

NASA Technical Memorandum 105239
AIAA-91-3533

1N-32
40127
p11

Data Compression for Full Motion Video Transmission

(NASA-TM-105239) DATA COMPRESSION FOR FULL
MOTION VIDEO TRANSMISSION (NASA) 11 p
CSCL 17R

N91-32247

G3/32 Unclass
0040127

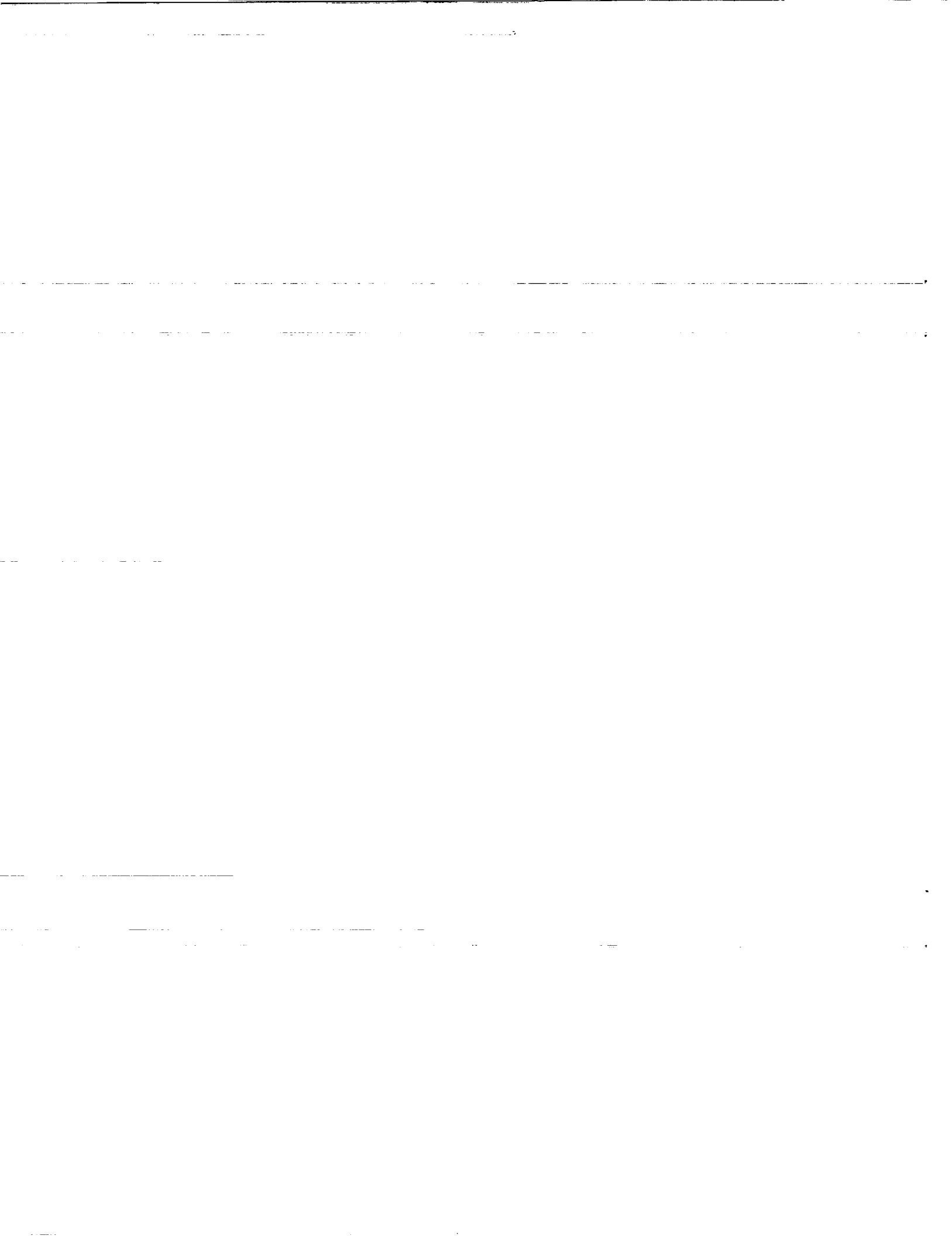
Wayne A. Whyte, Jr.
Lewis Research Center
Cleveland, Ohio

and

Khalid Sayood
University of Nebraska-Lincoln
Lincoln, Nebraska

Prepared for the
Conference on Advanced Space Exploration Initiative Technologies
cosponsored by the AIAA, NASA, and OAI
Cleveland, Ohio, September 4-6, 1991

NASA



DATA COMPRESSION FOR FULL MOTION VIDEO TRANSMISSION

Wayne A. Whyte, Jr.
NASA Lewis Research Center
Cleveland, Ohio

Khalid Sayood
University of Nebraska-Lincoln
Lincoln, Nebraska

Abstract

Clearly transmission of visual information will be a major, if not dominant, factor in determining the requirements for, and assessing the performance of, the SEI communications systems. Projected image/video requirements which are currently anticipated for SEI mission scenarios are presented. Based on this information and projected link performance figures, the image/video data compression requirements which would allow link closure are identified. Finally several approaches which could satisfy some of the compression requirements are presented and possible future approaches which show promise for more substantial compression performance improvement are discussed.

1.0 Introduction

Image/video data compression has been identified as a critical technology development element, needed to enhance throughput for data rate constrained lunar/Mars communications links, as part of the high rate communications program element of the Space Exploration Initiative (SEI). Technology assessment studies included in the 1989 Mission Analysis and Systems Engineering Final Report¹ identified uncompressed video and image data rate requirements in the range several hundreds of megabits per second (Mbps) to gigabits per second (Gbps). Additional studies² have shown that approximately 10 Mbps may be the maximum return rate available from Mars, at least for initial Mars-Earth communications links. The discrepancy between required and available data rates can be addressed through at least three means. One method is the use of advanced high-order modulation schemes, such as 8-ary or 16-ary PSK, which can allow transmission at multiple bits/second/Hz. These modulation schemes generally gain bandwidth efficiency at the expense of increased transmitter power. Lunar links could benefit from such schemes where ample margin exists to permit link closure. The Mars links, however, are power limited and therefore would not benefit from bandwidth efficient modulation. (Recent research in combined modulation and coding schemes³ shows promise of providing both bandwidth and power

efficiency improvements. Further study is needed to determine the performance of combined modulation/coding schemes on an already coded channel.) A second means of closing the gap between required and available data rates is mass data storage. High rate data can be buffered for subsequent transmission at a lower transmission rate. Mass data storage will undoubtedly be required to provide buffering during connectivity outages and periods when data volume exceeds the transmission capabilities of the communications system. To minimize data storage requirements, however, a third method for reducing the discrepancy between required and available data rates will need to be applied. This third method is data compression. Whether it is digital data, voice signals or image/video, data compression will be used to reduce the rate required to transmit the information.

This paper addresses data compression as applied to image and video information. Clearly transmission of visual information will be a major, if not dominant, factor in determining the requirements for, and assessing the performance of, the SEI communications systems. Section 2.0 examines the projected image/video requirements currently anticipated for SEI mission scenarios. Based on this information and projected link performance figures, section 3.0 identifies the image/video data compression requirements which would allow link closure. Finally section 4.0 presents several approaches which could satisfy some of the compression requirements and discuss possible future approaches which show promise for more substantial compression performance improvement.

2.0 SEI Image/Video Requirements

NASA studies have identified various lunar/Mars mission requirements that involve transmission of image/video data. In general these requirements were kept very austere in recognition of limited data return rates available from Mars. Image/video data is categorized into several data types; high rate video, edited high rate video, low rate video, science imaging data, and telerobotic video. A brief description of each data type follows.

High Rate Video: Proposed manned Mars mission scenarios involve long transit times and extended duration remote base station habitation. For the sociological and psychological benefit of the crew it is very desirable to provide two-way voice/video communication to mission operations personnel, relatives and friends. It is additionally desirable (probably essential) to provide entertainment video, news and video-based training to the crew during long duration missions. The news media and general public will require "live" video back from the Moon and Mars, as well, for educational benefit as well as to foster continued public support. Such video will need to be of suitably high quality.

High quality, full motion color video of a type similar to standard NTSC (National Television Systems Committee) video is needed to fulfill these requirements. Uncompressed video of this type would require a transmission rate of approximately 100 Mbps (megabits per second). It must be recognized, however, that over the next 20 years video will evolve from NTSC to HDTV (high definition television) with a corresponding evolution of viewer expectations. At the time of the first manned Mars missions, NTSC quality may no longer be acceptable within the mission scenarios, having been replaced by the need for HDTV quality. Uncompressed HDTV would require transmission rates around 1 Gbps. Video compression is needed to reduce these uncompressed rates to allow transmission within the available channel rates.

Edited High Rate Video: For applications such as remote monitoring or video mail a lower quality signal, comparable to video teleconference quality, will be appropriate. Such services currently require 1-2 Mbps and typically achieve these rates through editing (dropping) certain frames. The consequence of dropping the frames is reduced motion rendition. While these rates could be accommodated within the projected channel rates, additional compression research could allow more efficient utilization of the limited communications resources.

Low Rate Video: In some applications, such as occasional monitoring, a very low rate video signal will be sufficient. This signal type has been identified as being monochromatic with a resolution of 512 X 512 picture elements (pixels) at 8 bits per pixel. The frame rate is identified as 1 frame per second. The uncompressed data rate for such a signal is approximately 2 Mbps.

Science Imaging Data: Science imaging data requirements vary greatly. Far side lunar astrophysics experiments may require multiple gigabits per second

due to very high resolution and frame rates. In spite of the substantial channel rates proposed for the lunar links (350 Mbps), requirements of this type would undoubtedly overload the communications capabilities without some form of data compression and data buffering. Fortunately, scientists have indicated that experiments of this nature are best carried out on the Moon and are not currently being considered as part of the Mars mission scenarios where channel rates are even far more restricted. Imaging instrumentation packages for Mars missions will be used to conduct sample analysis, produce spectral plots, etc. and might produce one high resolution color image each minute (1024 X 1024 pixels, 8 bits per RGB color component). This would result in a data rate of approximately 0.5 Mbps.

Scientists believe that every single bit of information gathered is potentially invaluable and possibly irreplaceable. They are therefore very reluctant to consider any sort of compression being applied to their scientific data. It is likely then that only lossless (fully reversible) data compression techniques will be usable for most, if not all, science data.

Telerobotics Video: Telerobotics video usually involves stereoscopic video for depth identification. An unmanned rover needs to be able to negotiate a boulder field either autonomously or via remote control. To accomplish this task, distance information must be derived from a stereo image pair in much the same way that our brains analyze the image pairs received by our two eyes. Two high rate (i.e. 30 frame per second) color video channels are needed for telerobotics video. For resolutions of 512 X 512, two channels would require an uncompressed rate of 200 Mbps. Higher resolution image data (up to 2048 X 2048) may be desirable for some applications, which would drive the data rate requirements into the gigabit per second range.

3.0 Image/Video Data Compression Requirements

Several NASA studies have examined the channel data rates which can be supported for lunar and Mars missions. For both lunar and Mars mission scenarios the space-to-Earth links dominate the communications system requirements. This is due to the path length on these links and the attenuation through the Earth's atmosphere. Earth-to-space links are not as severely restricted because Earth-based transmitters are far less power restricted than space-based transmitters. Lunar links are also far less restricted than Mars links due to the enormous difference in path lengths. The Moon is approximately 405,000 km from Earth while Mars and Earth are approximately 2.5 AU (3.74×10^8 km) apart

at their maximum separation. Reasonable antenna sizes and transmitter powers can accommodate a 350 Mbps return link from the Moon at Ka-band (1 Gbps return data rate is not out of the question). For the Mars link, even when stressing antenna and transmitter parameters to reasonable limits, only about 10 Mbps can be supported on the return link at Ka-band. Development of optical technology may allow Mars return data rates as high as 100 Mbps, but such development is far into the future. Operational scenarios are currently being examined with a 10 Mbps maximum return rate from Mars.

Transmission of video/image information is clearly a major driver in setting the transmission requirements. Most, if not all other data transmission requires significantly less bandwidth than picture data. (Some non-imaging scientific instruments also produce large quantities of data. This data, however, can generally be more readily buffered and transmitted in a "non real time" manner.) Video/image data compression is therefore required for efficient information management in the lunar and Mars exploration missions. Data compression is essential for transmission of high rate video data on the Mars links. Even for data types such as low rate video, which could be accommodated within the available channel capacity, transmission efficiency could be substantially enhanced through use of efficient data compression techniques, thereby allowing maximum utilization of the available capacity.

The data compression requirements can only be considered as minimum requirements since any additional compression translates directly to increased communication system capability. In other words, if a 10:1 compression factor would allow a single channel to be transmitted within the available channel capacity, 20:1 compression providing equal quality would double the effective capacity of the communications channel. Table 1 presents the various image/video data types discussed in section 2.0. The table lists the uncompressed data rates along with the minimum required compressed data rates for each data type. In general, minimum requirements call for 10:1 compression for most data types with the exception of high resolution video which requires significantly greater compression. Current lossy compression techniques achieve reasonably good quality at 10:1 compression ratios using a combination of spacial and temporal processing. Higher compression ratios are achievable at reduced quality in spacial resolution and motion rendition. Additional research and development is needed to improve the quality at the higher compression ratios.

4.0 Compression

In this section we will discuss various image compression schemes which could be used in addressing the Space Exploration Initiative compression requirements discussed previously. Image compression schemes can be classified as lossless (invertible, noiseless), or lossy (non-invertible). As implied by their name, lossless coding schemes provide a compressed representation which can be inverted to obtain a reconstructed image which is identical to the original. In case of the lossy coding schemes, while the reconstructed images may look identical to the original, the pixel values of the reconstructed image are not identical to the pixel values of the original image (in cases of high compression the reconstructed image may look appreciably different than the original). The selection of which type of compression scheme is to be used depends on a number of factors, chief among which are, bandwidth availability, and user acceptance. The lossless coding schemes generally require significantly larger bandwidth than lossy coding schemes, however, user insistence might dictate the rejection of any coding scheme which throws away any information. In the following sections we will discuss both lossy and lossless coding schemes.

We will also examine these schemes from the viewpoint of implementation. Whether they can be implemented in the short term, or whether their implementation depends on technology which can realistically be expected to appear within the next decades. Development of all compression schemes can, in a sense be divided into two parts: the development of a model for the source output, and a coding scheme developed with reference to the model⁴. This is especially true for lossless coding schemes. The information to be transmitted to the receiver includes description of the model, and the information sequence coded with reference to the model (we will elaborate on this later in the paper). Whether a scheme can be implemented in the short term or will have to wait for further development of technology depends to a great extent on how complex a model is required. If a static model is to be used for all images, then this information can be made available to the receiver during initialization, and does not need to be transmitted. If an adaptive model is to be used (the model adapts to the data) then, first the model has to be extracted, developed at the transmitter, and sent to the receiver. If the model is complex, the level of technology required to extract it may be significantly higher than is currently feasible.

4.1 Lossless Image Compression Algorithms

The best a lossless compression algorithm can do is code at a rate equal to the entropy of the source

$$H(S) = \lim_{n \rightarrow \infty} G_n$$

where

$$G_n = -\sum P(X^n) \log P(X^n)$$

and X^n is a sequence of length n from the source. If the source is memoryless then

$$H(S) = -\sum P(X) \log P(X)$$

The estimate of the entropy depends on the model for the source sequence. Consider the following sequence

1 2 3 2 3 4 5 4 5 6 7 8 9 8 9 10

Assuming the frequency of occurrence of each number was reflected accurately in the number of times it appears in the sequence, the entropy for this sequence assuming a memoryless model would be 3.25 and the best scheme we could find for coding this sequence could only code it at 3.25 bits/sample. However if we assumed that there was sample to sample correlation between the samples and removed the correlation by taking differences of neighboring sample values, we would arrive at the sequence

1 1 1 -1 1 1 1 -1 1 1 1 1 -1 1 1

This sequence is constructed using only two values and its entropy is 0.70! Knowing only this sequence would not be sufficient for the receiver to reconstruct the original sequence. The receiver must also know the process by which this sequence was generated from the original sequence. Or, in other words, the receiver has to know the model being used. The model is a static one and the receiver can be informed about it on initialization. Therefore, the total cost is still 0.70. Consider now, the following sequence:

1 2 3 2 3 4 5 4 5 6 8 10 12 14 16 18 20 22 24 26

Now if we take differences, we would obtain the sequence:

1 1 1 -1 1 1 1 -1 1 1 2 2 2 2 2 2 2 2 2 2

Obviously, the model changed in the middle of transmission, and knowledge of this change would help decrease the coding rate. There are two processes

involved here, detecting and estimating the model change, and informing the receiver of the change. Taking the second process first; informing the receiver may involve the use of a "side" channel, and therefore increase the synchronization requirements. Detecting and estimating the model change, however, can be quite difficult, when the models in question are complex, and it might not be possible to do this in real time with current technology.

We present in the sequel two schemes, one which can be (and has been) implemented with current technology⁵, and the other which may require some advances in technology before being economically practical⁶.

That the pictures in almost all images are heavily correlated is rather obvious. Most lossless image coding schemes use a simple model similar to the one described above. That is, they take pixel to pixel differences, which are then coded using an entropy code. The model for the first scheme we describe is slightly more complex, but is still simple enough to function in real time.

The Differential Lossless Coding Scheme (DLCS) functions by comparing the current pixel (byte) value with a reference pixel to obtain a prefix and suffix value for each pixel in the image. The prefix and suffix values comprise the noiseless code for the pixel. The prefix value is the number of MSB (upper bits) in a byte that are identical to the reference pixel. For example:

reference pixel	=	1 1 0 1 0 1 1 0
current pixel	=	<u>1 1 0 1 1 0 1 0</u>
prefix value 4	=	(1 1 0 1)

Before being sent the prefix value is Huffman encoded. A given prefix value is assigned a predetermined Huffman code. The prefix value can range from zero to eight. The suffix consists of the bits of the current pixel that are not identical to the reference pixel minus the most significant bit (MSB) of the nonidentical bits. The MSB (of the nonidentical bits) is not sent because it is obviously the opposite of the reference pixel (otherwise it would be the same as the reference and be included in the prefix value). The actual data sent for each pixel is the Huffman code for the prefix value, and the suffix sent as is (bit for bit). In the previous example, if the Huffman code for 4 is 1 0 the code sent for the current pixel given would be 1 0 0 1 0. Due to the Huffman code and the fact that the suffix length is directly dependent on the value of the prefix, the compressed code sent is a variable length code. The next problem is to actually transfer the new code. Data is transferred in bytes (eight bits). Therefore, bits are placed into

bytes and transferred as soon as a byte is filled. The decoding is done by reading the bytes bit by bit. The bit(s) are matched against the Huffman codes to determine the prefix value. Once a match is found, that many upper bits of the reference pixel are set in the current pixel being decoded. Then the next bit (bit # = 7-prefix value) value is flipped, from that of the reference pixel. Then, according to the prefix value, the suffix bits are set. If the prefix value is four, then the suffix must contain three bits. For example, reversing the first example:

code sent	= 1 0 0 1 0
first bit compared	= 1
(no match)	
add bit, compare	= 1 0
(matches prefix = 4)	
if, reference pixel	= 1 1 0 1 0 1 1 0
set current pixel	= 1 1 0 1
flip next bit	= 1
set the next three	= 0 1 0
(7-4 bit suffix)	
current pixel	= 1 1 0 1 1 0 1 0

The next bit read from the code would be the start of the next prefix value. The very first pixel of every image is always sent as is and is always the first reference pixel. The first line always sets the reference pixel to be the previous pixel, to the left. For the first pixel on each line the reference pixel is always the pixel directly above the current pixel. These reference pixels are always true no matter how the rest of the image is referenced. The algorithm flips the reference pixel between above and to the left depending on a threshold value. The threshold value is set at the beginning of the program. If a prefix value drops below the threshold value, the reference pixel is switched (from above to left or vice versa). For example, if the reference pixel currently being used is to the left and the threshold value is three and the current prefix value is two, then for the next pixel, the reference pixel used will be above.

The algorithm was tested with a database containing nine images. Three of these images are from the USC Image Database, three were acquired using an IVG 128 frame grabber with a SONY CCD camera, and three are NTSC images. The USC images are of size 256 x 256, the IVG acquired images are of size 384 x 512, and the NTSC images are of size 768 x 512, thus there is an increase in resolution between the USC images, the IVG images, and the NTSC images. As we are generally interested in sending high resolution images, it is of interest to note how the system performs with increase in resolution. A value of 4 was used as the threshold value in the algorithm. The results were compared to the results obtained by using the commercially available

program PKARC, and to the theoretical best (entropy of differences). PKARC uses a total of six different coding schemes including Huffman coding and several versions of LZW. Table 2 shows the comparison between the various schemes. The performance measure used was percent compression, which is defined as

$$\%compression = \frac{R_o - R_c}{R_o} \times 100$$

where R_o is the number of bits in the original image and R_c is the number of bits in the compressed image. As can be seen from the results in the table the DLCS encoding scheme performs consistently better than PKARC, with compression very close to the theoretical maximum (for the simple difference model - note that for two of the NTSC images the DLCS scheme performs better than the entropy of the differences. This is probably because the adaptive nature of the DLCS algorithm provides a better model for the images than the static model used to compute the entropy). Furthermore, the performance tends to improve with improving resolution.

Notice that the model used here is relatively simple and therefore implementation at realtime rates is feasible. (A prototype version of this system has been implemented at the University of Nebraska-Lincoln.) Most scientists are reluctant to consider the use of compression of any kind for their data. However, the use of this algorithm could result in a doubling of the amount of science imaging data that can be transmitted, while maintaining complete integrity of the original data values. This algorithm could accommodate some of the SEI compression requirements, however even more efficient schemes must be developed to address the large volume of scientific data envisioned for the missions. To accommodate more of the requirements we have to look at techniques that are considerably more complex, with the hope that developments in technology will make the techniques feasible in the coming decades.

With this in mind, we next look at a system which uses a somewhat more complicated model for the image and provides correspondingly more compression. This approach is based on the idea that a top-down, left to right scan may not capture the structure in an image. There have been several attempts to generate scans that would capture more of the structure of the image. Or, in other words provide a better model for the image. Most of these efforts have been directed at constructing fixed scans of some sort. The same scan can then be used for each image. However images are sufficiently different that these efforts have not been of much use.

In this work we try to generate scans that are particular to the image. This of course means that we have to somehow inform the receiver about which scan we are using. More on that later. We start by representing the scan as a graph in which the pixels are the nodes and the weight on the edges are differences between pixels. If we consider the graph to be a directed graph then the weight on the edges need only be positive. If the edges are not directed edges, then the weights can be positive or negative. We can have two different types of graphs, a four-difference graph in which each node is connected to each of its vertical and horizontal neighbors, or an eight-difference graph in which each node is connected to its eight vertical, horizontal and diagonal neighbors. We can then obtain a spanning tree for the graph which is a possible scan of the image. Our goal is to find a model or scan, with respect to which the entropy of the image is a minimum. We would like to generate a scan based on spanning trees, which would provide the minimum entropy residuals. The generation of a minimum entropy spanning tree is most likely an intractable problem, so we have developed a set of heuristics which are reasonably efficient.

If we assume that the prediction errors are clustered around zero, then it seems reasonable to suspect that minimizing the sum of absolute errors will also tend to minimize the entropy. Given an image of size $M \times N$ a scan that minimizes the sum of absolute prediction errors can be obtained in time $O(MN \log MN)$ ⁷ and therefore is quite practical. We could also replace the errors with their frequency of occurrence on the graph, and compute the maximal spanning tree. This should also provide an approximation to the minimum entropy spanning tree. The maximal spanning tree can also be computed in time $O(MN \log MN)$. Finally we could simply use a greedy heuristic which uses the known information (past history) about the frequency of occurrence to progressively construct the tree.

Using the maximum scans on the test set we come up with the entropy values shown in Table 3. If we could code the images, with these rates, for most of the images this would mean a doubling of the percent compression. However we have not taken into account the rate required to code the model. How efficiently we encode the model will determine how much compression we finally get. To a large extent this depends on how much complexity we can handle. To reduce the complexity of the problem we could divide the image into smaller blocks and code these smaller blocks. In the third column in Table 3 we present the results of encoding 8×8 blocks using a codebook of trees containing 256 "codes" or spanning trees. While the results are better than those obtained using the DLCS, we are still quite far from the best achievable. Most of

this loss is due to breaking the image up into 8×8 blocks. Theoretically, if we spent just .25 bits per pixel we could store a codebook of 2^{16384} spanning trees, then when we wished to code an image we would simply send the label corresponding to that spanning tree. However, practically, with current technology this is not feasible. At the moment we cannot even simulate the effect of using a set of spanning trees of size 2^{16384} instead of all possible scanning trees on the rate. We are therefore limited by the current technology in our search for means of transmitting the spanning tree or an approximation without substantially increasing the overhead.

A technique which initially seemed to show a great deal of promise and generated quite a bit of publicity was the use of self-similarity to compress an image (fractals). While the previous technique defines the model as a scan, the fractal techniques concentrate more on some repetitious properties of the image. Though the fractal approach by itself has been somewhat of a disappointment, it is possible that as a class of models together with other classes of models, they may be useful in algorithms which provide significant compression. Such algorithms would view the source as being composed of several sub-sources. Each sub-source would be of a form that could be efficiently compressed by one of several approaches. Note that the switching information would have to be transmitted as side information, or derived from the transmitted data. This composite source approach requires the existence of a sophisticated segmentation/classification algorithm, which could alert the encoder as to which particular model was active at a given time. It would also require some somewhat complicated control logic, depending of course, on how complicated the source model is. Fractal techniques and other algorithms, which seem to work very well only on restricted classes of data would find use in such an approach.

4.2 Lossy Compression

While lossless compression is essential in applications where complete data integrity must be maintained, it is apparent that the amount of compression achievable is very limited. Fortunately, many of the SEI video/image compression applications do not have the requirement for reversible data recovery, and therefore lossy compression techniques can be considered. The lossy image compression area has seen much more activity in the last two decades. However it is more difficult to quantify the progress because of a lack of an accurate objective measure of performance. The objective measures used are generally, Signal-to-Noise Ratio (SNR) and its variants, compression ratios, and number

of bits per pixel. The SNR measure is known to not always correlate with perceptual evaluations. The compression ratio will vary depending on whether the original pixels were coded using eight bits or twenty four bits, or maybe even thirty two bits. And, finally all results may change if the algorithm is used on a test set different than the one presented. Subjective evaluations depend very much on the viewers personal experiences and biases. They are also difficult to verify from results in the published literature because the results are presented using small pictures which tend to mask errors. Recently, the concept of transparent coding has begun to gain some measure of acceptability, where an image is said to be coded in a transparent fashion if under specified viewing conditions, an observer looking at an uncoded and coded image side by side, takes more than 20 seconds to find a coding distortion⁸. Some standardization has begun to occur with the various international standard making bodies proposing standards at various rates (CCITT - H.261, CCITT/CCIR - CMTT/2, CCITT/ISO - JPEG, ISO - MPEG). Comparison with these standards will be a useful measure of the performance of any new algorithms.

At present, *transparent* schemes which are close to or undergoing implementation include a differential encoding system being developed at NASA LeRC for encoding NTSC (National Television Systems Committee) video with transparent quality (image size 768 X 512) at 15 to 20 Mbits/sec⁹, and the proposed MPEG standard (image size (512 X 486) at about 10 Mbits/sec¹⁰. These are described below.

The differential encoding system under development at NASA LeRC (patent pending) is an intrafield coding scheme designed to encode full motion NTSC video in a transparent manner. The technique is based upon a two dimensional predictive differential pulse code modulation scheme with data rate reduction enhancements. A non-uniform scalar quantizer in conjunction with multi-level variable length Huffman code sets provides significant increase in compression performance over conventional DPCM schemes without significant increase in implementation complexity. A non-adaptive predictor is used to reduce edge degradation, thereby improving the subjective quality of the reconstructed video image. No temporal processing is incorporated which allows perfect motion rendition.

The MPEG standard uses transform coding using the DCT and two different types of motion compensation to remove the temporal redundancy. There are three types of frames in the MPEG coded sequence, intraframes, predicted frames, and interpolated frames. Intraframes are transmitted at regular intervals, and are coded using

the DCT. No information from neighboring frames is used to reduce the redundancy in the frame. While this increases the bit rate, it allows for random access applications. The predicted frames are generated by coding the prediction error between that frame and a motion compensated prediction based on the previous frame. The interpolated frames are generated by averaging the motion compensated prediction from the next frame as well as from the previous frame. The ratio of the three types of frames is left up to the user. The prediction errors in the case of the predicted and interpolated frames, and the intraframes are coded using the DCT. The DCT implementation is according to the proposed JPEG standard. In the JPEG standard, the quantization of the coefficients is accomplished using a quantization table, and a specific variable length coding strategy.

While these systems provide substantial compression at high quality (for most sequences), as seen from the previous sections, there is need for providing substantially more compression without compromising the quality. The quality issue will become more and more important as people become accustomed to better and better quality video in their everyday lives. To provide high quality video at lower compression rates there is need for research in a number of different areas.

An integral part of any lossy compression scheme is the quantizer, and a substantial amount of effort has been devoted to finding quantizers, which achieve the theoretical limits. This effort has paid off resulting in quantizers which come very close to the theoretical limit. These include vector quantizers¹¹, trellis coded quantizers¹², and recursively indexed quantizers¹³. Depending on the amount of complexity acceptable, one or more of these quantizers can provide performance very close to the theoretical limit. Any improvement in this area will be incremental at best. However, the measure used to evaluate performance has been the SNR which, as mentioned previously does not necessarily correlate with perceived quality. While there has been some work done on quantization in the perceptual domain, more work still needs to be done in this area.

Another major aspect of any lossy image compression scheme is the model. Depending on the model one can come up with a variety of different coding schemes. These include differential encoding schemes, transform based schemes, vector quantization based schemes, and combinations of these. Again, as in the lossless coding area, it is doubtful that any one model will ever completely describe an image. The best approach will again probably be a composite source model, which uses

all available models including the relatively new approaches of fractals, and prediction trees.

An important component in many of the video coding algorithms being proposed is motion compensation. The basic idea behind motion composition is to use knowledge of the trajectories of various objects in the scene to remove interframe redundancy, thus increasing the compression rate. Most compensation techniques divide the image into rectangular blocks. Motion is assumed to be constant over the entire block, and a displacement vector is estimated for each block. Even though this approach seems somewhat primitive, it can result in a doubling of the compression rate for sequences with relatively little motion (it can also result in the generation of artifacts if the actual motion is greater than that assumed, during design). The prediction in motion compensation techniques is from one frame to the next, that is, it is a first order predictor. One can speculate, that if the restrictions in shape and order could be lifted, the gains from motion compensation can be substantially increased. Currently, the cost of removing these restrictions is prohibitive in terms of computation, and memory. However with improvements in technology, this could well be within reach in the near future. Consider, for example the use of a three dimensional transform, which would partially remove the order restriction (generally one would use an $N \times N \times N$ cube, giving an effective "prediction" order of N). This would require about 70 million multiplies and 40 million adds per second¹⁴. This would have seemed an excessive a decade ago. However, it is not totally unreasonable given current technology.

From this brief and admittedly selective overview of the image compression area, we can see that while significant progress has been made in the past years, there is still substantial progress that needs to be achieved in the future.

5.0 Conclusions

This paper has examined video compression requirements to support Space Exploration Initiative missions and discussed a number of potential video data compression techniques which could be used in addressing the requirements. These requirements must be viewed as a minimum set, since they are based upon fitting within the maximum projected channel capacities for the SEI mission communications links. While several compression techniques exist today which can fulfill some of these minimum SEI image/video requirements, additional research is needed to develop more efficient compression approaches. New methods for modeling the source information through use of

spanning trees can lead to more efficient lossless compression techniques, but these will require technology advances to handle the increased computational complexities. For lossy compression applications, additional research into quantizer design, source modeling, and motion compensation is needed to provide high quality video at lower compression rates.

References

1. Mission Analysis and Systems Engineering 1989 Final Report - Volume 6
2. D. S. Ponchak, J. E. Zuzek, W. A. Whyte, R. L. Spence, and P. Y. Sohn, "A Technology Assessment of Alternative Communications Systems for the Space Exploration Initiative", NASA Lewis Research Center Report to the Office of Space Operations.
3. Deng, R. H. and Costello, D. J. Jr., "High Rate Concatenated Coding Systems Using Bandwidth Efficient Trellis Inner Codes," IEEE Transactions on Communications, vol. 37, May 1989.
4. J. Rissanen, *Stochastic Complexity in Statistical Enquiry*, World Scientific, 1989.
5. K. Sayood and K. S. Anderson, "A Differential Lossless Coding Scheme", Proc. 31th Midwest Symposium on Circuits and Systems
6. N. D. Memon, S. Magliveras, and K. Sayood, "Prediction Trees and Lossless Image Compression" Proc. Data Compression Conference, Snowbird, UT, April 1991
7. J. B. Kruskal, "On the Shortest Spanning Subtree of A Graph and the Travelling Salesman Problem," *Proc. Amer. Math. Soc.*, pp.48-50, 1956.
8. V. Ramamoorthy and N. S. Jayant, "High Quality Image Coding With a Model-Testing Vector Quantizer and a Human Visual System Model," *ICASSP 88*
9. M. J. Shalkhauser and W. A. Whyte Jr., "Digital CODEC for Real Time Signal Processing at 1.8 bpp," *Global Telecommunication Conference*, Dallas, TX, Nov. 1989. Also *NASA Technical Memorandum 102325*.
10. Communications of the ACM, April 1991

11. Y. Linde, A. Buzo, and R. M. Gray, "An Algorithm for Vector Quantizer Design," IEEE Trans. Commun., vol. COM-28, pp. 84-95, Jan. 1980.
12. M. W. Marcellin and T. R. Fischer, "Trellis Coded Quantization of Memoryless and Gauss-Markov Sources," IEEE Trans. Commun. vol. COM-38, pp. 82-93, Jan. 1990.
13. K. Sayood and S. Na, "Recursive Quantization," Proc. Asilomar Conf. on Circuits Systems, and Computers, Nov. 1990.
14. M. Bauer, Group Meeting Presentation

Table 1 - SEI Image/Video Minimum Compression Requirements

Data Type	Raw Data Rate	Compressed Data Rate	Comments
High Rate Video	100 Mbps	10 Mbps	Single channel, color, 512 X 512 pixels, 12 bits/pixel, 30 frames/sec requiring 10:1 data compression with no perceptible quality degradation. Includes intraframe and interframe compression.
Edited High Rate Video	20 Mbps	1-2 Mbps	Quality similar to teleconferencing with some frames dropped (transmit < 30 frames/sec, display at 30 frames/sec)
Low Rate Video	2 Mbps	0.2 Mbps	Single channel, monochrome, 512 X 512 pixels, 8 bits/pixel, 1 frame/sec
Science Imaging Data	300 Mbps	30 Mbps	Not well defined as yet, 1024 X 1024 pixels, 8 bits/pixel, RGB signal. Desire for lossless compression for many applications. Variable frame rate requirements.
Telerobotics Video	200 Mbps	20 Mbps	Two channels, color, 512 X 512 pixels, 8 bits/pixel, 30 frames/sec, requiring 10:1 data compression and no perceptible quality degradation to teleoperator.

Table 2 - Comparison of Compression Results

Image Name	PKARC	DLCS	Entropy
USC Girl (256X256)	6.08 bpp (24%)	5.68 bpp (29%)	5.04 bpp
USC Couple (256X256)	5.84 bpp (27%)	5.12 bpp (36%)	4.83 bpp
Aerial (256X256)	7.44 bpp (7%)	6.72 bpp (16%)	5.97 bpp
Andy (512X384)	5.20 bpp (35%)	3.76 bpp (53%)	3.91 bpp
Karanne (512X384)	5.92 bpp (26%)	4.24 bpp (47%)	4.16 bpp
Eweek (512X384)	4.72 bpp (41%)	3.60 bpp (55%)	3.46 bpp
Beach (512X768)	5.36 bpp (33%)	3.52 bpp (56%)	3.87 bpp
Makeup (512X768)	4.24 bpp (47%)	2.79 bpp (65%)	2.91 bpp
Soap Opera (512X768)	4.64 bpp (42%)	3.20 bpp (60%)	3.11 bpp

Table 3 - Lowest Possible Lossless Compression Using Minimum Entropy Trees

Image Name	DLCS	Codebook Method	Spanning Tree Entropy
USC Girl (256X256)	5.68 bpp	4.46 bpp	3.02 bpp
USC Couple (256X256)	5.12 bpp	3.90 bpp	2.83 bpp
Aerial (256X256)	6.72 bpp	5.68 bpp	4.49 bpp
Andy (512X384)	3.76 bpp	3.06 bpp	1.97 bpp
Karanne (512X384)	4.24 bpp	3.22 bpp	2.21 bpp
Eweek (512X384)	3.60 bpp	2.60 bpp	1.70 bpp
Beach (512X768)	3.52 bpp	2.68 bpp	1.92 bpp
Makeup (512X768)	2.79 bpp	1.98 bpp	1.23 bpp
Soap Opera (512X768)	3.20 bpp	2.30 bpp	1.55 bpp

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
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 September 1991		3. REPORT TYPE AND DATES COVERED Technical Memorandum
4. TITLE AND SUBTITLE Data Compression for Full Motion Video Transmission			5. FUNDING NUMBERS WU-643-10-01	
6. AUTHOR(S) Wayne A. Whyte, Jr. and Khalid Sayood				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Lewis Research Center Cleveland, Ohio 44135-3191			8. PERFORMING ORGANIZATION REPORT NUMBER E-6564	
9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, D.C. 20546-0001			10. SPONSORING/MONITORING AGENCY REPORT NUMBER NASA TM-105239 AIAA-91-3533	
11. SUPPLEMENTARY NOTES Prepared for the Conference on Advanced Space Exploration Initiative Technologies cosponsored by AIAA, NASA, and OAI, Cleveland, Ohio, September 4-6, 1991. Wayne A. Whyte, Jr., NASA Lewis Research Center; Khalid Sayood, University of Nebraska-Lincoln, Lincoln, Nebraska 68508. Responsible person, Wayne A. Whyte, Jr., (216) 433-3482.				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Unclassified - Unlimited Subject Category 32			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) Clearly transmission of visual information will be a major, if not dominant, factor in determining the requirements for, and assessing the performance of, the SEI communications systems. Projected image/video requirements which are currently anticipated for SEI mission scenarios are presented. Based on this information and projected link performance figures, the image/video data compression requirements which would allow link closure are identified. Finally several approaches which could satisfy some of the compression requirements are presented and possible future approaches which show promise for more substantial compression performance improvement are discussed.				
14. SUBJECT TERMS Data compression; Mars communication; Lunar communication; Lossless coding; Lossy coding			15. NUMBER OF PAGES 12	
			16. PRICE CODE A03	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT	

