Digital Compression Algorithms for HDTV Transmission

Kenneth C. Adkins. Mary Jo Shalkhauser[†], Steven B. Bibyk

Dept. of Electrical Engineering, The Ohio State University
Columbus, Ohio 43210

[†]Digital Systems Technology Branch, NASA Lewis Research Center
MS 54-8, Cleveland, Ohio 44135

January 31, 1990

Abstract

Digital compression of video images is a possible avenue for HDTV transmission. Compression needs to be optimized while picture quality remains high. Two techniques for compressing digitized images are explained, and comparisons are drawn between the human vision system and artificial compression techniques. Suggestions for improving compression algorithms through the use of neural and analog circuitry are given.

I Introduction

Digital compression algorithms are being examined closely for possible use in High Definition Television (HDTV) transmissions. For digital transmission of HDTV pictures to become a reality, broadcast quality video needs to be compressed in order to remain within typical bandwidths allotted for TV. Figure 1 shows a block diagram of a general transmission of digital information. The compression can be done in either a lossy or lossless manner. Lossy compression involves the use of estimating a "continuum" of values with a discrete set of points, each of which represents a range of values. This estimation, or quantization, causes some loss of information. Lossless compression involves no loss of information; it makes use of statistical properties of the data being transmitted. A Huffman code is a good illustration of lossless compression: code words having higher a priori probabilities are assigned shorter word lengths, thus minimizing the code length.

II Vector Quantization

One popular form of digital compression is vector quantization (VQ). VQ is a lossy algorithm that involves dividing the screen into regular blocks of pixels. e.g. 2 pixels on a side. If each pixel is quantized into an 8 bit word, then the 4 pixel block, or tile, can be represented as a 32 bit vector. This vector can be used as a search key in an associative memory lookup [1] to find the closest match amongst a predetermined

Figure 1: Typical HDTV Transmission

1

"codebook" of vectors. The codebook is a set of relatively few vectors representative of all possible tiles. Since the encoder and decoder both have a copy of the codebook, only the index of the chosen code vector is sent. Thus, if the codebook contained 64 codewords, then only 6 bits are needed to specify the representative vector, giving a compression of 32:6 = 5.3 and a bit per pixel ratio of 6 bits/4 pixels = 1.5 bpp.

Proper codebook generation is a significant factor in determining VQ performance [2]. In principle, the codebook could be updated to match the image it was representing; however, in practice synchronizing the encoder to the decoder too frequently would negate the bandwidth reduction of VQ compression. The manner in which the algorithm handles edge features is also significant. The human eye places a heavy emphasis on edge detection, thus it is important that image reconstruction near edge features be sharp and consistent.

Generally, codebooks are generated using the most frequently repeated tiles in an image. This method of codebook generation performs well on average, but features that appear infrequently within the image are not handled well, since they are not represented in the codebook. Unfortunately, edge features may often fall into this category. The larger the codebook can become, the more accurately it is able to describe the image to be compressed; but at the same time, the vector matching hardware becomes more complex, and the number of bits sent per codebook index needs to be increased. The ability to trade off picture quality for compression performance may desirable in some applications, but not necessarily HDTV, where both properties need to be optimized.

Figure 2: Enhanced DPCM CODEC

III Enhanced DPCM

The enhanced DPCM video data compression CODEC [3] utilizes four techniques to achieve data compression: differential pulse code modulation (DPCM), non-adaptive prediction, non-linear quantization, and multi-level Huffman coding. The compression algorithm incorporates both lossy and lossless mechanisms in these four techniques. The encoder portion of the CODEC performs the data compression, and the decoder reconstructs the image from the compressed image data. A block diagram of the digital portion of the CODEC is shown in Figure 2. The input to the encoder is a digitized NTSC video signal, sampled at four times the color subcarrier frequency (nominally 14.32 MHz) with a resolution of 8 bits per sample (pixel).

DPCM and non-adaptive prediction are predictive techniques used to reduce the amount of redundancy in the video image. DPCM predicts the value of the current pixel based on the average of two neighboring pixels of the same color subcarrier phase, the 4th previous pixel on the same line and same pixel from two lines previous. Redundancy is removed by subtracting the DPCM predicted value (PV in figure 2) from the current pixel value (PIX) resulting in a difference value. The non-adaptive predictor attempts to remove more redundancy by predicting the difference value resulting from the DPCM prediction. Using the fact that the difference values for neighboring pixels are similar, the non-adaptive predictor values are based on the difference values of the previous pixel, and these were determined using statistical analysis on a number of sample images with widely varying picture content. The predicted value (NAP),

2

which is contained in a look-up table (see Figure 3) in the hardware, is subtracted from the DPCM difference value resulting in a new difference value (DIF).

The resultant difference value (DIF) is then quantized into one of 13 quantization levels (QL). Each quantization level (QL) has an associated quantization value (QV) which is representative of the difference value (DIF). The difference values are quantized non-uniformly to allow more resolution for the quantization levels with small difference values, and less resolution for large difference values. This non-uniform distribution is chosen because the human eye is more sensitive to small variations in smooth regions of an image (where the difference values will be small) than it is to large variations at transition boundaries (large difference values). Like the NAP values, the quantization values were determined by statistical analysis on sample images and the QL and QV values are contained in look-up tables in the encoder hardware memory.

The final data compression technique used in the enhanced DPCM CODEC is multilevel Huffman coding [4]. The quantized difference values are assigned Huffman codes based on the probability of occurrence of the quantization level. The most probable quantization levels are assigned the shortest length Huffman codes. Multilevel Huffman codes are used to take advantage of the fact that neighboring pixels fall into the same or close to the same quantization level. Fourteen Huffman code sets are used – one for each of the 13 quantization levels and one for start-up purposes. The Huffman encoder is implemented in a look-up table addressed by QL(N) and QL(N-1). The output of the Huffman encoder is a Huffman code up to 12 bits long and a 4-bit code indicating the length of the Huffman code.

The enhanced DPCM algorithm enjoys an advantage over vector quantization in the manner in which it handles irregular features. Whereas VQ has no good means of representing features that appear infrequently, the DPCM scheme is able to simply transmit longer codewords for unexpected events. The effect is a slight rise in the average bpp ratio in return for higher resolution and better image quality near edges.

The enhanced DPCM CODEC is able to compress a digitized NTSC image from 8 bits/pixel to an average of 1.8 bits/pixel. The reconstructed image quality is excellent and is indistinguishable from the original image.

IV Retinal Image Mapping

Due to the proven efficiency of the retina's image processing, it seems reasonable to match artificial image compression to the natural compression used by the human vision system. The photoreceptors on the retina form a signal proportional to the logarithm of the incoming light intensity. The signal is then spatially and temporally averaged with neighboring photoreceptor signals. A difference between the photoreceptor output and the averaged signal is formed, and it is a function of this signal on which the brain operates.

The DPCM compression technique is similar to the manner in which actual physiological systems "encode" visual information. The non-linear quantization can be compared roughly to the logarithm function, the predicted value (PV) corresponds to a spatial average, and the difference value is the final output in both systems. Qualitative performance of the two systems is also similar. If the video

image is shifted slightly, both the DPCM algorithm and the retinal mapping produce an output similar to but slightly shifted from that of the original image. On the other hand, the VQ system places artificial boundaries in the image; any shift of the video image could result in a very different ordering of the transmitted code indices. We will explore methods of removing redundancy with loss in a manner that is similar to the losses incurred in the retina.

V   Channel Error Effects

In addition to compressing video data, the manner in which an algorithm can detect and correct bit errors introduced by a noisy communications channel is another important performance criterion. Errors must occur at rates that are not visible or bothersome to the viewer.

The VQ compression algorithm has perhaps its biggest advantage in terms of error performance. Since each block is represented by a fixed length codeword, any errors that occur are localized to the block in which the error occurred; there is no cascading effect. This is analogous to the biological system in that each fixed length codeword maps to an optical nerve. The small amount of information that is lost in a single block of data or along a single nerve is so negligible that processing can continue uninterrupted.

One drawback to the DPCM algorithm is its non-graceful performance degradation due to bit errors. Due to their variable length nature, Huffman encoded data is extremely vulnerable to bit errors. When an error occurs in the Huffman code data, proper decoding is unlikely and will result in loss of synchronization of codeword boundaries. In the case of video data compression, loss of synchronization will result in poor image quality.

The enhanced DPCM CODEC employs line and field resynchronization techniques to reduce the effects of channel errors on the reconstructed image quality. A unique word is inserted into the compressed data stream at the beginning of each video line and each field (a different unique word is used for the lines and the fields) by the encoder. The decoder detects the occurrence of the unique words and also counts the number of Huffman codes received for each line and each field. If the location of the unique words and the beginning of the video lines or fields do not occur at the same point, then an error has occurred somewhere in the Huffman code data.

Control of a decoder FIFO buffer is used to maintain synchronization on a line and field basis. The FIFO controls work in conjunction with the unique words to reduce the impact of channel errors to merely streaks in a video line rather than a total loss of synchronization. In low error rate channels (less than 10^{-6}), the streaks caused by the errors may not be visible or bothersome to a viewer. This error rate corresponds to approximately one bit error per frame in NTSC transmissions. However, in a higher error rate channel (greater than 10^{-6}) the image degradation will be unacceptable. A better form of error detection/correction needs to be employed when using variable length codes like Huffman codes in a noisy communication channel.

VI   Algorithm Improvements

One approach for reducing the effects of channel errors on image quality is to introduce redundancy in the form of error checking bits. A high rate convolutional encoder coupled with a maximum likelihood decoder at the receiver would give a coding gain of 2-4 db. depending on the complexity of the decoder. Such

a system lends itself well to serial transmission and would reduce error rates in virtually all cases to acceptable viewing levels: however in addition to slightly larger bandwidth requirements, increased circuit complexity is one price paid for such a gain. For example, winner take all circuits, which are characteristic of maximum likelihood estimation, would have to be included. Also, neural networks have been shown to produce good non-linear estimates of maximum likelihood sequences [5].

Since the VQ algorithm performs so well in localizing errors to single pixels, the enhanced DPCM algorithm could be modified to take advantage of this characteristic. If the manner in which the NAP value was derived could be changed to an adaptive form, the difference value could (almost) always be driven to one of four quantization levels. If this were the case, then two bits per pixel could be sent, and the transmission would be very robust against channel disturbances. Unfortunately this method could introduce a further loss of information since we are limiting the possible values of DIF to a subset the actual values that it can attain.

Another approach to creating constant word length representations of difference values with the enhanced DPCM algorithm is to include more of the neighboring pixels in the prediction of the current pixel. In most cases this would improve the predicted value's estimate of the current pixel. The retina is able to change the size of the neighborhood depending on the activity of patterns within a given pixel's range. For example, in a field of constant pixel values, the neighborhood will expand to average in more of the surroundings; whereas in areas where there are rapidly changing pixel values, the neighborhood is shrunk, since values further away are not representative of the pixel of interest. Thus we would want to include some mechanism for varying the neighborhood size. Harris et al. [6]

accomplish this task through the use of resistive fuses in their hardware implementation of the retina. The resistive fuse acts as a resistor as long as voltage differences between pixels remain low; if they grow too large, the component becomes a fuse, effectively cutting off communication between pixels.

Another aspect of the retina is its ability to remove spatial distortion resulting from the changing media through which light signals pass. This distortion is analogous to the temporal smearing of a transmitted digital pulse as it passes through a network. Shown in Figure 4 is Lucky's decision directed transversal filter [7]. This filter removes temporal smearing of a digital signal by setting the weights so that they unconvolve the dispersive channel through which the signal passed. The incoming signal enters into a tapped analog delay line; the tapped values are weighted, summed and compared to some threshold value. The difference between the weighted sum and the binary comparator output forms the error value that drives the weight adapting algorithm. Figure 5 shows a modification to this idea so that it implements a spatial filter. Rather than being taps on an analog delay line, the incoming lines are the outputs from similar neighboring pixel filters. The weighted sum of neighboring pixels forms a prediction of the center pixel X. When the prediction matches the actual value, adaptation of the weights will stop. The modified spatial filter uses adaptive means to unconvolve any spatial smearing. Networks similar to this spatial filter are also useful for storing vectors in an associative memory [8].

Figure 4: Decision Directed Filter

5

Figure 5: Modified Spatial Filter

VII Summary

Digital compression of video images does indeed seem to be a viable means of transmitting HDTV signals. We examined two compression algorithms and made some comparisons to the human vision system. Vector quantization gives superior error correction performance due to the fixed length information block it sends. The enhanced DPCM algorithm compresses images more like the retinal image mapping and produces high quality reconstructions even near edges and other irregular features. The goal for HDTV of course is to produce the highest quality picture given the allotted bandwidth while still being robust against channel disturbances. Neural circuitry can help to block up the data output from the DPCM algorithm into more regular sized "chunks", thereby incorporating the advantages of both techniques into a single process.

VIII Acknowledgement

References

[1] R. Kaul. S. Bibyk. M. Ismail, and M. Andro, "Adaptive filtering using neural network integrated circuits," in *IEEE Symposium on Circuits and Systems*, May 1990.

[2] S. C. Ahalt, P. Chen. and A. K. Krishnamurthy, "Performance analysis of two image vector quantization techniques." *International Conference on Neural Networks*. vol. 1, pp. 169 – 175, 1989.

[3] M. J. Shalkhauser and W. A. Whyte, Jr., "Digital codec for real-time processing of broadcast quality video signals at 1.8 bits/pixel," in *IEEE Global Telecommunications Conference*, 1989.

[4] D. A. Huffman, "A method for the construction of minimum redundancy codes," in *Proceedings of the IRE*, pp. 1098 – 1101, vol. 40 1952.

[5] J. D. Provence, "Neural network implementation for an adaptive maximum-likelihood receiver," in *IEEE Symposium on Circuits and Systems*, pp. 2381 – 2385, May 1988.

[6] C. Mead and M. Ismail, eds., *Analog VLSI Implementation of Neural Systems.* ch. 2 Resistive Fuses: Analog Hardware for Detecting Discontinuities in Early Vision, pp. 27 – 55. Kluwer Academic Publishers, 1989. J. Harris, C. Koch, J. Luo, and J. Wyatt, authors.

[7] R. W. Lucky, "Automatic equalization for digital communication," *Bell System Technical Journal*, vol. 44, pp. 547 – 588, Apr. 1965.

[8] J. S. Denker, "Neural network models of learning and adaption," *Physica*, vol. 22D, pp. 216 – 232, 1986.

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE May 1990	3. REPORT TYPE AND DATES COVERED Conference Publication

4. TITLE AND SUBTITLE

Digital Compression Algorithms for HDTV Transmission

5. FUNDING NUMBERS

650-60-21

6. AUTHOR(S)

Kenneth C. Adkins
Mary Jo Shalkhauser
Steven B. Bibyk

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Department of Electrical Engineering
The Ohio State University
Columbus, Ohio 43210

8. PERFORMING ORGANIZATION REPORT NUMBER

P. 7

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

NASA Lewis Research Center
Cleveland, Ohio 44135

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Grant Contract Number NAG3-1082

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Unclassified - unlimited

12b. DISTRIBUTION CODE

STAR Category 32

13. ABSTRACT (Maximum 200 words)

Digital compression of video images is a possible avenue for HDTV transmission. Compression needs to be optimized while picture quality remains high. Two techniques for compressing digital images are explained and comparisons are drawn between the human vision system and artificial compression techniques. Suggestions for improving compression algorithms through the use of neural and analog circuitry are given

14. SUBJECT TERMS

HDTV
Video
DPCM
Data Compression
Vector Quantization
Human Vision System
Associative Memory

15. NUMBER OF PAGES

6

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102