Loughborough
University

This item was submitted to Loughborough's Institutional Repository (https://dspace.lboro.ac.uk/) by the author and is made available under the following Creative Commons Licence conditions.

For the full text of this licence, please go to:
http://creativecommons.org/licenses/by-nc-nd/2.5/

Error Concealment Algorithms for an ATM Videoconferencing System

L.T.Chia, J.W.R.Griffiths, D.J.Parish, I.W.Phillips.
Department of Electronic and Electrical Engineering,
Loughborough University of Technology,
Loughborough, Leics, LE11 3TU, United Kingdom.

Abstract

The paper describes a videoconferencing workstation using ATM for the transmission of all data and Motion-JPEG for compression of the video. Of particular interest is the work on algorithms to help conceal the errors that arise from transmission over an ATM Network. Several methods are discussed and the results of some subjective assessments of typical images are presented.

1. Introduction.

In an ATM network cells may be lost due to random bit errors in a cell header and perhaps more seriously, to network control when network traffic is congested. In most cases, an absolute guarantee that video cells will not be lost under any circumstances cannot be given and the video service must accept this possibility. There are many methods for dealing with this problem but most of them require that extra information is transmitted thus increasing the bandwidth required. This paper detail several methods of error concealment which do not require the transmission of additional information. In section 2 the experimental system is described. The video coding algorithm, Motion-JPEG, is discussed in section 3. An important aspect of the JPEG algorithm is the availability of marker codes which are described in section 4. Section 5 gives a general discussion of methods of dealing with the ATM errors, while section 6 describes a particular set of error concealment algorithms based on prediction from the neighbouring blocks. Section 7 describes the subjective experiments and the conclusions are contained in section 8.

2. Overview of System.

The basis workstation was an IBM PC running Windows equipped with adapter cards to provide multimedia and networking functionality. The connection to the ISDN network was via a propriety exchange referred to as the MATMX (Model ATM eXchange) developed by the Rutherford and Appleton Laboratories under a RACE project.

2.1 Multimedia Hardware and Software.

The workstation uses a complementary pair of multimedia cards, made by VideoLogic: the DVA-4000 for providing an image capture service, and the MediaSpace which can compress captured images using JPEG compression. The MediaSpace incorporates compression hardware -the CL550- and can encode video at 25 frames per second. Use of JPEG in this way is called Motion-JPEG. The MediaSpace card can also record and playback ADPCM or PCM audio.

This hardware was specifically designed to record and playback only to and from a local disk. Extensive modification of the control software took place to provide network support. Additionally, the MediaSpace incorporates an Inmos transputer and can be programmed to receive and send its data over a spare transputer link rather than via the PC bus. This facility was exploited in order to provide a data path from camera to remote display, across a network, independently of the main PC CPU and bus.

2.2 Network Hardware and Software.

The workstation incorporates a third adapter card to provide a network interface. As the local interface to the MATMX is via a transputer link, a TRAM motherboard is used to provide a network interface from the PC. TRAMs can be placed on this motherboard to provide enhanced levels of workstation support. These transputers are programmed in occam. The use of transputers and occam provided flexibility of development, design and implementation, and additional functionality. Typical operations performed by the transputers include video cell loss error concealment which is discussed later. It should be noted that for a commercial product a dedicated IC could replace the transputers in order to reduce system cost.

3. Motion-JPEG and JPEG.

The JPEG Still Picture Compression Algorithm is defined in the ISO Draft International Standard 10918 [1] and the C-Cube CL550 Chip implemented the Baseline System, a DCT based system that is a mandatory part of the standard.

The Baseline Process is a DCT-based algorithm that compresses images having 8-bits per pixel per component. The baseline system operates in sequential mode, i.e., the image is processed from left to right and from top to bottom in a single pass. Minimum buffering is required and this permits inexpensive implementation [2,3]. The JPEG algorithm has been

837

designed to become a general purpose standard that may be used by a board range of applications that could include file distribution. JPEG also specifies an Interchange Format that may be used to transfer compressed image files from one application to another.

4 Marker Codes in JPEG Images.

Compressed image data is described by a uniform structure and set of parameters for all mode of operations. The various parts of the compressed image data are identified by specific two-byte codes called markers. Some markers are followed by particular sequences of parameters such as table specifications and headers. Others are used without parameters, such as start-of-image and end-of-image.

The entropy coded data can be segmented and a reset marker is used to isolate entropy-coded data segments. The encoder outputs the reset markers, intermixed with the entropy-coded data at regular reset intervals. Reset markers can be identified without having to decode the compressed data as all markers start at a byte boundary. If the entropy-coded data ends before a byte boundary, it is padded to the next byte boundary with the required number of '1' bits. Because the segments can be independently decoded, the segmented mode of operation leads to other application-specific uses, such as parallel encoding and decoding, isolation of data corruption, and semi-random access of entropy-coded segments. The CL550 JPEG Image Compression Processor provides the option of reset markers. The Reset Interval can be stored in the Coding Interval Registers and this Coding Interval defines the numbers of Minimum Coded Units (MCU) per coding interval.

5 Treatment of Video Cell Loss.

There are two main methods of dealing with the loss of a video segment;

- Error Correction :- in which the missing cell is detected and replaced exactly using extra information inserted into the data stream.
  - Error Concealment :- in which a missing cell is detected and replaced by an approximation of its content derived from existing information in the data stream. This requires no extra coding information.

An ATM cell may contain several MCUs (depending upon the information within that part of the image). Therefore to avoid bursts of loss MCU blocks it is necessary to interleave the blocks before transmission. Various methods of replacing the loss blocks were investigated [4]:

- Replacement of the lost block with a grey block or a suitably chosen colour.
- Replacement of the lost block with a block from the previous frame (correlated frames only).

- Replacement of the lost block with the previous block within the frame (uncorrelated frames).
- Replacement using information from the neighbouring blocks within the same frame to estimate the contents of the block in error.

The first method produces very poor results particularly for moving sequences. The second method is very satisfactory but requires that successive frames are correlated. In a moving sequence this is normally the case unless there is a lot of motion or a change of scene but for applications where the individual frames are completely different, e.g., in a catalogue browsing application, this method does not work.

The third method is for moving sequences with uncorrelated frames or still images. This is a simple method that requires no computation, only slight buffering is required to store the previous block. It works well around areas which are similar, for example the background of a scene, but not when there is a sharp transition of the pixel intensities between the lost block and the previous block. The replacement block will be visually out of place but the method is still an improvement over the first method.

The fourth method can produce quite satisfactory results for many image types but does require some extra computation at the receiver. The processing was implemented in the transform domain since this enabled advantage to be taken of the energy concentration properties of the DCT. It is this method which has formed the basis in this paper and some examples of the results obtained are shown in figure 3. The next section will therefore concentrate on this fourth method.

6 Improving Error Concealment

In this section the significance of the coefficients in the transform domain is examined and the findings incorporated in the error concealment algorithm.

For images with high inter-element correlation, most of the energy is represented in the lower-order spectral components of the DCT (situated towards the top left corner of the transformed block). Higher order components will be small and may be coded with very few bits or deleted completely. For JPEG images, the quantisation of the transform coefficients according to the human visual system would remove most of the higher order spectral components. It was thus decided that the prediction methods should concentrate on the first four transform coefficients only and a study was made of the first and second order statistics of these coefficients.

The first order distribution of the pixel values is unlikely to be Gaussian but since the DCT coefficients are a weighted linear sum of the pixel values in a block it might be expected that the first order distribution of the transform coefficients might be Gaussian. However the correlation between pixels in the image domain for a typical head and shoulder image is very high, for

838

example a typical 1-step correlation is about 0.9 and for a lag of 8 pixels the correlation is still about 0.7 [5]. Because of this the distribution of the DC coefficients is similar to the pixel distribution. The distribution of the AC coefficients is a nearer approximation to the Gaussian but in fact has a sharper peak and longer tails.

As expected the DC coefficients show a strong correlation for small block separation. However the F(0,1) and F(1,0) coefficients show a strong correlation in one direction only, namely vertical for F(0,1) and horizontal for F(1,0). The coefficient F(1,1) has little correlation in either direction. A reason for this can be found from a study of the basis images for these coefficients shown in figure 1.
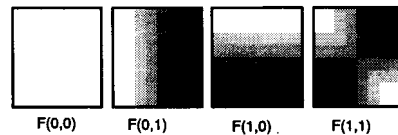

| F(0,0) | F(0,1) | F(1,0) | F(1,1) |

Figure 1: First 4 Basis Images for the DCT

For the F(0,1) basis image there is continuity of intensities along the edge when two F(0,1) basis image are placed side by side in the vertical direction but there is a discontinuity when placed side by side in the horizontal direction. Similar reasoning applies for the F(1,0) basis image but in the opposite directions. From consideration of the basis image for F(1,1) it will be seen that there will be a discontinuity when two F(1,1) basis images are placed side by side in any direction and hence the correlation might be expected to be low [6].

From this information it is possible to reduce the amount of computation required in the prediction of a missing block by considering only the coefficients which could influence the lost coefficients.

7. Estimation of Image Quality.

The purpose of this part of the work was to re-enforce the argument used earlier that only a few of the DCT coefficients are required to give a good estimate of the missing blocks. Experiments on the human visual response to sine-wave and square wave spatial distributions, viewed at a range of spatial frequencies, were conducted as early as the 60s [8,9] and the results indicated the importance of low spatial frequencies in the human visual contrast phenomena. The emphasis here was to look for transform coefficients that are essential for generating horizontal, vertical and diagonal edges

Image quality is, of course, very much a subjective measure. A number of images were constructed using a very limited set of transform coefficients and these were compared by a number of observers using the following impairment scale [7]:

1) Very annoying.
2) Annoying.

3) Slightly annoying.
4) Perceptible, but not annoying.
5) Imperceptible.

7.1 Test Images.

A total of 3 test images were selected for the assessment. They were standard images used in various experiments and the images were then compressed using the JPEG still picture compression algorithm. The three images used for testing were:

IMAGE 1:"LENA"-A head and shoulder image of a girl with a hat.
IMAGE 2:"MANDRILL" - The face of a large baboon, an image with high AC energy around the face.
IMAGE 3:"PEPPERS" - An image filled with red and green capsicums and chilli pods.

Eight images were then generated from each of the test images. For each of the eight images, a different number of transform coefficients were selected for decoding the compressed images according to the descriptions in the table below

Image.	Description.
(a)	Original.
(b)	Only the DC coefficient.- F(0,0).
(c)	DC and the first two AC coefficients.- F(0,0), F(0,1) & F(1,0).
(d)	DC and the nearest three AC coefficients.- F(0,0), F(0,1), F(1,0) & F(1,1).
(e)	DC and 5 selected AC coefficients.- F(0,0), F(0,1), F(1,0), F(1,1), F(2,1) & F(1,2).
(f)	First 6 coefficients in zigzag scan.- F(0,0), F(0,1), F(1,0), F(1,1), F(2,0) & F(0,2).
(g)	First 10 transform coefficients in zigzag scan.
(h)	First 21 transform coefficients in zigzag scan.

A subjective assessment was conducted and twenty subjects were involved in the experiment, six were considered as experts and fourteen as non-expert. An expert is defined as a person who has been involved in television, photographic or visual arts work and may sometimes be involved in critical examination of pictures.

As far as possible, the picture conditions specified in CCIR recommendation 500 [6] were followed. The subject was helped to "anchor" his quality ratings by informing him or her that Image (a), the original image, is the best and should be given a quality rating of 5. The subject was presented with the pictures in random order and was asked to rate them according to quality. Each picture was viewed at a distance of four to six times the image height and for as long as the subject wished.

Since several subjects were used in the evaluation process, the mean rating is given by:

$$R = \sum_{k=1}^{n} s_k n_k \Big/ \sum_{k=1}^{n} n_k$$

where s_k is the score associated with the kth rating, n_k is the number of observers with this rating, and n is the number of grades in the scale.
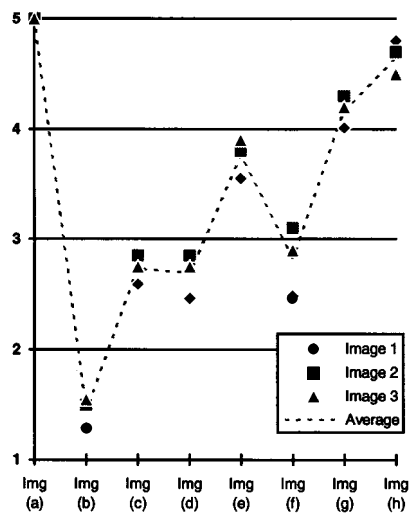

Figure 2: Results of Subjective Testing

The average results for the three images are shown in figure 2. From the results, it can be seen that image (b) is subjectively the worst. Image (c), (d) and (f) produced an impairment scale of around "2" and "3", which are classified as "*annoying*" and "*slightly annoying*".

Both image (e) and (g) have an average rating of "4", which is a much higher rating than most of the other images. Image (g) has the higher average of the two, which is hardly surprising, as the number of coefficients used for decoding was much higher. The highest overall rating in each case is image (h) with an average rating of "5", which is considered as "*imperceptible*" in the impairment scale. The most interesting result is the high average score achieved by image (e). Both images (e) and (f) use six transform coefficients in the decoding of the image. Image (f) uses the first six transform coefficients in the zigzag scan order, namely $F(0,0)$, $F(0,1)$, $F(1,0)$, $F(2,0)$, $F(1,1)$ and $F(0,2)$. Whereas image (e) uses the 6 transform coefficients $F(0,0)$, $F(0,1)$, $F(1,0)$, $F(1,1)$, $F(2,1)$ and $F(1,2)$.

8 Conclusions

A method of replacement by linear interpolation of the transform coefficients has been improved by an initial study of the transform coefficients which resulted in a reduction in the number of coefficients necessary for predicting the missing block.

Working with DCT blocks with an edge in either the horizontal, vertical or diagonal direction provided some hints on the relative importance of the various transform coefficients and with this knowledge, six coefficients were selected and subjective studies used to show that favourable reproduction of the original images can be achieved with just these coefficients. The findings were then incorporated in the error concealment algorithm.

The main conclusion from this work is that error concealment works well and that because of the energy compaction efficiency of the DCT only a few of the coefficients of the neighbouring blocks are required to predict the parameters of the missing blocks.

9 References.

[1] (JPEG1) ISO DIS 10918-1, "Digital Compression and Coding of Continuous-tone Still Images." ISO/IEC JTC1/SC2/WG10, Jan 1992.

[2] G.K. Wallace - "Overview of the JPEG (ISO/CCITT) Still Image Compression Standard.", SPIE, Vol. 1244, Image Processing Algorithms and Techniques, pp.220-233, 1990.

[3] A.Leger, T.Omachi, G.K.Wallace - "JPEG Still picture compression algorithm", Optical Engineering, Vol30, No.7, pp 947-954, July 1991.

[4] L.T.Chia, D.J.Parish, P.J.Coventry, I.W.Phillips, J.W.R.Griffiths - "Motion-JPEG on a Network and the Treatment of Video Cell Loss", 5th Int. Workshop on Packet Video, Berlin, Germany, 1993.

[5] R.J.Clarke, "Transform Coding of Images", Academic Press, London, 1985.

[6] L.T.Chia, D.J.Parish, J.W.R.Griffiths - "On the Treatment of Video Cell Loss in the Transmission of Motion-JPEG and JPEG Images", Computers and Graphics Vol. 18, No. 1, Jan 1994.

[7] J.Allnatt, "Transmitted-Picture Assessment", John Wiley & Sons, 1983.

[8] J.J. DePalma and E.M. Lowry, "Sine-Wave Response of the Visual System. II. Sine-Wave and Square-Wave Contrast Sensitivity", J. Opt. Soc. Amer., Vol. 52, pp. 328-335, Mar. 1962.

[9] D.H. Kelly, "Visual Contrast Sensitivity", Opt. Acta, Vol. 24, pp. 107-130, Feb. 1977.

(3a) : Original Image	(3b) : Image with 5% MCUs Loss	(3c) : Previous Block Replacement of Missing Block.
	(3d) : Linear Interpolation of all DCT coefficients	(3e): Linear Interpolation of the first 3 DCT coefficients

Fig 3 : Error Concealment
of Loss DCT Blocks