USING AN INACCURATE MODEL FOR EFFICIENT COMMUNICATION OF A SONAR IMAGE

W J Mullarkey (1) and R C Chivers (2)

(1) DB Electronics Limited, Systems House, 62-68 Strand Road, Bootle, Merseyside L20 4BG
(2) Department of Physics, University of Surrey, Guildford, Surrey GU2 5XH

1. INTRODUCTION

The Concise Oxford Dictionary defines information as "knowledge, items of knowledge". Modern scientists, who work within information theory (after HARTLEY and SHANNON [1 & 2], would probably define it as the logarithm of the probability of some event.

When designing schemes to communicate information, we should remember that the commodity we handle is described by the first definition, even though we may use the second one to help in the design process. Shannon made it very clear that his definition was a mathematical convenience, not a philosophical truth. To clarify discussion we will allow the word *information* its dictionary definition and define *data* as being that which represents it in the communications channel. In practice the information will have been transformed by a sensor and will have some smallest part which we can resolve as having its own existence. We call that smallest part an *element of the information.*

Two of the main parameters, relevant to the communication of information from one place to another, are the bandwidth and the dynamic range of the communications channel. For electromagnetic channels, increasing either significantly can incur severe financial penalties. For some other types of communication channels (such as the underwater acoustic one) they are determined by uncontrollable physical constraints: extra quality cannot be bought at any price.

In order to optimise the transfer of information, advantage can be taken of the statistical properties of and redundancy in the information, so as to reduce the amount of data which needs to be transmitted. In this context, *redundancy* refers to a part of the information which is considered to be unnecessary or unusable.

There has been considerable research into the data compression of electronically processed and displayed images since the early work of HARRISON [3]. However, most of the published results have been concerned with television. The papers by SCHAMEL [4] and CONSIDINE et al [5] are typical: the review by JAIN [6] is instructive.

Recent developments in a number of technological areas have led to the emergence and exploitation of imaging techniques which are based on different types of radiation [7]. The schemes which may have been developed for the efficient transmission of television images may not be appropriate or suitable for such "instrument images".

2. ANALYTICAL CONTEXT

2.1 The requirement for accuracy in the transmission of an image.
Any communication of an image from a source observer to a receiving observer, via some communications channel, needs to take account of both technical and non-technical aspects of imaging. Fundamental to this is the fact that interpretation depends upon the conceptual framework of the observer. Specifically a source observer, seeking to reduce the amount of information to be transferred may discard that which he or she regards as unimportant. This irretrievable rejection of information which is considered to be redundant is called *data reduction*. The problem is that, because the observers have different conceptual frameworks, they may have different opinions about what is and is not important about the image. Any data reduction, performed by the source observer may result in loss of real information vital to the receiving observer. Similarly, the engineer who provides the communications process cannot assume that the two conceptual frameworks are identical. Therefore, he should not remove, prior to the transmission, any part of the information which Shannon would have accepted as redundant, unless he can re-insert it after reception.

Instrumented images, of the kind produced by sonar, radar or medical equipment are essentially different from television pictures. Even without making judgements about the relative consequences of misinterpretation, there is a clear distinction between images which, by definition, will only be observed by the human eye and those which may be collaboratively observed by human being and computer. It is therefore essential that whatever processing is performed on an instrument image should result eventually in its faithful communication to the receiving observer.

In the present discussion it will be assumed that the communication channel is noiseless so as to concentrate on the effects of processing. For the same reason, it will be assumed that the transducers used at each end of the channel have been engineered to cause no intrinsic loss of information. Reducing the number of data bits required to represent the information can ultimately aid performance in noise by allowing additional error correcting bits to be added to the data to give a more robust code.

In order to communicate an image with complete accuracy it is clearly possible to transmit the state, s, (where s is one of S possible states) of each pixel (picture element or image point). If there are n points this will require $n\log_2 S$ bits. If we are to reduce this figure there are two distinct approaches - efficient encoding or data compression.

Although both terms feature strongly in the appropriate literature, they are used in a wide variety of ways. To clarify the argument, we will define *efficient encoding* (in the sense of statistical encoding) as the class of processes in which the occurrence probability of a given state of a particular pixel (or ensemble of pixels) is mapped to the length of the code which is used to represent it. The classic example of this is Huffman Coding [8]. *Data compression*, on the other hand, describes a class of processes which identifies notional redundancy (in the sense of "unnecessary" rather than "unusable") so as to reduce the amount of data which needs to be transmitted.

It will be observed that both of these definitions suggest a lack of accuracy in transmission, by the use of the words "probability" and "notional redundancy". In order to clarify this it is necessary to formalise a procedure which is used in most, if not all, schemes of image compression: that procedure is often implicit rather than explicit. The formalisation has the advantage not only of clarifying the discussions of the procedures but, more importantly, of permitting them to be extended in a systematic way. The key to the formalisation is the existence of a model of the information contained in the image: that model is necessarily inaccurate.

2.2 Inaccurate models and their relationship to image transmission

The term *Inaccurate Model (IM)* includes both qualitative and quantitative aspects. It is the basis on which redundancy in the information comprising the image is defined. Indeed, the instantaneous accuracy of the model is a measure of the instantaneous level of redundancy. It is assumed that, if data compression is used, the receiver is aware of the inaccurate model so that the redundancy can be reinserted after reception of the image so as to provide a faithful reproduction of the original. Specifically the *IM* describes the state of an element of the information as a function of the states of some other elements. It has a specific probability of being correct which we call the *conformance probability*.

In the process of data compression, the actual image is compared with the *IM* (which, in this case acts as a "predictor" [9] and only the differences are transmitted. The more accurate the model the fewer data bits need to be transmitted. A useful figure of merit is the number of bits which would have been transmitted ($n\log_2 S$) without compression, divided by the number required after the redundancy has been stripped away.

Efficient encoding, in contrast, requires the whole of the information (including parts which may be considered to be redundant) to be transmitted but the number of data bits used to represent it is minimised. The design process uses *a priori* knowledge of the occurrence probability of the different states of the elements (or ensembles of elements) by allocating shorter codes to the more common states. The receiver clearly needs to know the code, but not the inaccurate model which was used to design it. The fact that this model of the information includes a probability distribution implies essentially that it is not an accurate description of a particular image and is therefore an "inaccurate model". However, the receiver has no need to know what the model is.

Whilst we distinguish between data compression and efficient encoding, it is clearly possible first to use the inaccurate model to reduce redundancy so as to achieve data compression and then to use it again so as to encode efficiently the data which remain. It is also possible first to carry out efficient encoding and then look for redundancy in the resulting code.

One very important advantage of the concept of the inaccurate model is that it permits the use of qualitative *a priori* knowledge. A good example is the transmission of a sonar image, where different approaches can be seen to be indicated depending upon whether the source observer has only the image in front of him from which to construct the model, or whether he knows that it is a sonar image produced in some given way.

If we are to consider only pure data compression, we must ensure that, even if the model is totally inaccurate, the communications process will still reproduce the image with 100% integrity. In such a case, it may take a long time, because at least the $n\log_2 S$ data bits will be transmitted, but information will be conserved.

3. APPLICATION TO SONAR IMAGES

3.1. The nature of the image

Consider the image which represents the information about a seabed and the water column above it, as transformed by an echo sounder.

The image is a Cartesian display having co-ordinates x,y,t. A single pixel (which is the transformed element, e) is labelled by its co-ordinates and has some colour or value on a grey scale. That colour/grey scale value will be one state, s, which is a member of a finite set containing S states. Figure 1 illustrates.

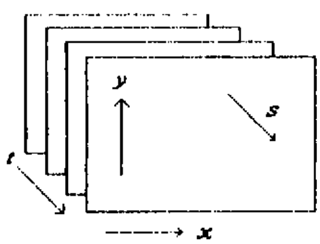


Figure 1.

For a practical system, typical values are:-
$0 \leq x \leq 511$
$0 \leq y \leq 255$
$1 \leq S \leq 4$
$s \in S$.

At each transmit pulse data are written to the right hand edge of the screen and all existing pixels move one value of x to the left. For the purpose of analysis, x,y, and t are integers: t relates to the current image and t-1 to the previous one.

$$s_e(x,y,t) = s_e(x+1,y,t-1) \quad \text{for:-} \quad 0 \leq x \leq x_{max}-1, \quad 0 \leq y \leq y_{max} \qquad \text{(Equn. 1)}$$

The new data are written to the column x_{max}.

3.2 A "conventional" approach to data compression of the image

One of the standard procedures for the data compression of images is interframe comparison using movement vectors [10]. This might produce a good general purpose scheme to identify redundancy in the image. It would use (possibly implicitly) the inaccurate model:-

$$s_e(x,y,t) = s_e(x+1,y,t-1) \bmod x_{max} \qquad \text{(Equn. 2)}$$

The movement vector would then be $x,t \rightarrow x+1,\ t-1$.

Note that Equn. 2 defines the state s, of an element $e(x,y,t)$, when the model is accurate. A fuller treatment, not needed for this discussion, would also require us to establish the conformance probability and to consider what happens when Equn. 2 is false.

It can be shown, given that the information is to be communicated (subject to noise) with 100% integrity, that

$$D = n(1-p_m)\log_2\left\{\frac{S-1}{1-p_m}\right\} \quad \text{Bits} \qquad \text{(Equn. 3)}$$

Where D is the minimum number of data bits required in the channel, n is the total number of elements (pixels in this case) and p_m $(0 \leq p_m \leq 1)$ is the conformance probability of the inaccurate model used as a predictor.

For a sonar image, using the Equn. 2 model, $p_m \geq 511/512$, so that $D > 2710$ bits.

Of this, only $256 \times \log_2(4-1) = 405.7$ bits are required to communicate the state of the elements; the remainder is an addressing overhead. It is generally true that the better the inaccurate model (that is the greater is its p_m), the greater is the proportion of addressing overhead in the transmitted data.

For this inaccurate model, a figure of merit, related to the "compression ratio" [11] is the "Quality of the model" given by:

$$Q_m = \left\{ \frac{(x_{max}+1)(y_{max}+1)\ \log_2 S}{n(1-p_m)\ \log_2[(S-1)/(1-p_m)]} \right\} \approx 96.7 \qquad \text{(Equn. 4)}$$

3.3. The use of an inaccurate model of the transformed information to design a data compression scheme for a sonar image

The approach described in 3.2 considers the image as the source of information: the inaccurate model does not take cognisance of what information the image represents or the mechanisms of its production. Clearly a two dimensional predictor [12] would do better if there is redundancy in more than one dimension. Any such improvement would increase the algorithmic complexity.

However, if we were to trace backwards from the image to the information it represents, and include our knowledge of that information and of the characteristics of the devices which transformed it, we can greatly improve on the accuracy of the inaccurate model. One improved model might use the information given by Equn. 2 but without the "mod x_{max}" condition; we can, however, do much better. We know from our *a priori* knowledge that the probability (p_m) of the first 511 columns being correct has $p_m = 1$. Thus we know that we need only consider inaccuracies in the final column $x = x_{max}$.

It would be possible to use an inaccurate model to search for redundancy in this column and the use of

$$s_e(x_{max}, y, t) = s_e(x_{max}-1, y, t-1) \qquad \text{(Equn. 5)}$$

would be the obvious choice. However, it is not clear that this would have significant advantages over a procedure which used efficient encoding to transmit the whole of the final column.

As a means of achieving this, we could use Equn. 2 for all values of y and t combined with $0 \leq x \leq x_{max}-1$; and

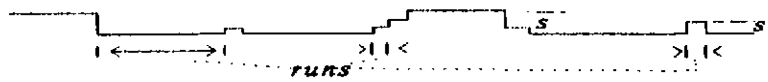
$$s_e(x_{max}, y, t) = s_e(x_{max}, y-i, t-1), \qquad \text{(Equn. 6)}$$

for all values of y, i and t combined with x_{max}. Given that $1 \leq i \leq \ell$, where ℓ is the run length of a given state. The model would further recognise that all values of ℓ are not equally probable.

The run length code needs ℓ discrete messages to indicate length, S messages to indicate the state of the first run and S-1 for all subsequent runs. This latter is because the current state is, by definition, different from the proceeding one. Figure 2 illustrates this.

USING AN INACCURATE MODEL FOR COMMUNICATION

Figure 2.



The code thus contains $(S\text{-}1)\times\ell$ messages and needs a minimum of

$$D = \log_2 \ell(S-1) \text{ Bits} \qquad \text{(Equn. 7)}$$

However, whilst Equn. 7 may give the absolute minimum required, it suffers from two practical problems. Since the code is of variable length, loss of synchronisation will cause an exaggerated loss of information in the event of noise [13], and it would need a different procedure for the first run (to include S rather than S-1 messages).

A practical code would use a fixed $\log_2 (\ell_{max})$ bits, rather than the variable $\log_2 \ell$. It would also always allow for S states rather than S-1. The latter is required from the former because if, for some given run, $\ell_{max} < \ell$ then the next run would have the same state. The final part of the inaccurate model is then simply $\ell - \ell_{max}$.

For any run ℓ, where $\ell \leq \ell_{max}$, the bits required to encode the run length are

$$D = \log_2 \ell_{max} S \quad \text{Bits} \qquad \text{(Equn. 8)}$$

If $\ell > \ell_{max}$ then

$$D' = \left[\left(\mathcal{J}\,\frac{\ell_{max}}{\ell}\right) + \mathcal{P}\right] \log_2 \ell_{max} S \quad \text{Bits} \qquad \text{(Equn. 9)}$$

Where $\mathcal{J}\frac{\ell_{max}}{\ell}$ is the integer portion of $\frac{\ell_{max}}{\ell}$. $\mathcal{P} = 0$ if $\ell_{max} \bmod \ell = 0$, and
$\mathcal{P} = 1$ if $\ell_{max} \bmod \ell \neq 0$.

The quality of the model, Q_m, in so far as it applies to the communication of the single run, can be defined as

$$Q_m = \frac{\ell \log_2 S}{\left[\left(\mathcal{J}\,\frac{\ell_{max}}{\ell}\right) + \mathcal{P}\right] \log_2 \ell_{max} S} \qquad \text{(Equn. 10)}$$

Where $\ell \log_2 S$ is the number of bits which would have been required if the states of elements within the run had been independently encoded.

Equations 8, 9 and 10 describe single runs. The data set:- $s_e(x_{max}, y = 0$ to $y = y_{max}, t)$ will consist of m runs, the last of which will terminate at $y_{max}+1$. Then

$$\sum_{i=1}^{m} \ell_i = y_{max} + 1 \qquad \text{(Equn. 11)}$$

The equations for the whole column, which correspond to Equations 8 and 9 are:

$$D = \sum_{i=1}^{m} \log_2 \ell_{max} . S \qquad \text{(Equn. 12)}$$

and

$$D' = \sum_{i=1}^{m} \left[\left(\mathcal{J} \frac{\ell_{max}}{\ell_i} \right) + \mathcal{L} \right] \log_2 \ell_{max} . S \qquad \text{(Equn. 13)}$$

respectively. Whilst the compression ratio of the whole image (cf Equn. 10) can be calculated from

$$Q_m = \frac{(y_{max}+1)(x_{max}+1)\log_2 S}{\sum_{i=1}^{m} \left[\left(\mathcal{J} \frac{\ell_{max}}{\ell_i} \right) + \mathcal{L} \right] \log_2 \ell_{max} . S} \qquad \text{(Equn. 14)}$$

Consider Figure 3, which is a tracing made from a photograph of the screen of a hydrographic sounder.

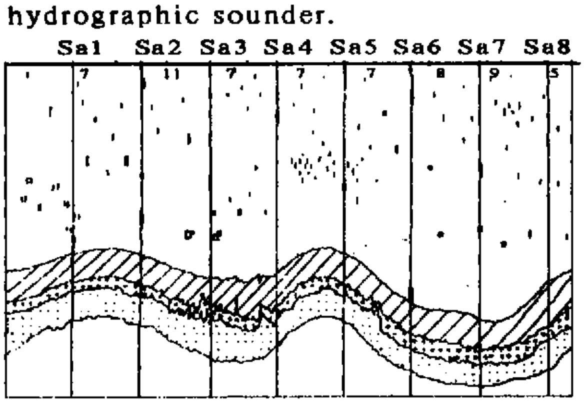


Figure 3

By way of experiment, we have taken eight monotonic time samples (Sa1 to Sa8) across the image: they give an average number of run length of 7.75 with a standard deviation of 1.83. For such a practical machine, $y_{max} = 255$, $x_{max} = 511$ and $S = 4$. If we use an eight bit code, then the $S = 4$ needs two bits and we have a convenient $\ell_{max} = 64$. For safety's sake let us estimate 12 runs per column.

From Equn. 13, we can calculate the number of bits needed to transmit the information contained in the image as $D' = 96$ bits. From Equn. 14, we can calculate the compression ratio (or quality of the model) as $Q_m = 2731$.

Equn. 11 can be used to check easily for errors. It is a necessary (but not sufficient) condition that there is an integer solution for "m".

4. CONCLUSION

The subject of data compression of dynamic images is currently receiving considerable attention. The figure of merit normally used is "bits per pixel" (bpp). For images with an eight bit grey scale current achievements are typically 1.0 bpp and not less than 0.1 bpp. The first (conventional) procedure suggested for analysing and compressing sonar images gives a result of 0.021 bpp, assuming a perfect movement vector. Bearing in mind that this is for a two bit grey scale, we would expect $0.021 \times (8/2)^2 \approx 0.33$ for a corresponding eight bit image, from the same inaccurate model. It is the particular structure of the sonar image which causes the figure to be so low.

Formal introduction, analysis, and application of the concept of an inaccurate model, not of the image but of the information and how it was transformed by the sonar to produce the image, has a profound effect on the efficiency of the schemes we can design. In other words, a model which extends to include exact *a priori* knowledge of the information source and mechanisms involved in image production, dramatically increases our ability to find redundancy for individual sources and mechanisms. In particular it enables an inaccurate model to be constructed which will make possible a quantifiable improvement on the conventional approach. In the particular case chosen, the compression ratio is improved by a factor of over 28. This results in a figure of 0.0007 bpp, corresponding to 0.01 bpp for an eight bit image!

The inclusion of noise leads to some additional, but manageable, complexity in the calculation of the various figures of merit for different schemes. Detailed analysis of the effects of noise, from the viewpoint of additional errors introduced into the inaccurate model, provides scope for further research.

It might be argued that the second approach is one which any sensible engineer would have adopted: it could then be countered that, if he were given only the image, he would not know that he always could. An engineer who did adopt the second approach could only do so by using *a priori* knowledge. Accepting, then, that he would be implicitly using an inaccurate model: it is proposed that he should do so explicitly and analytically.

Clearly sonar images have turned out to be particularly suitable candidates for analysis by the *IM* approach, even though the authors did not know that until the analysis had begun. The way that the images are generated permits the construction of an inaccurate model with a very high conformance probability. The fact that the performance of schemes, which the same approach might indicate for other kinds of images, are unlikely to be so good, should not be allowed to detract from the significance of the approach. It is probable that in most, if not all, cases, it will offer a means for assessing and improving image compression procedures.

The particular advantage of the novel, formal description of the conceptual framework which includes an *IM* is that it permits one to identify the strategies for image compression that are likely to lead to the greatest improvement. It is then relatively straightforward to quantify the improvement which can be expected.

5. REFERENCES

[1] R V L HARTLEY, 'Transmission of information', BSTJ **7** pp535-539 (1928)

[2] C E SHANNON, 'A mathematical theory of communications', BSTJ **27** pp379-623 (1948)

[3] C W HARRISON, 'Experiments with linear prediction in television', BSTJ **31** pp764-783 (1952)

[4] G SCHAMEL, 'Three dimensional sub-band coding of HDTV signals with 140 Mbits/s', in Conf. Proc. of 3rd Int. Conf. on image processing, conf. pub. **307** (IEE) pp270-274 (July 1989)

[5] V CONSIDINE, A S BHANDEL and J N GOODING, 'Single chip motion estimator for video CODEC applications', in Conf. Proc. of 3rd Int. Conf. on image processing, conf. pub. **307** (IEE) pp285-289 (July 1989)

[6] A K JAIN, 'Image data compression - a review', Proc. IEEE **69** No.3 pp349-389 (1981)

[7] CONFERENCE PROCEEDINGS, 3rd Int.Conf.on image processing, Conf. pub. **307** (IEE), 142 papers U.of Warwick (July 1989)

[8] D A HUFFMAN, 'A method for the construction of minimum length codes', Proc. IRE **40** pp1098-1101 (1952)

[9] P ELIAS, *'Predictive coding'*, PhD Thesis, Harvard University Press, Cambridge Mass. (May 1950)

[10] T KOGA et al., 'Motion compensated interframe coding for teleconferencing', Proc. Nat. Telecom. Conf. **4** G5-3 (1981)

[11] G HELD, *'Data Compression'*, John Wiley and Sons, London ISBN 0-471-26248 (1983)

[12] A LEMPEL and J ZIV, 'Compression of two dimensional data', IEEE Trans.on Info. Theory **IT32** No.1 pp2-8 (Jan 1986)

[13] J C MAXTED and J P TOBINSON, 'Error recovery for variable length codes', IEEE Trans. on Info. Theory **IT31** No.6 pp794-800 (Nov 1985)