A WORKBENCH FOR THE COMPUTER SIMULATION OF PICTURE CODING SCHEMES

M A Ould, N D Birrell, P R Radford, N Tucker & D I Crawford*

Logica Ltd, UK & *British Telecom Research Laboratories, UK.

INTRODUCTION

At the British Telecom Research Laboratories (BTRL) at Martlesham Heath there is increasing activity in picture coding research to meet the demands for more efficient visual telecommunications. Picture coding techniques are being investigated for the coding of single frames (static pictures) and multiple frames (moving picture sequences).

The work on single frame coding is required for applications involving still picture (freeze frame) television transmission systems such as Picture Prestel (videotex). In these systems the static picture information is transmitted at rates compatible with available telephone network transmission bandwidths. Optimisation and comparison of various coding techniques are performed with a view to reducing the transmission times for a given picture quality, when the bit rate is limited to the range 1.2 kbit/s to 64 kbit/s.

Research into the coding of moving picture sequences and their transmission in real time involves applications ranging from relatively low resolution monochrome videoconference transmissions at 2 Mbit/s to studies of future high resolution colour picture transmissions at much higher rates.

For still picture working, single frame stores have been built to capture and hold the video data for coding. A dedicated microprocessor has then been employed to process the data held in the store. Various coding algorithms, such as intraframe differential pulse code modulation (DPCM) and two dimensional transform coding, have been studied in this way. However, such systems have proved to be relatively slow both in programming and processing times.

Because of the limited available storage (one frame), the above technique is unsuitable for work on realistic moving picture sequences. As a result, for work in this area, flexible prototype codecs have been built in hardware, for optimisation in real time. This has proved time-consuming and costly, because of the necessary complexity of the codec designs.

In order to serve the needs of research groups in both areas, it is proposed to build up a computer simulation system suitable for all picture coding research at the laboratories. The system is required to be:

1. Multi-user, able to operate on several tasks simultaneously.

2. User-friendly, operating with a high level image processing language and supplying meaningful diagnostics.

3. Suitable for work on both still pictures and moving picture sequences.

4. Based on a central processor that is well-supported and reasonably powerful.

5. Flexible, to operate on a range of video formats and sampling frequencies.

6. Suitable for expansion in terms of number of users, moving picture sequence length, and processing power.

VERSATILITY AS THE KEY DESIGN CRITERION

BTRL's aim is to set up a multi-user image processing system for the investigation of coding techniques for still and moving TV pictures. There are two separate dimensions to this requirement:

- firstly, to be versatile in the simulations of coding schemes (the "logical" dimension).
- secondly, to be versatile in the TV environment in which coding schemes can be tried out (the "physical" dimension).

The immediate implications are:

- the system must provide good basic image processing facilities plus the tools with which users can build their own, more complex tools
- the system must be versatile in the capture and replay of video; in particular it must be capable of supporting a variety of sampling frequencies (eg 13.5 MHz, 16 MHz, 20 MHz), and a variety of sampling algorithms for RGB and YUV coding formats.

The responsibilities for supporting the two dimensions divide naturally between the software and hardware in the system currently being implemented. Firstly, users are supplied with the necessary image processing toolkit through Logica's image processing system INSIGHT. This software package (first implemented on PDP-11 computers and now available on the VAX) has been enhanced by incorporating facilities specific to this application - moving sequence capture and replay for instance. Secondly, a specially designed hardware "workbench", comprising two microprocessor controlled framestores and a high-speed digital storage system provide the necessary physical versatility.

The system being developed

Figure 1 shows the major components of the BTRL Image Processing System (BTRLIPS) hardware. A one megabyte Digital VAX-11/750 mini-computer supports the image processing package INSIGHT which provides access to and processing of images. The VAX is connected to two digital framestore engineered by Logica Limited - the Single Framestore (SFF) and the Moving Sequence Framestore (MSF).

These framestores can each hold up to

0.5Mpixels of TV picture in three bands (RGB or YUV); each accepts analogue input via three analogue to digital converters (ADCs) and generates analogue output via three digital to analogue converters (DACs). The MSF has the same capabilities as the SFF but is additionally attached via a high speed data channel to a pair of Ampex Parallel Transfer Disks (PTDs) that are together capable of sinking or sourcing digitised TV at video speed - see below.

The PTDs hold up to about 800 frames of video, ie a sequence of around 30 seconds. The space on them can be divided up into named chunks or "sequence files" into which a user, via INSIGHT on the VAX, can read full speed video from a suitable video source, having it digitised and buffered through the MSF. Sequences held on the PTDs can be replayed at video speed back through the MSF to video sinks.

The SFF can catch single frames of video and output a currently held frame. Like the MSF, it also allows a sequence of consecutive part-frames to be captured in the framestore or replayed from the framestore without the use of the PTDs, the length of the sequence being determined by the size of the part-frames. This feature allows a new coding scheme to be tried out first on a small section of each picture in a sequence. These small sections can then be "animated" cyclically before processing of full pictures in a long sequence is committed to.

Physical versatility

BTRLIPS is designed to handle 625 line TV with 525 and 313 line standards being a straightforward enhancement.

Different base sampling frequencies are supported on each framestore by individual Clock Pulse Generator circuits (CPGs) that can be inserted in the framestore as required.

A number of different "sampling algorithms" are supported for video input and output. All three channels can be sampled at the base sampling frequency of the CPG ("4-4-4") or all at three-quarters (3-3-3) or one-half (2-2-2) of that rate - these algorithms are intended for work with RGB. For YUV, the chrominance U and V channels can be sampled at half the Y signal rate (4-2-2) or one third (3-1-1), 2-1-1 is also supported - the Y signal being sampled at half the base sampling frequency and the U and V signals at half of that. Monochrome algorithms (4-0-0, 3-0-0 and 2-0-0) simply require the user to disconnect or ignore all signals except one - Y. Other algorithms such as 3-1-0, 1-1-1 etc can be simulated by suitable processing via INSIGHT.

Users therefore have considerable flexibility when capturing or replaying video in their choice of "sampling regime" - the combination of TV line rate, base sampling frequency and sampling algorithm.

Windowing on both input and output is an important facility. The user can place a window on his video source if he is not interested in capturing the whole active picture in a framestore. Likewise he can place a window on the digitised picture in the framestore if he wishes to display only a portion of that picture on the analogue outputs.

Image processing versatility

INSIGHT allows users at terminals to process images held in files on the VAX RM03 disks. Using an interactive BASIC-like language a user can call up "primitives" that perform specific processing operations on images. A user can supplement the supplied primitives by writing his own in FORTRAN. INSIGHT also makes available to writers of primitives a number of subroutine libraries that facilitate image accessing (at image and line level) parameter passing and intertasking. Standard supplied primitives include ones for controlling the capture and replay of still frames, moving sequences and animation sequences by the SFF and MSF, and for adding and subtracting images, edge detection, 3X3 linear spatial filtering, 2-D Fourier transformations, thresholding etc. Taken together, these primitives and the lower level subroutine libraries provide a powerful toolkit for the experimenter in coding techniques.

REAL TIME IMAGE CAPTURE

Requirements

The requirement to store and play back sequences of television frames in digitised real time places severe constraints on the data storage system that can be used. The specifications of such recording systems which are of most importance in this respect are:

- data recording/playback rate (what sampling regime can be used ?)
- storage capacity (how many frames can be stored?)
- data accessibility (can single frames be retrieved?)
- availability and cost (can it be bought now at a reasonable price?).

The requirements for the BTRLIPS storage system are:

- a total recording rate of at least 18M bytes/s. This data rate allows in particular the recording of sequences of full frames of 576 lines sampled at 12 MHz in Y and 4 MHz in U and V, as well as frames with fewer pixels of course.
- recording capacity for about 30 seconds of digitised television. This corresponds to a storage requirement of at least 606 Mbytes for the storage of 30 seconds of YUV-coded 625 line TV samples at 13.5 MHz
- individual frames must be retrievable from a sequence of frames to allow individual frame processing. It should then be possible to restore the processed frame at the position in the sequence from which it was originally retrieved
- the storage system should be available now and must not have an exorbitantly high price.

Storage Systems

The types of storage media that could be considered as the best contenders for meeting the requirements given above are:

- solid state random access memory (RAM)
- magnetic tape
- magnetic disk.

For each of these media, we consider below the merits and the limitations of such a storage system in relation to the real time image capture requirements.

RAM. RAM has been rapidly decreasing in price in recent years and might now be thought of as a serious contender for real time image sequence capture. The simplest solution for using RAM in this way is to pipe digital video directly into the main memory of a computing system where it can be manipulated by the computer's processor. In such a solution, meeting the target data recording rate is obtained by using a computer with a sufficiently high data rate on its memory access (DMA) channels. Suitable (large) mini-computers are now available for such applications. For example, the Perkin-Elmer 3244 supports up to four DMA channels each with a maximum data rate 10 Mbytes/s into a corresponding memory module of up to 4 Mbytes in size.

Such a system would allow approximately 10 frames, or about 0.45s, of video to be captured. This does not meet the 30s target set above, but similar solutions are likely to become increasingly viable in the near future.

Tape. High-speed high-density digital magnetic tape units have been used in aero-space and military systems for sometime now and units with data recording rates of 18 Mbytes/s or more are available as virtually standard items (from, for example, Honeywell). The principal disadvantage of most such tape units is that they take some seconds to come up to speed and thus do not allow single frames to be readily retrieved from and then returned to a sequence. The only tape unit that we have found on the market at the present time that does not suffer from this problem is the Bell and Howell System 600. This machine has a 2.5 ms start time and is capable of recording 75 Mbytes/s, but it also has a price in seven figures. In the near future, several manufacturers are likely to introduce high data rate tape units based on the helical scan principle. These recorders will not suffer from the problems associated with presently available units. In particular Ampex have a 37.5 Mbytes/s helical scan recorder under development which is expected to be released within about two years.

Until helical scan recorders are available, the only possible use of a high speed tape recorder in a system such as BTRLIPS is as a buffer for image data which is subsequently to be transferred in less than real time to low-speed storage devices such as conventional computertapes or disks.

Disks. A conventional computer disk storage unit typically records data at a rate of about 1.2 Mbytes/s via a single (active) recording head. To achieve the target recording rate of 18 Mbytes/s or more, one must either use 15 or more conventional disk units running in parallel, or a smaller number of unconventional disk units having more than one recording head active at a time, eg two disk units each recording to 8 heads in parallel, which have a combined data rate of 2 x 1.2 x 8 = 19.2 Mbytes/s.

The method of using a number of conventional disks in parallel has been used in a trial system, consisting of only two disks, at Bell-Northern Research Ltd (1). The technique used in the Bell-Northern system could in principle, be applied to larger numbers of disks. Indeed, the decreasing price of fixed pack (Winchester) disk units makes a system with 15 or 20 units appear financially quite attractive. The main disadvantages of a multi-disk system are its size and the relatively low MTBF of such a large collection of disks. The technical problem of synchronising data coming from 15 or more independently rotating disks is also not insignificant.

A more attractive solution is to use a disk which records via a more than one head at a time. Such a parallel transfer disk drive is available from Ampex in the form of its model PTD-9309. This disk drive (in standard form) allows data to be recorded via one, eight or nine heads in parallel.

Two such disk drives are used in BTRLIPS, each controlled by a MATRA HE-34 controller. The HE-34 uses all nine heads of each Ampex PTD in parallel, allowing a data transfer rate of 10.88 Mbytes/s per disk to be achieved.

In BTRLIPS each frame is stored on a separate disk cylinder, giving a maximum sequence recording length of some 800 frames of 32s. Two of the three bands (U and V) which make up a YUV image in BTRLIPS are multiplexed onto one disk, giving a maximum data storage capacity for the two bands of 362,880 bytes per frame. The third band is stored by itself on the other disk providing it with 362,880 bytes of storage per frame.

Framestores

For the framestores, a design is required that provides the following features:

- digitising of TV signals coded in up to three algorithms: RGB, YUV and monochrome.
- reconstitution of analogue signals from digital equivalents.
- dual ported store able to take in data at up to 20 MHz from ADCs and output simultaneously at the same rate to either the Ampex discs or to DACs; alternatively, to take in data at up to 20 MHz from the Ampex discs and output at the same rate to DACs.
- an alternative fast port into the store for transfer of frames between the framestore and the VAX at up to UNIBUS speeds.
- access to the framestore and other peripherals by the 8085 microprocessor that controls the framestore.
- hardware to implement the different sampling algorithms on the three bands (R, G, B or Y, U, V) under software control.

The hardware being developed is based around the technology of Flair, the video painting system manufactured by Logica under licence from the BBC [2].

SOFTWARE FOR IMAGE PROCESSING

To furnish BTRLIPS with a comprehensive image processing facility, Logica is providing a package called INSIGHT. This was originally developed to provide both an in-house development facility and a marketable software system. It is a sophisticated portable package which is simple to program and rapid in operation.

INSIGHT handles images as entities. This enables the user to concentrate on the way in which he wishes to process his images rather than on the way the computer is to perform the task.

To free the user from the programming, INSIGHT incorporates a high level language similar to BASIC. This enables him to write "procedures" that link together basic processing functions provided by the system. These basic functions are called "primitives" and a wide selection is provided with the package. The user can also write his own primitives. These are currently written in ANSI FORTRAN and have a standardised format that simplifies programming and reduces the risk of introducing errors.

To complete the user interface, a friendly error reporting technique is used that provides information on where an error occurred, what it was due to and, if the information is available, what should be done to remove it.

The basic set of primitivies covers elementary operations such as adding or subtracting one image from another, and more complex processes such as various filters. Fourier transforms and Karhunen-Loeve Principal Components Analysis.

We have identified some 18 primitives required for the system in addition to the standard set. These will provide control commands and data transfer requests for all the special hardware - framestores and PTDs. The new functions implement a complete set of operations for moving data around the system and thereby attempt to meet the future needs of the user.

A typical procedure for a user at his INSIGHT terminal might be:

i . create a sequence file on the PTDs.

ii . capture a frame sequence in this file from a camera attached to the MSF.

iii. process each frame in the sequence by

a. transferring it to an RM03 host computer "image" disk file.

b. simulating the coding technique being experimented with by processing the image with INSIGHT.

c. transferring the processed frame back to its place in the sequence file.

iv . observe the results of the coding simulations by replaying the processed sequence from the PTDs through the MSF out on to a monitor.

THE FUTURE

As to the future, enhancement of BTRLIPS can be achieved in two directions.

The first direction lies within the context of the existing hardware. Users will develop and collect new primitives and new procedures via INSIGHT. We expect that after the system has been delivered there will be a rapid realisation of the system's content capabilities as users begin to experiment.

The second direction of enhancement lies with extension of the system's hardware and software. Experience in using the system will naturally suggest improvements and extensions but enhancements seen at present include:

- the addition of an array processor to the VAX complex; provision exists within INSIGHT for an array processor - this will extend the range of image processing possible by speeding up typical manipulations.
- other sampling frequencies.
- longer sequences by upgrading the Ampex PTD sub-system.
- archiving of sequences on magnetic tape.
- supporting more users.
- other forms of user interaction, eg graphics tablets, touch screens, tracker balls and joysticks.
- support for high definition work with higher frequencies, faster line rates and a bigger framestore.
- the addition of further framestores.
- the conversion of the SFF to an MSF.

The BTRL project is part of a European collaborative research project involving equivalent organisations in Belgium, France, Holland, Italy, Sweden, the United Kingdom and West Germany [3].

It is expected that collaboration can be furthered by the exchange of video in digital form - eg through the exchange of PTD disks or magnetic tapes.

We hope to report on the progress and outcome of the project to develop this versatile image processing system in the future in an appropriate forum.

ACKNOWLEDGEMENTS

The authors wish to thank BTRL and Logica for their kind permission to publish this paper.

REFERENCES

1. Johnston, R., Mastronardi, J. and Mony, G., 1978, "A digital television sequence store", IEEE Trans Comms, COM-26, 594-600.

2. Tanton, N., 1981, ""FLAIR" - A Microprocessor-Based Tool for Graphic Design", International Broadcast Engineer, 12/178, 12-13.

3. Thompson, J. E., 1981, "European Collaboration on Picture Coding Research for 2 Mbit/s Transmission", IEEE Trans Comms, COM-29, 2003-2004.

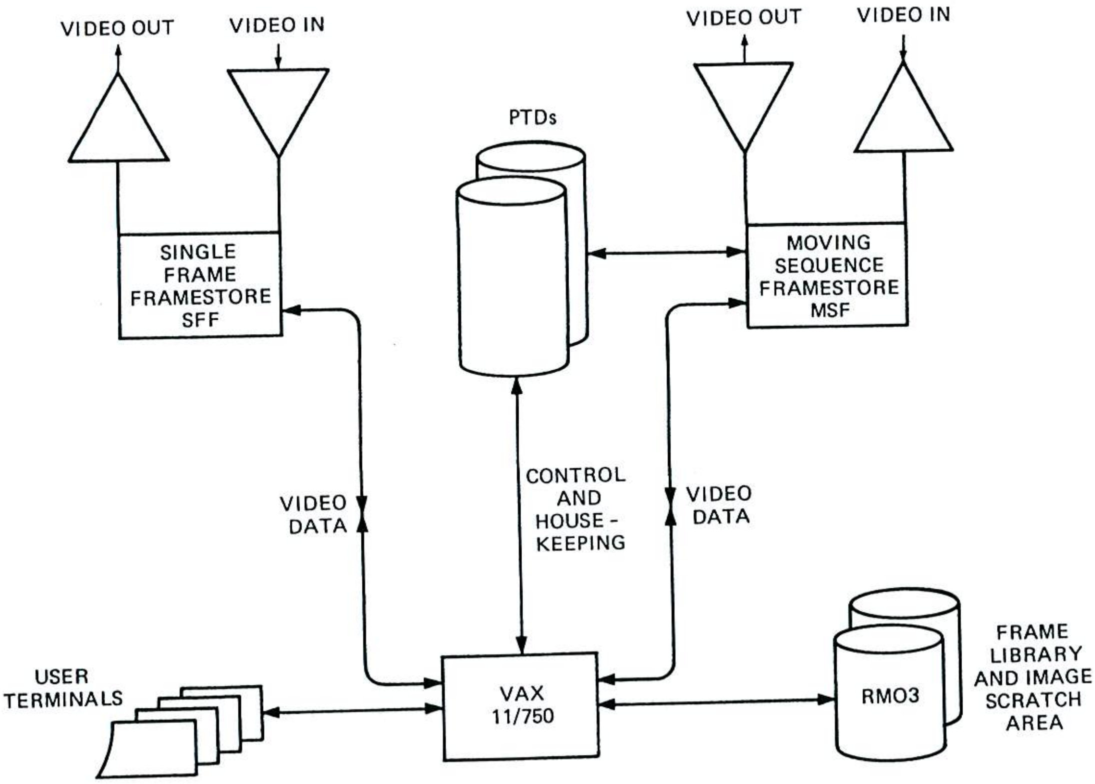

FIGURE 1 BTRLIPS - Major Components of the Physical System