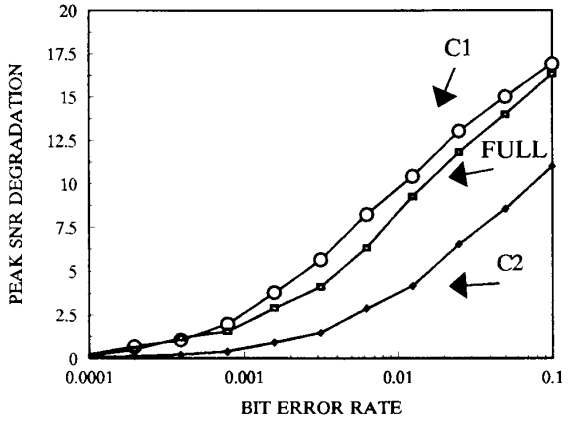

Figure 3: PSNR Degradation versus Bit Error Rate for Codec A

be arbitrarily adjusted by using different BCH forward error correction (FEC) codecs in both subchannels. In the selection of the source-matched FEC codecs we have to ensure however that the number of FEC-coded bits in both subchannels is identical. This implies that when using different FEC codecs for the protection of the more and less vulnerable video source bits transmitted over the higher and lower integrity C1 and C2 QAM subchannels, respectively, the subchannels deliver a different number of video source bits.

Explicitly, the increased number of redundancy bits of stronger FEC codecs provide a monotonously decreasing capacity, increasing integrity subchannel for the transmission of the more sensitive source bits. Hence there will be an optimum FEC coding power, above which the system's robustness is reduced upon increasing the FEC coding power due to directing too low a number of high importance bits over the high integrity route. This inevitably relegates too many comparatively important source bits to the lower integrity subchannel, whose coding power must be reduced in order to be able to accommodate a higher number of source bits.

Initially we therefore divided the video source bits in two subclasses, which contained an equal number of bits from both video source bit classes and evaluated the PSNR degradation due to inflicting an identical fixed bit error rate using random bit corruption in both classes. These results are depicted for Codec A in Figure 3, suggesting that an approximately three to four times lower bit error rate is required for the more vulnerable source bits in order to guarantee similar PSNR degradations to those due to the more robust corrupted source bits. Similar results were obtained also for Codec B. Since this transmission integrity requirement coincided with that provided by the C1 and C2 16-QAM subchannels, this system was initially implemented using the systematic binary BCH codecs [9] BCH(122,66,9) in both subchannels. These codecs encode 66 source bits using 122 channel coded bits and can correct 9 errors per frame, corresponding to an error correction capability of about 7 %.

In an attempt to perfectly match the FEC coding power and the number of bits in the distinct protection classes to the video source sensitivity requirements we also evaluated the performance of a variety of different schemes, while maintaining the same overall coding rate and bit rate. For example, when using the the BCH(122,80,6) and BCH(122,52,11) codecs in the C1 and C2 16-QAM subchannels to protect the more and less sensitive video bits, respectively, the overall coding rate of R=(66+66)/(122+122)=(52+80)/(122+122) $\approx$ 0.54 was maintained.

Using the video bit rates of 71.28 and 285 kbits/s derived in Section 2, the corresponding FEC coded rates are $71.28/R \approx 132$ and $285/R \approx 527$ kbits/s, corresponding to signaling rates of 132/4=33 kBd and $527/4 \approx 132$ kBd respectively. The 16-QAM bursts are constituted by 61 information symbols, 7 pilot symbols according to the pilot spacing of $P = 10$ and 4 ramp symbols, yielding a burst length of 72 16-QAM symbols. Consequently, the signaling rates become $33 \cdot 72/61 \approx 39$ kBd and $132 \cdot 72/61 \approx 156$ kBd.

In this treatise we follow the Digital European Cordless Telecommunications (DECT) system using a bandwidth of 1728 kHz, but adopt a time division multiple access (TDMA) scheme. The maximum possible signaling rate can be computed from the 2.4 bits/s/Hz bandwidth efficiency of our 16-QAM modem, which implies a filtering excess bandwidth of 50 % and a modulated spectrum attenuation of 24 dB at the transmission band edge [7]. Then the maximum channel rate is $2.4 \cdot 1728$ kHz=4147.2 kbits/s $\approx$ 1037 kBd, accommodating about 26 and 6 video subscribers, when using the 39 kBd Codec A and 156 kBd Codec B, respectively. At this signaling rate micro- and pico-cellular cordless systems typically exhibit flat fading and hence require no channel equaliser.

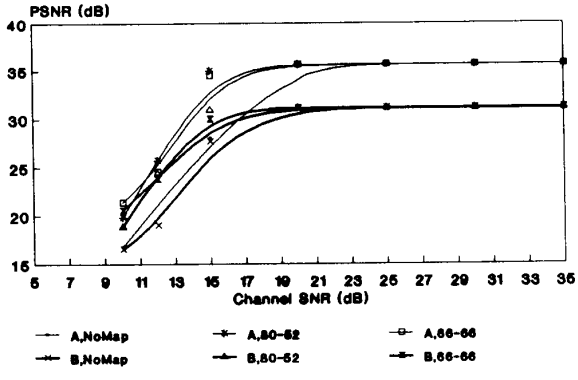

Figure 4: PSNR Degradation versus Channel SNR for Codecs A and B

4 Transceiver Performance

The PSNR versus channel SNR performance of the fractal video transceiver proposed is portrayed in Figure 4 for both Codecs A and B when using three different FEC schemes. These performance curves were evaluated over a Rayleigh-fading channel using a propagation frequency of 1.8 GHz, a pedestrian speed of 2 mph and a signaling rate of 1037 kBd. Observe that the previously mentioned source-matched twin-class BCH(122,66,9) coded scheme has a significantly better performance in case of both Codec A and B than the corresponding arrangements using no mapping. Explicitly, there is an approximately 5 dB channel SNR gain, when using the source-matched mapping schemes associated with the BCH(122,66,9) codecs.

The best overall performance was attributable to the arrangement where the more and less vulnerable video bits were protected by the BCH(122,80,6) and BCH(122,52,11) codecs, respectively. This performance curve is also shown in Figure 4 for both Codec A and B. The slightly superior performance of this scheme over that of the BCH(122,66,9) coded systems was achieved by re-allocating some of the parity bits from the inherently higher integrity C1 16-QAM subchannel to protect the lower integrity C2 subchannel. Observe however from Figure 4 that there is a slight performance penalty towards low channel SNR values, where the higher sensitivity video bits would require slightly more FEC protection.

5 Summary and Conclusions

A range of QCIF fractal video codecs was studied in terms of image quality, compression ratio and complexity. Two fixed-rate codecs, codec A and B were selected for further investigations in a video-telephone transceiver. Both codecs were subjected to rigorous bit sensitivity analysis and it was found that the integrity requirements of the more vulnerable bits are about three to four times higher than those of the less sensitive video bits, when the bits are equally split in two classes. This approximately coincided with the integrity differences of the 16-QAM subchannels, hence both bit protection classes required a similar FEC coding power. This prompted us initially to use the source-matched BCH(122,66,9) codecs in both 16-QAM subchannels, although our further experiments revealed that the best system performance can be achieved in case of both Codec A and B, when using the BCH(122,80,6) and BCH(122,52,11) codes in the C1 and C2 16-QAM subchannels, respectively. The minimum channel SNR and signal to interference ratio (SIR) must be in excess of 15 dB in order to maintain unimpaired image quality for both codecs, although Codec B has a quadrupled bandwidth and complexity, while providing better image quality. The signaling rates of Codec A and B are about 39 kBd and 156 kBd, requiring a video user bandwidth of about 60 kHz and 240 kHz, respectively.

6 Acknowledgement

The financial support of the SERC, UK (GR/J46845) is gratefully acknowledged.

References

[1] D.M. Monro, F.Dudbridge: Fractal Block Coding of Images, Electr. Let. 21st of May 1992, Vol. 28, No. 11, pp 1053-1055

[2] A. E. Jacquin, "Image Coding Based on a Fractal Theory of Iterated Contractive Image Transformations", IEEE Trans. Image Proc. ,Vol 1, Jan. 92, pp 18-30

[3] Michael F. Barnsley, "A Better Way to Compress Images", BYTE, Jan. 88, pp 215-222

[4] J. M. Beaumont, "Image data compression using fractal techniques", BT Technol. Vol. 9 No 4, Oct. 91, pp 93-109

[5] D.M. Monro, D.L. Wilson, J.A. Nicholls: High Speed Image Coding with the Bath Fractal Transform, Proc. of IEEE Symposium on Multimedia Technologies and Future Applications, 21-23 Apr. 1993, Sothampton, UK.

[6] B. Ramamurthi and A. Gersho, "Classified Vector Quantization of Images", IEEE Trans. Commun., Vol 34 No 11, No 86, pp 1258-1268

[7] W.T. Webb, L. Hanzo: Quadrature Amplitude Modulation: Principles and Applications for Fixed and Wireless Communications, IEEE Press-Pentech Press, 1994

[8] R. Stedman, H. Gharavi, L Hanzo, R. Steele, *Transmission of Subbband-coded Images via Mobile Channels*, IEEE Tr. on Circuits and Systems for Video Technology, Febr. 1993, Vol. 3, no.1, pp 15-27

[9] K.H.H. Wong, L. Hanzo: Channel Coding, pp 347-488, Chapter 4 in R. Steele (Ed.) Mobile Radio Communications, Pentech Press, London, 1992