



INVESTIGATION OF DIGITAL VIDEO COMPRESSION STANDARDS

SUMMARY

This report sets out the results of an investigation aimed at comparing the characteristics of various digital video compression standards with the requirements for compression in military avionics as set out in a number of ASSC reports.

The need for compression of data representing video images is discussed together with possible application areas, benefits, disadvantages and trade-offs.

The criteria for selection of standards for consideration are set out and a set of standards proposed for investigation.

The requirements set out in various ASSC documents, together with some additional generalised requirements and applications from other sources are summarised. It is concluded that only one of the ASSC documents gives any definitive requirements that may be compared with the characteristics of standards. Tabulated comparisons of those requirements and the characteristics of the selected standards are provided with supporting text.

A discussion is provided on the wider considerations regarding requirements versus standards, and a possible pro-forma for presenting the results of such comparisons is set out, with discussions of the various requirement issues identified. This is followed by a

* This report is published by the Avionic Systems Standardisation Committee (ASSC) to advance the role of standardisation in avionics. The use of this is entirely voluntary, and its applicability and suitability for any particular use, including any patent infringement arising therefrom, is the sole responsibility of the user. Copies of this paper are obtainable from the ASSC Agency as an Avionic Systems Standardisation Committee publication.

discussion of the features of various candidate standards with particular regard to the requirements issues already identified.

Conclusions are presented followed by a list of references, including URLs for web sites identified as useful sources of information during the investigation.

Readers of this report may also wish to review the follow on work set out in ASSC Document No. ASSC/130/2/163 Issue 1 March 2002 summarising the results of a brief study aimed at identifying and defining the issues effecting the application of the MPEG-2 digital video compression/decompression standard in Uninhabited Aerial Vehicle (UAV) systems.

Table of contents	Page no.
SUMMARY	i
RELATED DOCUMENTS	
viii	
GLOSSARY	ix
1 INTRODUCTION	VII
2 BACKGROUND	1
2.1 The need for compression	1
2.2 Application areas for compression	2
2.3 Benefits of compression	2
2.4 Disadvantages of compression	2
2.4.1 Latency	2
2.4.2 Loss of information and adverse effects on image quality	3
2.4.3 Susceptibility to errors	3
2.4.4 Potential problems with mixing and merging of compressed images	3
2.4.5 Visual images in automated processing applications	4
2.5 Trade offs	4
3 COMPRESSION STANDARDS	5
3.1 Criteria for Selection	5
3.2 Standards Selected for Consideration in this Investigation	6
4 CRITERIA FOR COMPARISON WITH REQUIREMENTS	6
5 SPECIFIC REQUIREMENTS	6
5.1 Requirements for the distribution and transmission of video images for airborne platforms (ASSC/130/2/116)	6
5.1.1 Data Rates	6

Table of contents (cont)	Page no.
5.1.2 Latency	7
5.2 ASSC Advisory Document on Avionic Video Compression Technology (ASSC/130/2/77)	8
5.3 Digital Video Techniques and the Possible Adoption of Commercial Video Standards in Future Military Avionic Systems (ASSC/130/2/134)	8
5.4 Applications and Requirements Identified from other Sources	9
5.4.1 Discussions with an airborne video supplier	9
5.4.2 Use of MPEG-2 to Compress Unmanned Aerial Vehicle (UAV) Gathered Enhanced Definition TV Video	10
5.5 COTS Requirement	11
5.6 Conclusions regarding Definition of Requirements	11
6 COMPARISON OF COMPRESSION STANDARDS WITH REQUIREMENTS SET OUT IN ASSC/130/2/116	11
6.1 Introduction	11
6.2 Re-statement of Requirements set out in ASSC/130/2/116 and Comparison with Standards	11
6.2.1 Sensor to Processing (Processing Link)	11
6.2.2 Processing to Display (Display Link)	21
6.2.3 Direct Video Link from Sensor to Display (Direct Link) ASSC/130/2/116	24
6.3 Latency	24
7 WIDER CONSIDERATION OF REQUIREMENTS VERSUS STANDARDS	25
7.1 Tabulation of requirements versus capabilities	25
7.2 Table 5 Entries	28
7.2.1 Application	28
7.2.2 Compression Standard	29

Table of contents (cont)	Page no
7.2.3 Requirement - Image Format	29
7.2.4 Requirement - Frame Rate	29
7.2.5 Requirement - Compression Ratio	29
7.2.6 Requirement - Lossless versus lossy compression	29
7.2.7 Requirement - Latency	30
7.2.8 Requirement - Image Quality	30
7.2.9 Requirement - Data Rate	31
7.2.10 Requirement - Error Resilience	32
7.2.11 Requirement - Functionality	32
7.2.12 Requirement - Maturity	32
7.2.13 Requirement - Hardware/software availability	33
7.2.14 Remaining Cells	33
8 DISCUSSION OF CANDIDATE COMPRESSION TECHNIQUES AND STANDARDS	33
8.1 General	33
8.2 Compression Techniques	33
8.2.1 Intraframe versus Interframe coding	33
8.2.2 Techniques for Intraframe Coding	34
8.2.3 Closed Systems	36
8.3 MPEG-1	37
8.3.1 Overview	37
8.3.2 Video Stream Composition	38
8.3.3 Comparison with Requirements from Table 5	38
8.3.4 Compression Ratio	38
8.3.5 Latency	39
8.3.6 Image Quality	39

Table of contents (cont)	Page no.
8.3.7 Data Rate	40
8.3.8 Error Resilience	40
8.3.9 Functionality	40
8.3.10 Maturity	40
8.3.11 Hardware/software availability	40
8.4 MPEG-2	41
8.4.1 Overview	41
8.4.2 Image Format	42
8.4.3 Frame Rate	42
8.4.4 Compression Ratio	42
8.4.5 Latency	42
8.4.6 Image Quality	42
8.4.7 Data Rate	42
8.4.8 Error Resilience	43
8.4.9 Functionality	43
8.4.10 Maturity	43
8.4.11 Hardware/software availability	43
8.5 MPEG-4	43
8.5.1 Overview	43
8.5.2 Applicability of MPEG-4 to Military Avionics	45
8.5.3 Maturity	45
8.6 MPEG-7	45
8.7 JPEG and its Variants	45
8.7.1 Overview	45
8.7.2 JPEG	46
8.7.3 L-JPEG	50

Table of contents (cont)	Page no
8.7.4 JPEG-LS	51
8.7.5 MJPEG	53
8.7.6 JPEG 2000	55
8.7.7 Motion JPEG 2000	58
8.8 H261 and H263	58
8.8.1 Overview	58
8.8.2 H261	58
8.8.3 H263	60
9 CONCLUSIONS	62
REFERENCES	63

Table of Tables

Table 1: Table 1 from ASSC/130/2/116 Video data rates for Sensor to Processing Link	13
Table 2: Comparison of Requirements for Sensor to Processing Link with Characteristics of Standards	15
Table 3: Table 2 from ASSC/130/2/116 Requirements for Processing to Display Link	21
Table 4: Comparison of Requirements for Processing to Display Link with Compression Standards	22
Table 5: Application Requirements versus Characteristics of Standards	27

Table of Figures

Figure 1: Comparison of JPEG CODEC image degradation against Bit Error Rate with and without RS error protection	49
--	----

RELATED DOCUMENTS

ASSC/130/2/77 Draft 5 April 1997, ASSC Advisory Document on Avionic Video Compression Technology (now integrated into Chapter 4 of ASSC/130/2/97 - Issue 2)

ASSC/130/2/134 Issue 1 May 2000, Digital Video Techniques and the Possible Adoption of Commercial Video Standards in Future Military Avionic Systems

ASSC/130/2/116*-Issue1 December 1999, Requirements for Distribution and Transmission of Video Images for Airborne Platforms

ASSC/130/2/122 Image Fusion

ASSC/130/2/38 Characteristics of Avionic Display Performance

ASSC/130/2/56 F-22 Cockpit, Lockheed Boeing Controls and Displays IPT

ASSC/130/2/97*-Issue 2 June 2000, Guide to Avionic Video Systems

ASSC/130/3/191 Overview of the MPEG-4 Standard, ISO.IEC JTC1/SC29/WG11 N1730 Stockholm, July 1997

ASSC/130/3/235 French Helmet Visors Clear the Skies for Combat Pilots, Signal Magazine 1996

ASSC/130/3/216-draft 6 Project Paper 661, Cockpit Display System Interfaces to User Systems, May 2000, ARINC

ASSC/130/3/6 Cockpit Displays, Avionics Magazine May 1996

ASSC/130/6/1-draft 4, Chapter Five: Video Data Processing**

* These documents are available on the ASSC web site at www.era.co.uk/assc/ASSCdocs.htm

** This document has now been included in ASSC/130/2/97-Issue 2.

GLOSSARY

16CIF	16 x Common Interface Format
4CIF	4 x Common Interface Format
AVO	AudioVisual Objects
B	Bidirectional coded picture
BCH	Bose-Chaudhuri-Hochquenghem code
BER	Bit Error Rate
bpp	Bits per pixel
CATV	Cable TV
CD	Compact Disk
CD-ROM	Compact Disk-Read Only Memory
CGI	Computer Graphics Interface
CIF	Common Interface Format
CMY	Cyan Magenta Yellow Colour space
CODEC	Coder/DECoder
COTS	Commercial off the Shelf
dB	deciBel
DCT	Discrete Transform Coding
DIB	Device Independent Bitmap
ED	Enhanced Definition
FVDR	Flight Video Data Recorders
Gb/s	10 ⁹ bits per second
GOP	Group of Pictures
HDTV	High Definition TeleVision
HSI	Hue, Saturation and Intensity colour space
I	Intraframe coded picture
IEC	International Electrotechnical Commission

IR	Infrared
IRST	Infrared Search and Track
ISO	International Standards Organisation
J2K	JPEG 2000
JFIF	JPEG File Interchange Format
JP2	File format defined for JPEG 2000
JPEG	Joint Photographic Experts' Group
JPEG 2000	JPEG standard for still images using wavelet compression.
JPEG-LS	Lossless and near-lossless JPEG.
kB	Thousand bytes
L-JPEG	Lossless JPEG
MAW	Missile Approach Warner
MB	Megabytes
Mb/s	Megabits per second
MB/s	Mega bytes per second
Mbytes	Megabytes
MHz	Million Hertz
MJPEG	Motion JPEG
MJPEG 2000	Motion JPEG 2000
MPEG	Motion Picture Experts' Group
P	Predicted picture
PC	Personal Computer
PSNR	Peak Signal to Noise Ratio
QCIF	Quarter Common Interface Format
RAM	Random Access Memory
RBBER	Residual Bit Error Rate
RGB	Red Green Blue
ROI	Region of interest

RS	Reed Solomon
SDI	Serial Digital Interface
SIF	Standard Input Format
SPIFF	Still Picture Interchange File Format
SQCIF	Sub Quarter Common Interface Format
SVGA	Super VGA
SXGA	Super XGA
TCDL	Tactical Common Data Link
UAV	Unmanned Aerial Vehicle
UXGA	Ultra XGA
VCS	Visually Coupled System
VGA	Video Graphics Array
WASAD	Wide Area Surveillance and Detection
XGA	Extended Graphics Array
YcBr	Colour space defined by Recommendation ITU-R BT.601
YUV	Colour space used by PAL

Collation page

1 INTRODUCTION

This report sets out the results of an investigation aimed at identifying available digital video compression standards and comparing these with the requirements outlined in the following ASSC documents:

- Requirements for the distribution and transmission of video images for airborne platforms (ASSC/130/2/116)
- ASSC Advisory Document on Avionic Video Compression Technology (ASSC/130/2/77) (now integrated into Chapter 4 of ASSC/130/2/97 - Issue 2)
- Digital Video Techniques and the Possible Adoption of Commercial Video Standards in Future Military Avionic Systems (ASSC/130/2/134).

Readers of this report may also wish to review the follow on work set out in ASSC Document No. ASSC/130/2/163 Issue 1 March 2002 (Ref. 1) summarising the results of a brief study aimed at identifying and defining the issues effecting the application of the MPEG-2 digital video compression/decompression standard in Uninhabited Aerial Vehicle (UAV) systems.

2 BACKGROUND

2.1 The need for compression

Digital video offers enormous advantages to designers of avionic systems in its ability to generate, propagate, synthesise, analyse, fuse, merge, overlay and synchronise images and other data.

However, one of the major disadvantages of digital video lies in the vast quantities of data needed to represent images. For example, a bandwidth of approximately 160Mb/s is required in order to transmit broadcast quality video, and hence a memory capacity of 200Mbytes is needed to store just a 10 second sequence. The solution to these problems of data volume versus bandwidth and memory capacity is to reduce the volume of data by compression. A variety of techniques for compressing video signals have been developed and many of these are formalised in standards and utilised in commercial products.

Scope for compression is provided by the high degree of redundancy which exists in the majority of natural images. For example, natural images frequently contain large areas in which the texture, colour and intensity are approximately constant.

Reduction of the quantity of data may be achieved using compression techniques to remove redundant or unimportant data prior to transmission, while still retaining

sufficient information to allow the original image to be reproduced with an acceptable quality for the intended purpose. The multimedia revolution relies heavily upon so called image-coding technology that reduces the volume of data required to represent information in its various forms. It is the major advances in this technology that have allowed the multimedia revolution to succeed.

2.2 Application areas for compression

Application areas for compression in military avionics may include:

- Image storage.
- Transmission within the aircraft (where available bandwidth is less than that required).
- Off-platform transmission.
- After the event viewing.

2.3 Benefits of compression

Despite the considerable advances in the performance of data transmission and memory systems in recent years coping with the volumes of data generated by digital video represents a severe challenge.

As already indicated compression techniques offer a solution to this problem by reducing the volume of data needed to represent digital video images, and hence the bandwidth required to transmit, and the memory capacity required to store such images.

2.4 Disadvantages of compression

Although compression offers major advantages, there are also significant potential disadvantages that need to be addressed. These include the following:

2.4.1 Latency

Compression and subsequent decompression of images takes time and the delay this introduces is known as latency. The degree of latency in a given system will depend on the image format (number of pixels and number of bits per pixel (bpp)), compression technique adopted (and configuration used), acceptable image quality requirements, robustness (ability to withstand errors) required and the processing power available to perform compression/decompression.

2.4.2 Loss of information and adverse effects on image quality

Some compression techniques are able to perform lossless compression i.e. produce compressed images from which it is possible (as long as data is not corrupted) to reproduce exact replicas of the original image. However, the degrees of compression available in lossless modes are almost invariably very modest (typically of the order of 2 or 3 maximum dependent on picture content). Whenever higher compression ratios are required some loss of information leading to a degree of corruption of the image is usually inevitable.

A well known example of the adverse effects on image quality is the classic 'blocking' effects seen with images that have been compressed using discrete cosine transform techniques. The designer must endeavour to determine the degree of corruption of image quality that is acceptable for the particular application e.g. high altitude reconnaissance versus real time pilotage.

2.4.3 Susceptibility to errors

Uncompressed digital video may be relatively unaffected by small numbers of errors in the data representing images. Occasional corrupted bits will usually result in a few corrupted pixels, and in many applications will have little practical effect. However, compressed digital video is much more susceptible to errors and these may have severe effects. For example a single bit error could cause large blocks of pixels to be displayed in a completely different colour to that intended, and this artefact may be extended to cover a significant number of frames.

Hence the designer must ensure that the compressed data is adequately protected against the environment (in particular electromagnetic) that the system has to endure. Protective measures will range from improving the electromagnetic performance of hardware to providing adequate error detection and correction mechanisms as part of data processing. The latter may or may not form part of a standard or family of standards for image compression/decompression.

Ref. 2 provides a useful discussion of related issues including quality of service, effects of bit errors and packet loss and forward error correction.

2.4.4 Potential problems with mixing and merging of compressed images

Mixing, merging and fusing images from different sources is fraught with potential problems that are likely to be aggravated by image compression. These include:

- Ensuring that artefacts and noise in one image do not corrupt the fused result.
- Fusion processing time.

- Differing latencies in images from different sources and/or compression/decompression systems.
- Misregistration of input images.
- Translating images to a common format for fusion and then subsequent compression after fusion.

2.4.5 Visual images in automated processing applications

Examples of automated viewing include target cueing and CGI/real image overlay. Compression may complicate certain forms of automated processing in that an uncompressed format may be required even to achieve relatively simple tasks (e.g. overlays), thus possibly requiring decompression and subsequent re-compression.

There is a need to understand the different constraints on the use of compression in different automated applications.

2.5 Trade offs

The implementation of digital video systems for military aircraft will involve trading off many of the different issues discussed above including:

- Data bandwidth/memory size.
- Compression ratio.
- Image resolution.
- Latency.
- Image update rate.
- Image quality.
- Processing power.
- Error Resilience.
- Complexity.
- Hardware/software availability.
- Functionality.
- Cost.

The term image quality is used here to imply the extent to which the image is not degraded by the compression/decompression process. Perceived image quality is also dependent upon the quality of the original image before compression. Image resolution is intended to mean the degree to which detail in the image may be resolved and is

dependent upon the number of pixels used to represent the image. Clearly image resolution and image quality both have important influences on the quality and acceptability (or otherwise) of the image as perceived by the human end user.

A better understanding of the relationships between all the issues listed above is required, and there is a need to define how they may be traded off to provide an optimum solution.

From the end users' viewpoint the essential parameters are image quality, resolution and update rate as well as latency; the other issues listed above not being of immediate concern to the end user in an operational situation.

It may be that many military avionic applications will fall into one of two categories i.e. those requiring:

1. High image quality and resolution, but with modest requirements for latency and image update rate. (e.g. aerial reconnaissance).

or

2. Low latency, high update rate, but modest requirements for image quality and resolution (e.g. video presented to the pilot of a fast jet who needs an immediate indication of circumstances and has no time to study the image in detail).

Clearly there will be other applications that require high performance in all the above areas. These applications may either need high cost investment in processing power and complexity if compression is to be used, or will need to be implemented without compression with investment being made in data bandwidth and/or in memory capacity instead.

3 COMPRESSION STANDARDS

There is a need to identify appropriate compression standard(s) for military avionic applications taking into account the issues listed above. It may be that different forms of compression will prove appropriate for different applications.

3.1 Criteria for Selection

The principal criteria used in selecting standards for consideration in this report were as follows:

- Open standard (i.e. specification is public and available for use by all)
- Maturity
- Performance
- Hardware/software availability

- Variety i.e. the desirability of considering a range of different techniques.

3.2 Standards Selected for Consideration in this Investigation

Taking into consideration the criteria listed above the following standards were selected.

- MPEG1 & 2
- MPEG 4
- JPEG
- MJPEG
- JPEG 2000
- MJPEG 2000
- H261/H263

As will be apparent from the inclusion of MPEG-4 and JPEG 2000, maturity has been treated as a desirable attribute rather than a deciding factor in the selection of standards for consideration.

Fractal compression is also of interest, but this investigation failed to reveal any existing standards for this technique.

4 CRITERIA FOR COMPARISON WITH REQUIREMENTS

A comprehensive list of criteria for comparison of the performance of compression standards with the requirements of military avionic applications would include all the issues listed in 2.5 above. However, as will be seen, the manner in which the requirements for video compression in military avionic applications are currently defined by ASSC makes comparison difficult.

5 SPECIFIC REQUIREMENTS

The requirements set out in the documents listed in Section 0 above are summarised in the following sections. Some additional applications and requirements from other sources are also summarised.

5.1 Requirements for the distribution and transmission of video images for airborne platforms (ASSC/130/2/116)

5.1.1 Data Rates

The document provides estimates of data rate requirements for the following links:

- sensor to processing

- processing to display
- direct video links from sensor to display

5.1.1.1 Sensor to Processing (Processing Link)

Data rates are derived from requirements for image resolution, frame rate and pixel clock frequency.

Data rate requirements are tabulated for thirty-two sensors with resolutions from 320x240 to 4000x4000 pixels, with frame rates from 25 to 200 frames per second and from 12 bpp to 48 bpp. From these ranges of requirements data rates (without compression) from 6.64Mb/s to 153.6Gb/s are computed. Presumably the main objective of utilising compression in these applications would be to reduce the above data rate requirements whilst retaining sufficient information for any subsequent processing.

5.1.1.2 Processing to Display (Display Link)

Data rate requirements are tabulated for displays with resolutions from 640x480 to 2048x2048 with frame rates from 30 to 75 frames per second and Pixel Clock rates of from 25 to 250MHz. From these ranges of requirements for data rates from 566Mb/s to 6Gb/s are computed (based on 24 bpp).

Again, presumably one of the objectives of utilising compression in these applications would be to reduce the data rate requirements. However, in this case only human operators will be involved and hence, provided that sufficient quality is preserved, accurate retrieval of an uncompressed image is not necessary.

5.1.1.3 Direct Video Link from Sensor to Display (Direct Link)

Typical data rate requirements are tabulated for direct sensor to video links ranging from 6Mb/s (compressed) to 270Mb/s.

Compression requirements would be similar to that for the processing to display link in Section 5.1.1.2

5.1.2 Latency

Maximum permissible overall system latencies are estimated for the following systems:

- VCS
- IRST (wide scan)
- WASAD
- MAW

- Data gathering and reconnaissance

Estimates range from <40ms to 'a few seconds'.

5.2 ASSC Advisory Document on Avionic Video Compression Technology (ASSC/130/2/77)

This document makes a distinction between the compression of images for human viewing and that for automated processing.

Examples of the former are listed as cockpit displays, remote control systems and remote surveillance systems. However, no information is given regarding the use of compression in these applications.

With regard to compression of images for automated processing, the need to prevent any detrimental effect on the main processing function is stressed. Examples of target cueing and CGI/Real Image overlay are given, but no details of requirements are provided.

5.3 Digital Video Techniques and the Possible Adoption of Commercial Video Standards in Future Military Avionic Systems (ASSC/130/2/134)

This document aims to identify key aspects, issues and problems associated with the application of digital video to military avionics, but it does not identify solutions.

With regard to compression it summarises potential application areas (as per 2.2 above). It goes on to list the potential adverse effects of compression (as per 2.4 above).

Latency is defined in the document as the time delay between the detection of a change by a sensor and its appearance on the display. This also includes persistence, i.e. the time for which a feature continues to be displayed after it ceases to be within the sensor's view. Clearly the latency associated with compression and/or decompression is part of a larger system latency that also includes delays due to transmission and other factors.

Acceptable magnitudes of overall system latency are discussed and 100ms is suggested for most human viewing applications. However, it is recognised that real time systems such as Infrared Search and Track (IRST) are particularly vulnerable to the effects of latency, and hence are likely to require much lower magnitudes.

It is suggested that consideration of the consequences of latency should include:

- Tracking errors.
- Blurring of images when sensor moved.
- Delay between aircraft manoeuvre and manoeuvre observed in the displayed sensor imagery.

5.4 Applications and Requirements Identified from other Sources

5.4.1 Discussions with an airborne video supplier

Discussions with an airborne video supplier identified some additional applications and requirements as summarised below.

5.4.1.1 Security

Concerns were expressed about the possibility of unauthorised viewing and/or use of images. Potential solutions include data encryption and/or use of private formats.

It is considered undesirable that the image should be in a standard format at any point in its digitised existence, especially where it might be liable to interception. For this reason a private format with key code generated in the codec is considered desirable.

Clearly such approaches militate against the use of open standards and are at odds with the aims of the current investigation.

5.4.1.2 Smart Image Processing for Ground Replay

Future ground replay systems for debrief, training etc. are envisaged as providing a high degree of integration between images/video from a number of aircraft together with other image content providing information such as aircraft positions in the form of graphics, maps, text etc.

Whether ground replay systems fall within the remit of ASSC is uncertain. However, it is likely that the requirements for video for use in such systems would have an impact on the requirements for airborne video systems.

5.4.1.3 Flight Deck Video Recording Application

This potential application is an extension of the 'black box' flight recorder concept using video to record events on the flight deck. It would be aimed initially at civil airliners, but could presumably also have application in military transport aircraft.

5.4.1.4 Flight Video Data Recorders (FVDR)

The FVDR system helps to record incidents and also allows the pilot to view the aircraft from the outside whilst in flight. e.g. Ref. 3 describes use of FVDR in an airliner (or military transport) where two cameras mounted in the trailing edges of the wings, plus two mounted on the rear fuselage section looking forward, allow the tail, rear of the engines, ailerons and flaps etc. to be viewed.

In the event of an incident (e.g. bird strike) the crew could immediately playback and review the video to assist in making a judgement on the probable severity of any damage

and hence a decision on whether to continue with the mission. FVDR can also be used as a maintenance aid.

The FVDR concept could also be adapted to monitor helicopter rotors and engines etc.

5.4.2 Use of MPEG-2 to Compress Unmanned Aerial Vehicle (UAV) Gathered Enhanced Definition TV Video

The abstract of a presentation (ASSC/130/2/147) given at the ASSC Video Systems Subcommittee meeting on 15 February 2001 describes the use of MPEG-2 to compress enhanced definition video acquired by UAVs.

It is anticipated that future UAVs will gather enhanced-definition (ED) and high-definition (HD) colour TV video. EDTV video is defined as colour images of at least 720 by 480 pixels, progressively scanned at rates of 50Hz or greater. HDTV video is defined as colour images of 1280 by 720 pixels, progressively scanned at 24Hz or greater. The resulting data rates of EDTV and HDTV are at least 414 Mb/s and 530 Mb/s respectively, and will overwhelm the capacity of the data links. For example the Tactical Common Data Link (TCDL) has a maximum data rate of 45 Mbit/s. Clearly the video will have to be compressed in order for it to be transmitted.

MPEG-2 video compression was considered because it is designed to be transmitted over noisy communication links, in contrast to MPEG-1, which is designed to be read from digital storage media (for example CDs and hard disk drives).

It was found that a data rate of 4Mb/s gave reasonable image quality with a Group of Pictures¹ (GOP) size of 12 and almost acceptable picture quality at 2Mb/s and a GOP size of 48 pictures. The uncompressed data rate is 497Mb/s, these data rates represent compression factors of approximately 125 and 250 respectively. The picture type was found to effect the image quality.

The communication link will be noisy and it is important to see the effect of noise on the compressed video. The effect of corrupting the compressed bitstream on the decompressed images was therefore explored. It was found there was an approximate linear relationship between the mean square error and bit error rate, although this relationship was strongest for higher data rates. The acceptable bit error rate was found to be below 10^{-6} and to be data rate dependent.

¹ Group of Pictures: a set or group of encoded frames representing part of a video sequence and starting with an I frame. Hence, generally speaking the smaller the GOP the more frequent the I frames, the lower the compression and the better the picture quality. (See also 8.3.)

5.5 COTS Requirement

A declared aim is to use Commercial Off The Shelf (COTS) hardware and software whenever possible. However, this approach brings significant risks principally in relation to environment and obsolescence. The issues needing to be addressed (but not as yet the solutions) are set out in ASSC/130/2/134.

5.6 Conclusions regarding Definition of Requirements

From the above it should be apparent that only ASSC/130/116 gives any definitive requirements that may be compared with the capabilities of various standards and an attempt to perform such a comparison is made in 6 below, subject to the limitations imposed by the requirement definitions.

6 COMPARISON OF COMPRESSION STANDARDS WITH REQUIREMENTS SET OUT IN ASSC/130/2/116

6.1 Introduction

This section attempts to compare the requirements listed in ASSC/130/116 with the compression standards listed in 3.2 above. A positive statement regarding the ability of a standard to meet requirements does not necessarily mean that implementations currently exist that also meet the requirement.

6.2 Re-statement of Requirements set out in ASSC/130/2/116 and Comparison with Standards

ASSC/130/2/116 provides estimates of data rate requirements based on image resolutions, frame rates etc. for the following links:

- sensor to processing
- processing to display
- direct video links from sensor to display

These are re-stated and compared with the constraints imposed by the video compression standards in the following subsections

6.2.1 Sensor to Processing (Processing Link)

Table 1 from ASSC/130/2/116 is reproduced below. Each sensor type has been allocated a number in the first column for ease of reference.

In comparing the requirements set out in Table 1 with various compression standards the essential parameters are:

- Image Array Size.
- Frame Rate.
- Pixel Resolution.

Pixel Resolution, expressed in bits per pixel, presents a problem since most of the standards under consideration define colour space formats (e.g. YCbCr 4:2:2). There are also complications resulting from issues such as variations in sample positions in different implementations of (for example) colour space formats 4:2:0. Furthermore the requirements do not address the issue of interlacing of fields. Hence it is not possible to make a definitive conversion from bpp to colour space format for comparison with the capabilities of various standards.

		Image Array Size									
Sensor Type		Width (pixels)	Height (pixels)	Total pixels (pixels)	Frame Rate (Frames/s)	Video Rate (Mpixels/s)	Sample rate (MHz)	Pixel Resolution (bits per pixel)	Data Rate (Mbits/s)	Timescale	Comments
1.	Thermal Camera	320	240	76800	30	2.30		14	32.26	Now	
2.	Thermal Camera	320	240	76800	60	4.61		14	64.51	Now	
3.	Thermal Camera	320	240	76800	200	15.36		14	215.04	Now	
4.	Thermal Camera	384	288	110592	25	2.76		14	38.71	Now	
5.	Thermal Camera	384	288	110592	50	5.53		14	77.41	Now	
6.	Thermal Camera	640	480	307200	30	9.22		14	129.02	Now	
7.	Thermal Camera	640	480	307200	60	18.43		14	258.05	Now	
8.	Thermal Camera	640	480	307200	200	61.44		14	860.16	Now	
9.	Thermal Camera	768	576	442368	25	11.06		14	154.83	2000	
10.	Thermal Camera	768	576	442368	50	22.12		14	309.66	2000	
11.	Thermal Camera	1280	768	983040	25	24.58		12	294.91	Now	
12.	Thermal Camera	1280	768	983040	50	49.15		14	688.13	2000	
13.	Thermal Camera	1280	768	983040	100	98.30		14	1376.26	2000	
14.	Thermal Camera	1280	1024	1310720	30	39.32		16	629.15	2003	
15.	Thermal Camera	1280	1024	1310720	60	78.64		16	1258.29	2003	
16.	Thermal Camera	1280	1024	1310720	200	262.14		16	4194.30	2003	
17.	Thermal Camera	1600	1600	2560000	30	76.80		16	1228.80	?	
18.	Thermal Camera	1600	1600	2560000	60	153.60		16	2457.60	?	
19.	Thermal Camera	1600	1600	2560000	200	512.00		16	8192.00	?	
20.	Davlight Camera	640	480	307200	30	9.22		24	221.18	Now	
21.	Daylight Camera	640	480	307200	60	18.43		24	442.37	Now	
22.	Daylight Camera	640	480	307200	200	61.44		24	1474.56	Now	
23.	Daylight Camera	768	576	442368	25	11.06		24	265.42	Now	625 line standard
24.	Daylight Camera	768	576	442368	25	11.06	13.50	30	331.78	Now	625 line, SDI, 4:4:4
25.	Daylight Camera	768	576	442368	50	22.12		24	530.84	Now	
26.	Daylight Camera	1280	1024	1310720	30	39.32		30	1179.65	2000	
27.	Daylight Camera	1280	1024	1310720	60	78.64		30	2359.30	2000	
28.	Daylight Camera	1280	1024	1310720	200	262.14		30	7864.32	2000	
29.	Daylight Camera	4000	4000	16000000	30	480.00		48	23040.00	2000	
30.	Daylight Camera	4000	4000	16000000	60	960.00		48	46080.00	2003	
31.	Daylight Camera	4000	4000	16000000	200	3200.00		48	153600.00	2003	
32.	WASAD System	768	360	276480	2	0.55		12	6.64	Now	

Table 1: Table 1 from ASSC/130/2/116 Video data rates for Sensor to Processing Link

The objective of the following sections 6.2.1.1 to 6.2.1.10 is to compare the requirements for the above parameters with the constraints/capabilities of the various standards. In the absence of definitive requirements for colour space formats, pixel resolution has been omitted from the comparisons.

The results of this comparison are shown in Table 2.

	Image Array Size Width	Image Array Size Height	Frame Rate	MPEG-1 Constrained Mode	MPEG-1	MPEG-2	MPEG-4	JPEG*	MJPEG**	JPEG- 2000***	MJPEG- 2000	H261	H263
	(nixels)	(nixels)	(Frames/s)										
1.	320	240	30	✓	✓	✓	✓	X	✓	X	✓	X	X
2.	320	240	60	X	✓	✓	✓	X	✓	X	✓	X	X
3.	320	240	200	X	X	✓	✓	X	✓	X	✓	X	X
4.	384	288	25	✓	✓	✓	✓	X	✓	X	✓	X	X
5.	384	288	50	X	✓	✓	✓	X	✓	X	✓	X	X
6.	640	480	30	✓	✓	✓	✓	X	✓	X	✓	X	X
7.	640	480	60	X	✓	✓	✓	X	✓	X	✓	X	X
8.	640	480	200	X	X	✓	✓	X	✓	X	✓	X	X
9.	768	576	25	✓	✓	✓	✓	X	✓	X	✓	X	X
10.	768	576	50	X	✓	✓	✓	X	✓	X	✓	X	X
11.	1280	768	25	X	✓	✓	✓	X	✓	X	✓	X	X
12.	1280	768	50	X	✓	✓	✓	X	✓	X	✓	X	X
13.	1280	768	100	X	X	✓	✓	X	✓	X	✓	X	X
14.	1280	1024	30	X	✓	✓	✓	X	✓	X	✓	X	X
15.	1280	1024	60	X	✓	✓	✓	X	✓	X	✓	X	X
16.	1280	1024	200	X	X	✓	✓	X	✓	X	✓	X	X
17.	1600	1600	30	X	✓	✓	✓	X	✓	X	✓	X	X
18.	1600	1600	60	X	✓	✓	✓	X	✓	X	✓	X	X
19.	1600	1600	200	X	X	✓	✓	X	✓	X	✓	X	X
20.	640	480	30	✓	✓	✓	✓	X	✓	X	✓	X	X
21.	640	480	60	X	✓	✓	✓	X	✓	X	✓	X	X
22.	640	480	200	X	X	✓	✓	X	✓	X	✓	X	X
23.	768	576	25	✓	✓	✓	✓	X	✓	X	✓	X	X
24.	768	576	25	✓	✓	✓	✓	X	✓	X	✓	X	X
25.	768	576	50	X	✓	✓	✓	X	✓	X	✓	X	X
26.	1280	1024	30	X	✓	✓	✓	X	✓	X	✓	X	X

Table 2: Comparison of Requirements for Sensor to Processing Link with Characteristics of Standards (contd next page)

	Image Array Size Width	Image Array Size Height	Frame Rate	MPEG-1 Constrained Mode	MPEG-1	MPEG-2	MPEG-4	JPEG*	MJPEG**	JPEG-2000***	MJPEG-2000	H261	H263
27.	1280	1024	60	X	✓	✓	✓	X	✓	X	✓	X	X
28.	1280	1024	200	X	X	✓	✓	X	✓	X	✓	X	X
29.	4000	4000	30	X	✓	✓	✓	X	✓	X	✓	X	X
30.	4000	4000	60	X	✓	✓	✓	X	✓	X	✓	X	X
31.	4000	4000	200	X	X	✓	✓	X	✓	X	✓	X	X
32.	768	360	?	✓	X	✓	✓	X	✓	X	✓	X	X

Table 2 : (continued): Comparison of Requirements for Sensor to Processing Link with Characteristics of Standards

Notes:

* JPEG, when subject to the constraints of the SPIFF, file format appears to be capable of meeting all of the Image Array Size requirements, but cannot meet Frame Rate requirements since JPEG is for still images only. Also SPIFF does not cater for 24bpp shown in Table 1 (see 6.2.1.5).

** Frame Rate for MJPEG will be implementation dependent.

*** JPEG-2000, when subject to the constraints of the optional file format defined in the standard appears to be capable of meeting all of the Image Array Size requirements, but cannot meet Frame Rate requirements since JPEG is for still images only.

6.2.1.1 MPEG-1 Constrained Parameter Mode

MPEG-1 has a Constrained Parameter Mode intended to promote compatibility as follows:

- Horizontal resolution - up to 768 pixels
- Vertical resolution - up to 576 lines
- Frame rate - up to 30Hz
- Colour space format - YCbCr 4:2:0

Comparison of the above with Table 1 shows that MPEG-1 in Constrained Parameter Mode fails to meet the requirements for image array size and frame rate for most of the applications listed. Those that are met are indicated in Table 2 thus '✓'.

6.2.1.2 MPEG-1

Examination of the video bit stream specification for MPEG-1 set out in Ref. 4 reveals the following limits on relevant parameters:

- Horizontal resolution - defined by Horizontal_size field having 12 bits and thus implying a maximum of 4095 pixels.
- Vertical resolution - defined by Vertical_size having 12 bits and thus implying a maximum of 4095 pixels.
- Frame rate - defined by 4 bits half of the possible bit patterns being reserved and leaving rates of 23.976, 24, 25, 29.97, 30, 50, 59.94 and 60.

The above constraints on parameters have been compared with the requirements in Table 2 and those that are met are indicated thus '✓'. It can be seen that nearly all of the listed requirements are met by MPEG-1, the exceptions being those with very high (100 and 200 FPS) or very low (2 FPS) frame rates.

Ref. 4 lists 'common MPEG-1 resolutions' as follows:

1. 352 x 240 at 29.97 FPS
2. 352 x 240 at 23.976 FPS
3. 352 x 288 at 25 FPS
4. 320 x 240 at 29.97 FPS
5. 384 x 288 at 25 FPS

4 and 5 are for square pixel resolution.

Comparing the above with Table 2 it is apparent that implementations that provide only the listed 'common MPEG-1 resolutions' would only meet requirement No.4.

6.2.1.3 MPEG-2

Examination of the video bit stream specification for MPEG-2 set out in Ref. 4 reveals the following limits on relevant parameters:

- Horizontal resolution - defined by Horizontal_size_value and Horizontal_size_extension having 12 bits and 2 bits respectively and thus implying a maximum of $2^{14}-1 = 16,383$ pixels.
- Vertical resolution - defined by Horizontal_size_value and Horizontal_size_extension having 12 bits and 2 bits respectively and thus implying a maximum of $2^{14}-1 = 16,383$ pixels.
- Frame rate - defined by Frame_rate_code, Frame_rate_extension_n and Frame_rate_extension_d having 4, 2 and 4 bits respectively. When the Frame_rate_extension values are zero rates of 23.976, 24, 25, 29.97, 30, 50, 59.94 and 60 are defined. A formula is provided from which it appears that non-zero Frame_rate_extension values encompass the 200 FPS requirement in Table 2, but not 2 FPS.

However, horizontal and vertical resolutions must be in multiples of 16 for frame coded pictures and vertical resolution must be in multiples of 32 for field coded pictures.

The above constraints on parameters have been compared with the requirements in Table 2 and those that are met are indicated thus '✓'. It can be seen that all the listed requirements are met by MPEG-2, with the exception of No. 32 which fails both in terms of permissible frame rate and the requirement for resolution to be a multiple of 16.

6.2.1.4 MPEG-4

MPEG-4 supports the coding of conventional images and video in a similar manner to MPEG 1 and 2. However, MPEG-4 also provides a set of technologies to support the coded representation of units of aural, visual or audiovisual content, called "audio-visual objects" or AVOs, together with means of manipulating AVOs. The standard also permits multiple codecs to be used to represent different objects within the image.

However, this obviously has implications for the complexity of the decompression algorithms.

Under the heading of Formats Supported Ref. 5 lists the following formats and bitrates that will be supported by MPEG-4 Version 1:

- Bitrates: typically between 5 kbit/s and 10 Mbit/s.
- Formats: progressive as well as interlaced video.
- Resolutions: typically from sub-QCIF to beyond HDTV.

It appears that MPEG4 can be used at any resolution between the extremes of QCIF and HDTV, since one of the constraints of MPEG2 i.e. that the borders must be straight, is removed when arbitrary objects must be considered, and hence there is no longer any size constraints on images.

An exchange of e-mails with a leading member of MPEG elicited a statement that there are no constraints on frame rate in MPEG-4. Hence MPEG-4 has been shown as meeting all the requirements in Table 2.

6.2.1.5 JPEG

Since JPEG is designed to handle still images Frame Rate is not specified in the standard. Furthermore JPEG refers only to a family of compression algorithms; it does not refer to a specific image file format.

A de-facto standard for JPEG files is the JPEG File Interchange Format (JFIF) originated by C-Cube Microsystems.

More recently the ISO JPEG committee has defined a complete file format specification called Still Picture Interchange File Format (SPIFF) in the new "Part 3" extensions to the JPEG standard. SPIFF is upward compatible with JFIF.

SPIFF defines a parameter 'HEIGHT: Image height – the number of lines in the highest component of the image.' Being a 32 bit unsigned integer, values 1 – 4,294,967,295.

SPIFF defines a parameter 'WIDTH: Image width – the number of samples per line in the widest component of the image.' Being a 32 bit unsigned integer, values 1 – 4,294,967,295.

Clearly the ranges of these parameters are more than adequate to accommodate the image array sizes shown in Table 2.

However SPIFF also defines a parameter BPS: Bits Per Sample – the number of bits per sample for the components of the image.' Having permissible values of 1, 2, 4, 8, 12, 16. Hence if the reference to 24 bits/pixel in the Data rate column of Table 1 is taken as a

requirement then it cannot be met by SPIFF. However, if the reference to 24 bits/pixel is taken as an example value inserted to allow data rate to be computed, and ignored for purposes of this comparison then SPIFF could be said to be capable of handling all the display modes listed in Table 1 and Table 2 except for Frame Rate.

6.2.1.6 MJPEG

MJPEG is not covered by an international standard, but is the subject of various proprietary or de facto standards. Since MJPEG consists of transmitting a sequence of JPEG still images it is subject to the same constraints as JPEG, see 6.2.1.5, and cannot take advantage of any commonality between frames during compression. Like JPEG, MJPEG may meet all of the Image Array Size requirements.

Frame rate is likely to be implementation specific and will, among other things, depend on the processing power available. If we assume an implementation capable of achieving the frame rate requirements in Table 2 then MJPEG may be said to meet all the requirements. However, the higher frame rate requirements (e.g. 200 FPS) may prove a challenge to the implementor.

Also see 8.7.5 for further discussion.

6.2.1.7 JPEG-2000

The JP2 file format specified for optional use with JPEG-2000 defines a 4 byte unsigned integer field for image height and width and hence these may be considered virtually unlimited for practical purposes. However, since JPEG-2000 is for still images it cannot by definition meet the frame rate requirements.

6.2.1.8 MJPEG-2000

MJPEG-2000 (Motion JPEG-2000) will form Part 3 of the standard and will, optionally use the same JP2 file format as JPEG-2000. Hence image height and width may be considered virtually unlimited for practical purposes. However, it is not clear what constraints apply to frame rate and this is likely to be implementation dependent. If we assume an implementation capable of achieving the frame rate requirements in Table 2 then MJPEG may be said to meet all the requirements. However, the higher frame rate requirements (e.g. 200 FPS) may prove a challenge to the implementor.

6.2.1.9 H261

H261 handles image sizes of 352x288 (CIF) and 176x144 (QCIF) with a frame refresh rate of 29.97Hz only. Hence H261 fails to meet any of the requirements listed in Table 2 for either image size or frame rate.

6.2.1.10 H263

H263 handles image sizes of 1408x11152 (16CIF), 704x576 (4CIF), 352x288 (CIF), 176x144 (QCIF) and 128x96 (SQCIF). Surprisingly none of these image sizes align with any of those listed in Table 2. Hence H263 fails to meet any of the listed requirements.

Furthermore H263 handles only one frame rate: 29.97 FPS and so fails to meet any of the frame rate requirements.

6.2.2 Processing to Display (Display Link)

Table 2 from ASSC/130/2/116 is reproduced below as Table 3 of this report, each display mode has been allocated a number in the first column for ease of reference.

	Mode	Definition (H x V)	Frame Rate (Hz)	Pixel clock (MHz)	Data rate (Mb s⁻¹) (@ 24 bits/pixel)
1.	VGA	640x480	60	25	600
2.	VGA	720x400	70	28	672
3.	SVGA	800X600	60	40	960
4.	XGA	1024X768	30	23.59	566.16
5.	XGA	1024X768	60	65	1560
6.	SXGA	1280X1024	30	40	672
7.	SXGA	1280X1024	60	112	2688
8.	UXGA	1600X1200	75	250	6000
9.	VXGA	2048X2048	30	120	2888
10.	HDTV	1280X720	30	82.92	663.36
11.	HDTV	1280X720	60	77.54	1860
12.	HDTV	1820X1080	30	62.2	1492.9
13.	HDTV	1820X1080	30(i)	77.2	1852.8

Table 3: Table 2 from ASSC/130/2/116 Requirements for Processing to Display Link

Presumably the objective of compression in these applications would be to reduce the Data rate requirement shown in the last column of Table 3. Hence the applicability of various standards may be judged in terms of Definition and Frame Rate requirements.

The objective of the following sections 6.2.2.1 to 6.2.2.9 is to compare the requirements for the above parameters with the constraints/capabilities of the various standards. The results of the comparison are set out in Table 4.

Mode	Definition (H x V)	Frame Rate (Hz)	MPEG-1 Constrained	MPEG-1	MPEG-2	MPEG-4	JPEG* (SPIFF)	MJPEG	JPEG- 2000	MJPEG- 2000	H261	H263
VGA	640x480	60	X	✓	✓	✓	X	✓	X	✓	X	X
VGA	720x400	70	X	X	X	✓	X	✓	X	✓	X	X
SVGA	800X600	60	X	✓	✓	✓	X	✓	X	✓	X	X
XGA	1024X768	30	X	✓	✓	✓	X	✓	X	✓	X	X
XGA	1024X768	60	X	✓	✓	✓	X	✓	X	✓	X	X
SXGA	1280X1024	30	X	✓	✓	✓	X	✓	X	✓	X	X
SXGA	1280X1024	60	X	✓	✓	✓	X	✓	X	✓	X	X
UXGA	1600X1200	75	X	X	X	✓	X	✓	X	✓	X	X
VXGA	2048X2048	30	X	✓	✓	✓	X	✓	X	✓	X	X
HDTV	1280X720	30	X	✓	✓	✓	X	✓	X	✓	X	X
HDTV	1280X720	60	X	✓	✓	✓	X	✓	X	✓	X	X
HDTV	1820X1080	30	X	✓	✓	✓	X	✓	X	✓	X	X
HDTV	1820X1080	30(i)	X	✓	✓	✓	X	✓	X	✓	X	X

Table 4: Comparison of Requirements for Processing to Display Link with Compression Standards

6.2.2.1 MPEG-1 Constrained Parameter Mode

Comparison of the constrained parameter mode limits set out in 6.2.1.1 above with the requirements in Table 4 shows that MPEG-1 in Constrained Parameter Mode fails to meet any of the requirements for image array size and/or frame rate for any of the applications listed.

6.2.2.2 MPEG-1

The limits on parameters set out in 6.2.1.2 have been compared with the requirements in Table 4 and those that are met are indicated thus '✓'. It can be seen that MPEG-1 can handle all the image resolutions listed and all but two of the frame rates i.e. 70 and 75 Hz.

6.2.2.3 MPEG-2

The limits on parameters set out in 6.2.1.3 have been compared with the requirements in Table 4 and those that are met are indicated thus '✓'. It can be seen that like MPEG-1, MPEG-2 can handle all the image resolutions listed and all but two of the frame rates i.e. 70 and 75 Hz.

6.2.2.4 MPEG-4

See 6.2.1.4 above.

6.2.2.5 JPEG

Comparison of the limits of SPIFF set out in 6.2.1.4 with the requirements in Table 4 shows that the ranges of these parameters are more than adequate to accommodate all the image array sizes shown in Table 4.

However SPIFF also defines a parameter BPS: Bits Per Sample – the number of bits per sample for the components of the image.' Having permissible values of 1, 2, 4, 8, 12, 16. Hence if the reference to 24 bits/pixel in the Data rate column of Table 1 is taken as a requirement then it cannot be met by SPIFF. However, if the reference to 24 bits/pixel was taken as an example value inserted to allow data rate to be computed, and was ignored then SPIFF could be said to be capable of handling all the display modes listed in Table 1 and Table 2.

6.2.2.6 MJPEG

Like JPEG, MJPEG may meet all of the Image Array Size requirements.

Frame rate is likely to be implementation specific and will, among other things, depend on the processing power available. If we assume an implementation capable of achieving the frame rate requirements in Table 2 then MJPEG may be said to meet all the

requirements. However, the 70Hz frame rate requirement may prove a challenge to the implementor.

Also see 8.7.5 for further discussion.

6.2.2.7 JPEG-2000

The JP2 file format specified for optional use with JPEG-2000 defines a 4 byte unsigned integer field for image height and width and hence these may be considered virtually unlimited for practical purposes. However, since JPEG-2000 is for still images it cannot by definition meet the frame rate requirements.

6.2.2.8 MJPEG-2000

As already state according to the JP2 file format image height and width may be considered virtually unlimited for practical purposes. Again it is not clear what constraints apply to frame rate and this is likely to be implementation dependent. If we assume an implementation capable of achieving the frame rate requirements in Table 4 then MJPEG may be said to meet all the requirements. However, the 70Hz frame rate requirement may prove a challenge to the implementor.

6.2.2.9 H261

Since H261 handles image sizes of 352x288 (CIF) and 176x144 (QCIF) with a frame refresh rate of 29.97Hz, none of the requirements listed in Table 4 can be met by H261.

6.2.2.10 H263

Similarly the image sizes and frame rates specified for H263 (see 6.2.1.10 above) are such that it cannot meet any of the requirements in Table 4.

6.2.3 Direct Video Link from Sensor to Display (Direct Link) ASSC/130/2/116

Typical data rate requirements are tabulated for direct sensor to video links ranging from 6Mb/s (compressed) to 270Mb/s (presumably uncompressed). There is no scope for comparing this information with the capabilities of compression standards. Hence these requirements are not considered further, although the issues involved are likely to be similar to those for the processor to display link discussed in Section 6.2.2.

6.3 Latency

ASSC/130/2/116 provides estimates of maximum permissible overall system latency for the following systems:

1. VCS pilot aid
 - Overall system latency : <40 ms
2. IRST [wide scan]
 - Overall system latency [to declaration] : a few seconds typical
 - Overall system latency [for visual display] : one frame typical
 - Distribution Latency : a few lines typical
3. WASAD system
 - Overall system latency [to declaration] : one second typical
 - Overall system latency [for visual display] : one frame typical
 - Distribution Latency : a few lines typical
4. MAW system
 - Overall system latency [to declaration] : much less than one second typical
 - Distribution Latency : probably less than one line typical
5. Data gathering/ reconnaissance
 - Overall system latency (to recorder) : not important (probably several frames acceptable, particularly if compression as per MPEG-2 is used)
 - Overall system latency (to download depends - off aircraft) : possibly as above (however on requirement, some may require near real time attributes).
 - Distribution Latency : consistent with not blocking other traffic that is more important.

Latency is highly implementation dependent. However, some generalised points regarding the relative performance of various standards can be made. See Section 8 for further discussion.

7 WIDER CONSIDERATION OF REQUIREMENTS VERSUS STANDARDS

7.1 Tabulation of requirements versus capabilities

Having, insofar as is possible, compared the requirements set out in ASSC/130/2/116 with the characteristics of various compression standards, this section discusses issues

that would need to be addressed in order to perform more in depth comparisons, given adequate definitions of requirements.

The following (as yet incomplete) Table 5 is an example of a pro-forma that might be used as to summarise requirements for video compression in various applications against the capabilities of various compression standards.

The main advantage of this approach is that it provides a checklist and a structured and regulated comparison of requirements and capabilities.

One disadvantage of this approach is that it does not readily allow for the various trade-offs that can be made between various performance aspects as discussed in 2.5 above. Furthermore many of the issues listed may not be readily amenable to a brief entry in a table cell, but this could be overcome by using separate (and if necessary larger) sheets/tables for each standard. However, despite these disadvantages it is probably still worth exploring the utility of this approach in order to exploit its advantages.

(Ref. 1 sets out the findings of a follow-on task in which the requirements for digital video systems for Uninhabited Aerial Vehicles were compared with the capabilities of MPEG-2 and efforts made to fill out a version of Table 5 for this combination of application and compression standard.)

Application:-		VCS Pilot Aid								
Compression Standard:->		MPEG-1	MPEG-2	MPEG-4	JPEG	MJPEG	JPEG-2000	MJPEG2000	H261	H263
Requirement										
Image Format	SVGA 800x600									
Frame Rate										
Compression Ratio (lossless)										
Compression Ratio (lossy)										
Latency	<30ms									
Image Quality										
Data Rate Mb/s	960									
Error Resilience										
Functionality										
Processing Power										
Complexity										
Maturity										
Cost										
Hardware/soft-ware										
Comments										

Table 5: Application Requirements versus Characteristics of Standards

When attempts are made to fill in tables like Table 5, basing requirements on the sources referenced in Section 0 above, the lack of relevant requirement definition in those sources becomes even more apparent. For example, ASSC/130/2/116 defines the following:

1. data rates for various links (sensor to processing, processing to display and direct video links) for various image formats and frame rates
2. estimates of latency requirements for various systems (VCS, IRST etc.).

However, no connection is provided between 1 and 2. e.g. whilst estimates of latency are provided for systems such as VCS, IRST etc., no indication is given as to the image formats, frame rates etc. used in these systems.

Without this information it is only possible to provide a generalised discussion of the features of various standards in terms of the likely requirements in military avionic systems and this is set out in 8 below.

However, it may be useful to first briefly explore the issues associated with definition of the various topics listed in Table 5 for various applications, and this is attempted in the following sections.

7.2 Table 5 Entries

The following sub-sections discuss some of the issues involved in trying to provide entries for the various cells in Table 5. The discussion concentrates mostly on definition of requirements for a given application in the column headed 'Requirement'.

7.2.1 Application

ASSC/130/2/116 names five different avionic systems (or applications) in what is described as a non-exhaustive list of examples. These are VCS pilot aid, IRST (wide scan), WASAD, MAW and data gathering/reconnaissance. ASSC/130/2/134 provides what is described as a crude summary of digital video applications (that may or may not require compression) including generalised topics such as image enhancement, surveillance, navigation, debrief etc.

There is clearly a need to define a set of applications that involve compression. It should be borne in mind that there may be significant variations in the detailed requirements for different implementations of the same generalised application. However, it should be possible to define a typical set of requirements with which the characteristics of various compression standards can be compared. It is possible that in some cases all that can be achieved is a 'best guess' at likely requirements made by suitably qualified and experienced parties.

7.2.2 Compression Standard

Table 5 includes the set of standards that have been initially proposed for this investigation. However, it is possible that this might need to change in the light of any further investigation. For example as the requirements for various applications are defined in more detail, it might become apparent that the set of standards initially chosen does not provide the required performance and other standards may need to be investigated.

7.2.3 Requirement - Image Format

Image formats (for sensors and displays) might be relatively easy to define if it is possible to discover the formats used in current systems or systems due to go into service in the next few years. However, this assumes that these decisions have been made and that the information is available, it might be necessary to return to first principles and estimate the required image format for various applications.

This definition needs to cover both numbers of pixels and the colour space specification.

7.2.4 Requirement - Frame Rate

Frame rates (for sensors and displays) might be relatively easy to define if it is possible to discover the formats used in current systems or systems due to go into service in the next few years. However, this assumes that these decisions have been made and that the information is available, again it might be necessary to return to first principles and estimate the required frame rates for various applications.

7.2.5 Requirement - Compression Ratio

In some cases this requirement may be relatively easy to fix once other requirements have been fixed. For example if image format and frame rate are defined and a given fixed bandwidth is available for transmission of the digital video data. However, compression ratio is more likely to be part of a complex multi-way trade off between a number of the topics listed under 'Requirement' in Table 5. See also 2.5 above.

7.2.6 Requirement - Lossless versus lossy compression

This is a fundamental issue in that if lossless compression is essential then typically only very modest compression ratios will be achievable. This issue is closely related to image quality since only lossless compression can provide a guaranteed retrievable image quality (i.e. the same as before compression). Once lossy compression is accepted, we are immediately immersed in the complex issue of assessment of image quality, see 7.2.8 below.

7.2.7 Requirement - Latency

ASSC/130/2/116 gives some estimates of latency requirements for various avionic systems (listed in 6.3 above). However, these are couched in terms such as 'one frame typical' and 'a few lines typical' and need to be related to frame rate and image format etc.

The above document also gives a simplified example of a VCS system and demonstrates that in this case the overall system latency (the only latency that matters to the end user) is made up of at least four elements i.e.:

1. Control Demand and Turret Inertia
2. Sensor to Processing
3. Processing
4. Processing to Display

It is clear that compression and decompression, if used, cause only parts of the overall system latency, and for the purposes of the current investigation needs to be separately identified and quantified. Furthermore, in the case of the VCS system shown as an example in ASSC/130/2/116, it is not clear what part, if any, compression plays in the overall process. This raises a fundamental issue for selection of example systems with regard to requirements for video compression. i.e. is compression appropriate at all and if so what is its role?

7.2.8 Requirement - Image Quality

This seems to be a particularly difficult issue, potentially involving definition of requirements, prediction of probable performance and assessment of actual performance.

Firstly the definition of image quality requirements for a particular application is likely to prove difficult and will need to take into account the principal objectives and purposes of the system under review. e.g. high altitude photo reconnaissance and real time pilotage will almost certainly have very different requirements for image quality.

Secondly predicting the likely actual quality of the pictures produced by various standards (not to mention various implementations and configurations) for various image subject matters may also prove difficult, particularly since in many cases it will be both scene and movement dependent.

Thirdly assessment of displayed image quality is often a very subjective and a poorly defined process. Those needing to assess the quality of digital video sometimes resort to using a panel of human viewers.

Methods for objective assessments of image quality are available, based on mathematical formulas that provide a numerical measure based on the difference between an original image and one that has been processed. Such methods have the advantage that they offer a definitive means of comparing the performance of different compression techniques and implementations. The disadvantage is that mathematical measures do not always indicate whether an image is acceptable to human observers. Ideally this subjective judgement should be made by the end-users under the actual viewing conditions and environment that apply in practice.

A widely used image quality measurement is the PSNR (peak signal to noise ratio). This relates to the sum of the squared differences between corresponding pixels of two images. Ref. 6 defines PSNR as follows:

$$\text{PSNR} = 20 \log_{10} \left[\frac{255}{\left[\frac{1}{\text{rows} \times \text{cols}} \sum_{i=1}^{\text{rows}} \sum_{j=1}^{\text{cols}} (P_{i,j} - Q_{i,j})^2 \right]^{1/2}} \right]$$

Where $P_{i,j}$ is a pixel in row i column j of image P and $Q_{i,j}$ is a pixel in row i column j of image Q . Clearly PSNR becomes smaller as the difference between P and Q grows

Under normal circumstances PSNR is considered a good indicator of image quality, but it has been shown to have serious potential limitations. For example Ref. 6 shows two images, one of which has an overall poor ‘noisy’ quality while the other is of much higher clarity and crispness except for a number of horizontal bands of severe interference that blot out the centre of the image. Both pictures have identical values of PSNR.

For current purposes it may be necessary to rely on the ‘best guesses’ of persons with appropriate experience in defining both requirement and performance. Indeed this unsatisfactory solution may also have to be applied in a number of other areas.

7.2.9 Requirement - Data Rate

This parameter could be applied to a fixed available data rate for transmission of compressed digital video in a given application, that might dictate the degree of compression required. Alternatively the term could apply to the data rate needed for an application using a certain compression technique, given other parameters such as image format, frame rate etc. Again we are likely to find ourselves in the realm of the complex multi-way trade offs.

Similar arguments might apply to memory size and sequence duration in the case of a video storage system.

7.2.10 Requirement - Error Resilience

Error Resilience may be defined as the ability of the coding system to limit the adverse effects of errors on image quality. This may be expressed in terms of PSNR v Bit Error Rate (BER). Hence to assess the error resilience of a coding technique in a given application it is first necessary to define the probable BER of the data that conveys the coded image(s) in the application. This would entail definition of (for example) the network standard used to convey the data representing the image/video, plus an estimate of its BER in military avionic applications.

7.2.11 Requirement - Functionality

Functionality is intended to define the specific features required for a given application versus the set of features offered by a given standard.

Ref. 7 includes under this term:

- Lossless compression capability.
- Lossy compression capability.
- Region of interest (ROI) coding.
- Arbitrary shaped objects.
- Random access.
- Low complexity.
- Non-iterative rate control.
- Genericity (the ability to efficiently compress different types of imagery across a wide range of bit rates).

A more generalised definition of functionality would be simply the provision of means by which the user may tailor the system to the specific needs of the application.

7.2.12 Requirement - Maturity

Traditionally the requirement for maturity would be taken to mean that the standard under consideration must not be so new that it might soon disappear due to lack of support, but not so old that support for it (especially in terms of available hardware and software) may soon cease.

If the benefits of digital video are to be exploited in military avionics then it will be essential to tap into current non-military sources of digital video technology.

However, military avionics, where system lifetimes are measured in decades, contrasts markedly with the telecommunications/commercial/entertainment/domestic worlds which

have spawned virtually all current digital video technology, and where product 'operational' lifetimes may be measured in a few years.

Bridging the gap between these two worlds will be a difficult task, particularly since it is most unlikely that the military world will be able to exert any influence whatever upon the producers of modern digital video technology.

7.2.13 Requirement - Hardware/software availability

The comments above with reference to maturity mostly apply equally to hardware and software availability.

7.2.14 Remaining Cells

The remaining cells in Table 5 are intended to be used to define the characteristics of candidate standards versus the requirements entered in the 'Requirement' column.

Given the very incomplete state of the requirement definition, the best that can be done at this juncture is to investigate the probable merits and demerits of various standards and that process is attempted in 8 below.

8 DISCUSSION OF CANDIDATE COMPRESSION TECHNIQUES AND STANDARDS

8.1 General

The following sub-sections contain generalised discussions of compression techniques, followed by discussions of the features of the candidate standards.

8.2 Compression Techniques

Brief summaries of the principal features of various compression techniques and standards are given in ASSC/130/2/77 (now integrated into ASSC/130/2/97 as chapter 4). Some additional discussion of the features of each standard is included below with an examination of the issues that may render various standards appropriate or otherwise for different avionic applications.

8.2.1 Intraframe versus Interframe coding

Compression (or coding) techniques may be divided into two very general categories intraframe and interframe.

Intraframe coding techniques code or compress each image or frame separately without any consideration of the contents of preceding or succeeding frames, if there are any. Intraframe coding takes advantage of spatial redundancy within each frame and examples include discrete cosine transform (e.g. JPEG), wavelet (JPEG 2000) and fractal

compression (this investigation has not identified a standard for fractal compression). Since each frame is processed separately there is no opportunity to take advantage of temporal redundancy, and compression ratios may be less than those achieved with interframe coding, also the processing task may be much more onerous (e.g. as in fractal encoding). Intraframe coding techniques are particularly appropriate to the capture of still images or slow sequences of images. However, they may also be used to compress real time video by simply treating it as a sequence of still images.

Interframe coding utilises temporal redundancy between frames i.e. it takes advantage of the often extensive commonality in the content of a sequence of frames. Standards such as MPEG utilise both intraframe and interframe coding and hence may, depending on image and sequence content, achieve higher levels of compression than intraframe only coding. However, they are less suitable for processing single images, or sequences where the image content changes greatly from frame to frame, e.g. fast changing images as might be captured from a fast jet flying at low altitude, or slow frame rate images that also have large changes in content.

Interframe coding has another significant disadvantage in that it operates on a sequence of frames. This implies that at the very start of the encoding/compression process there must be a delay while a number of frames are captured before coding can commence. Once that delay (or latency) has occurred it will not be possible to 'catch up', and the latency due to the time taken to capture the initial sequence of frames will be maintained thereafter. Further latency, due to the time needed to process the data representing the images, will also be added to these delays to give a total value of latency due to the complete capture and encoding and decoding process. As already discussed this will form part of a longer overall system latency. See 8.3.5 below for a discussion of latency with specific regard to MPEG.

Techniques that rely only on intraframe coding are not subject to this initial delay and may therefore, depending on the complexity of the technique and the processing power available, provide lower latency.

8.2.2 Techniques for Intraframe Coding

8.2.2.1 Discrete Cosine Transform, as applied by JPEG

A widely used method of spatial encoding is known as Discrete Transform Coding (DCT). This approach transforms the data into a frequency domain so that frequencies with low amplitudes can be eliminated to reduce the volume of data. DCT is used by JPEG and MPEG-1 and 2, for the intraframe coding part of the encoding process.

Ref. 8 provides an introduction to JPEG as follows:

The JPEG standard, which supports both lossy and lossless compression, specifies four modes of operation: sequential, lossless, progressive and hierarchical. The most widely used is the baseline mode, which is a subset of the sequential encoding based on the Discrete Cosine Form (DCT). The three latter modes are essentially an extension in order to cover a wider range of applications, and not used as extensively.

The general encoding algorithm for the baseline mode is as follows:

- The image is transformed to an optimal colour space. The algorithm is capable of encoding images that use any form of colour space (RGB, HSI, CMY, YUV, YCbCr) but best results are achieved with YCbCr.
- Some components are downsampled by averaging groups of pixels together.
- Each component of the source image is divided into 8x8 pixels of non-overlapping blocks. A complex DCT is applied to each block that converts the spatial image into a frequency map.
- Each DCT block is divided by a "quantisation coefficient" and the result rounded to an integer. It is this step where most of the visual information is discarded and hence is controlled by the quality setting.
- The quantised coefficients are ordered in a zigzag manner and then Huffman or arithmetic encoding is applied to remove redundant data.

Huffman or arithmetic coding is used as the final output of compressed data. The image quality will be the same in both cases but arithmetic coding usually produces a 5% to 10% smaller file.

8.2.2.2 Wavelet Compression

Ref. 9 describes wavelet compression as using a tree of filters and decimators working on the image in both horizontal and vertical directions. The luminance and chrominance components of the image in both axes are filtered into successively narrower bands and decimated (i.e. the number of samples is halved) at each stage. No compression has been achieved at this stage, the image has simply been broken down into a range of frequency bands. The resulting set of data is then passed through an adaptive quantiser, this element compresses the filtered image by exploiting the characteristics of the human visual system. The eye generally resolves high frequency elements of images to a lower level than low frequencies. Application of 'intelligent quantisation' of the filtered image achieves significant levels of compression without unduly degrading the image quality. The final stage in the compression process is to entropy encode the data set resulting from the quantisation process. This is typically achieved using run length encoding.

This process is considerably aided by the fact that the filtering and quantisation processes increase the numbers of zeros and strings of zeros in the data.

An industry source has claimed compression ratios of 65:1 on 640x480 resolution video for front line pilot debrief where pin point sharpness is not required. It is also claimed that hardware wavelet compression can readily achieve latencies below the 40ms discussed in 5.1.2 above.

8.2.2.3 Fractal compression

Fractal images are composed of repeated patterns on an ever decreasing scale. The fractal technique can be used to represent some natural objects, most famously a fern where each fernlet is supposedly a copy of the whole fern on a reduced scale, an assertion likely to be vigorously disputed by a botanist. The technique can also be used to produce naturalistic looking images of trees and certain types of coastline, as viewed from a satellite. Fractal techniques can be used to produce 'made up' representations of almost any image and, since they rely on similar features in different orientations and at different sizes, they can readily be compressed into a limited number of coefficients giving very high levels of compression when compared to a conventional digital image of the picture. Levels of compression of 10,000:1 have been claimed for images which are naturally favourable to the technique. In practice the big problem appears to be how to apply the technique to representing real life pictures rather than 'made up' ones. There has certainly been some success in this approach but the levels of compression achieved are far more modest - typically in the range 8:1 to 50:1. Furthermore considerable volumes of computation are needed to code fractal images leading to high latencies.

This investigation has not identified any standards for fractal compression and this technique has not therefore been considered further.

8.2.3 Closed Systems

Open systems are usually desirable in commercial and domestic applications since they may allow different implementations of encoder and decoder to work together successfully. However, this means that additional information must be transmitted with the encoded data stream (e.g. details of quantization tables etc.) in order for the decoder to decompress the image. For JPEG this data is normally conveyed in a header typically of 500 to 600 bytes.

MPEG-2 can also transfer additional coding information when it may be necessary to encode/decode several times in order to perform editing.

Some military applications might with advantage be implemented as closed systems if the encoder and decoder are under the control of the system designer and predefined defaults

used, removing the need for a header to carry additional information. Use of this 'abbreviated format' allows higher compression to be achieved. However, this approach severely restricts the portability of files.

8.3 MPEG-1

8.3.1 Overview

The Motion Picture Experts Group (MPEG) defined MPEG-1 primarily for the transfer of data at a rate of approximately 1.5 Mb/s. Thus allowing video to be replayed in real-time from CD-ROMs.

MPEG-1 is formally defined in ISO/IEC 11172. This defines a multiplexed structure for combining audio and video data and the means of representing the timing information needed to replay synchronised sequences in real-time. The standard specifies the coded representation of video and audio respectively, as well as the decoding process required to reconstruct pictures and audio. Audio is beyond the scope of this report and will not be considered further. However, it may be that ASSC should review the relevance of digital audio to military avionics.

8.3.1.1 MPEG Picture Types

The MPEG standard specifically defines three types of pictures (frames):

- Intra
- Predicted
- Bidirectional

8.3.1.2 Intra Pictures

Intra (or I) pictures, are coded using information only found in the image itself i.e. intraframe coded. I-frames provide potential random access points into the compressed video data. I-frames only use transform coding and provide moderate compression ratios, typically to 2 bits per coded pixel.

8.3.1.3 Predicted Pictures

Predicted (or P) pictures are interframe coded with respect to the nearest previous I or P-frame. This technique is called forward prediction. Like I-frames, P-frames serve as a prediction reference for B-frames and future P-frames. However, P-frames use motion compensation to get a higher compression ratio than is possible for I-frames. Since P-frames are predicted from previous reference (P or I) frames, they can propagate coding errors.

8.3.1.4 Bidirectional Pictures

Bidirectional (or B) pictures use both a past and future image as a reference. This technique is called bidirectional prediction. B-frames provide the most compression and do not propagate errors because they are never used as a reference. Bidirectional prediction also reduces the effects of noise by averaging two pictures.

8.3.2 Video Stream Composition

The MPEG algorithm allows the encoder to choose the frequency and location of I-frames. This choice may be dependent on the application's need for random access and the location of scene cuts in the video sequence. In applications where random access is important, I-frames are used twice per second.

The encoder also chooses the number of B-frames between any pair of reference (I or P) frames. This choice is based on factors such as the amount of memory in the encoder and the characteristics of the material being encoded.

The MPEG encoder re-orders pictures in the video stream to present the pictures to the decoder in the most efficient sequence. In particular, the reference pictures needed to reconstruct B-frames are sent before the associated B-frames.

8.3.3 Comparison with Requirements from Table 5

Image Format and Frame Rate have already been discussed in 6 above.

8.3.4 Compression Ratio

8.3.4.1 Lossless Compression

MPEG-2 does not provide lossless compression.

8.3.4.2 Lossy Compression

Achievable compression ratios will depend greatly on subject matter, image quality requirements and configuration etc. Compression ratios are very application/implementation dependent, but are typically (and very approximately) of the order of 100:1.

One configuration issue that will effect compression is the selection of the various coded frame types in a GOP. Ref. 4 provides estimates of the approximate numbers of bits per pixel for the various frame types as follows:

- I frame ~1 bpp
- P frame ~0.1 bpp

- B frame ~0.015 bpp

From the above it is clear that I frames will provide far less compression than P and B and that the overall compression achieved will depend very much on the make up of the GOP selected by the user.

8.3.5 Latency

MPEG uses interframe coding as discussed in 8.2.1 above and is thus subject to its limitations, in particular with regard to latency.

Before encoding can commence an MPEG encoder must be presented with the first I frame and first P frame, from which to encode the intervening B frames.

i.e. the latency (in seconds) due to the time taken to capture an initial sequence of frames

$$= \frac{\text{Number of B-frames between the first I-frame and first P-frame} + 2}{\text{Frame rate (FPS)}}$$

Hence this latency is dependent on the make up of the GOP. However, as stated above, the MPEG algorithm allows the encoder to choose the frequency and location of I-frames, and by implication the 'acquisition latency'.

This particular latency problem can be overcome by configuring MPEG to perform intraframe coding only. i.e. to produce I frames only. However, the loss of temporal compression results in a degradation of compression capability. With a fixed data rate, this is manifested in a severely reduced frame rate.

Processing or 'pipeline' latency, the time taken to encode and decode the image stream, must be added to the 'acquisition latency' discussed above to give a total latency due to the encoding and decoding process.

8.3.6 Image Quality

Ref. 4 states that at bit rates of 3 to 4Mb/s 'broadcast quality' is achievable with MPEG-1. (Note that MPEG-1 was not actually intended, and is not normally used, for broadcasting). However, sequences with complex spatial-temporal activity (such as sports) may require up to 5 to 6Mb/s. Motion estimator effectiveness determines motion artefacts, such as a reduction in video quality when movement starts or when the amount of movement exceeds a certain level. Poor motion estimation results in a general degradation in video quality. The higher the bit rate the fewer the image artefacts and/or the higher the resolution that can be achieved. Conversely reducing the bit rate does not result in a gradual reduction in quality of the decoded video, but rather a rapid degradation with 8x8 blocks becoming clearly visible below a certain threshold.

If MPEG-1 has problems handling rapid changes to the image inherent in sports material, then it will probably have even more problems with scenes captured from a fast low flying aircraft.

8.3.7 Data Rate

MPEG-1's constrained parameters mode defines data rate as up to 1.856Mb/s.

8.3.8 Error Resilience

Since MPEG uses interframe coding it is likely to be inherently less error resilient than standards that use intraframe coding.

8.3.9 Functionality

MPEG-1 provides a significant degree of functionality in the broadest sense. i.e. the provision of means by which the user may tailor the system to the specific needs of the application.

However, it is not able to provide all the functionalities listed as examples in 7.2.11 above. For example it does not provide lossless compression, Region of Interest (ROI) coding, arbitrary shaped objects or random access. Although editing capability, the main purpose of random access, may be enhanced by selecting suitable GOP configurations, possibly at the expense of other parameters. MPEG-1 does not qualify as a low complexity solution compared with JPEG.

8.3.10 Maturity

MPEG-1 was issued in 1992 and has found use in CD-ROMs etc. It was followed in 1994 by MPEG-2 aimed at applications including Cable TV (CATV). Hence both standards may be considered as mature, but it is likely that MPEG-2 will supplant MPEG-1 and MPEG-2 is the standard now used in the majority of digital TV systems. In the longer term both MPEG-1 and MPEG-2 are likely to give way to MPEG-4. However, it should be possible to regard MPEG-2 as a mature standard with a high degree of support with (by digital video standards) a significant ongoing life expectancy, not withstanding the comments in 7.2.12 above

8.3.11 Hardware/software availability

MPEG-1 is supported by hardware and software products from a range of suppliers, and therefore meets the requirements for hardware/software availability.

However, general concerns about environment and obsolescence apply as much to MPEG-1 as any other standard.

MPEG-1 was designed for 'one encode many decode' situations where low decoder cost is important. Hence MPEG is asymmetrical. That is to say the encoding process requires about two orders more computing power than the decoding. Hence encoders are likely to be significantly more costly than decoders.

8.4 MPEG-2

8.4.1 Overview

MPEG-2 is a second generation of the MPEG standard. MPEG-2 is formally defined in the international standard "ISO/IEC 13818 Information technology - Generic coding of moving pictures and associated audio".

It includes a large number of enhancements and extensions, the key ones being as follows:

- the ability to accommodate interlaced images
- a variety of raw image formats for a number of diverse applications ranging from low-quality SIF through to broadcast quality high definition television
- scalable modes (scaleable spatial resolution and scaleable image quality plus different quality levels in a single video stream)
- user definable DCT quantisation matrices and alternatives to the standard zig-zag scan
- non-linear DCT quantisation allowing a wider dynamic range than in MPEG-1
- enhanced motion prediction including half-pixel accuracy in motion vectors.
- Extended error redundancy due to special vectors in I frames
- Further prediction modes and macro blocks

To ensure compatibility between MPEG-2 encoders and decoders designed for different types of application, the standard defines a number of optional levels and profiles. These limit the format of the images, the bit-rate and the coding techniques associated with conformant bit-streams.

Overall, the standard offers many options and therefore considerable scope for optimisation in order to meet the requirements of a particular application. However, it is also very complex and in-depth understanding of what it has to offer requires a significant investment in time.

8.4.2 Image Format

Ref. 4 states that MPEG-2 supports resolutions up to 16,383x16,383 and 4:2:0, 4:2:2 and 4:4:4 YcbCr colour space specifications.

A display window within the encoded video may be defined. Conversely the encoded video may be defined as a window in a larger display.

8.4.3 Frame Rate

MPEG-2 provides for a wide range and variety of frame rates see 6.2.1.3.

MPEG-2 provides syntax for a 3:2 pull down process in the decoder to generate 30 FPS from 24 FPS input.

8.4.4 Compression Ratio

8.4.4.1 Lossless Compression

MPEG-2 does not provide lossless compression.

8.4.4.2 Lossy Compression

Refer to 5.4.2 above for examples of compression ratios achieved in a military application.

8.4.5 Latency

MPEG-2 uses interframe coding as discussed in 8.2.1 above and is thus subject to its limitations, in particular with regard to latency. However, it supports a Low Delay Mode in which no B frames are generated to eliminate frame reordering delay at the decoder, this is aimed at applications such as video conferencing.

8.4.6 Image Quality

Refer to 5.4.2 above for an indication of image quality achieved in a military application.

8.4.7 Data Rate

Ref. 4 states that the primary application targeted during definition of the standard was broadcast quality video at bit rates from 4 to 9Mb/s. However, MPEG-2 may be used for many other applications, including HDTV, at bit rates from 1.5 to 60Mbps.

MPEG-2 supports constant and variable bit rates, thus allowing use of higher bit rates to maintain video quality during difficult scenes.

Refer to 5.4.2 above for an indication of data rates used in a military application.

8.4.8 Error Resilience

MPEG-2 is claimed to be more resilient than MPEG-1. 5.4.2 above gives an indication of the error resilience of MPEG-2. A useful discussion of error detection and concealment for MPEG-2 will be found in Ref. 10.

8.4.9 Functionality

See 8.4.1 above.

8.4.10 Maturity

See 8.3.10 above.

8.4.11 Hardware/software availability

MPEG-2 is supported by hardware and software products from a range of suppliers, and therefore meets the requirements for hardware/software availability.

However, general concerns about environment and obsolescence apply as much to MPEG-2 as any other standard, although the large investment in digital TV implies that MPEG-2 systems will be existence for many years.

MPEG-2 was designed for 'one encode many decode' situations where low decoder cost is important. Hence MPEG-2 is asymmetrical. That is to say the encoding process requires about two orders more computing power than the decoding. Hence encoders are likely to be significantly more costly than decoders, although single chip encoders are now becoming available, thus considerably reducing cost.

8.5 MPEG-4

8.5.1 Overview

ASSC/130/3/191 states that MPEG-4 will support the coding of conventional images and video in a similar manner to MPEG 1 and 2.

Ref. 11 describes MPEG-4 as also providing a set of technologies to support:

- the coded representation of units of aural, visual or audiovisual content, called "audio-visual objects" or AVOs
- the way individual AVOs are composed in a scene

- the way AVOs are multiplexed and synchronised, so that they can be transported over network channels providing a Quality of Service (QoS) appropriate for the nature of the specific AVOs or the user requirements
- a generic interface between the application and the transport mechanisms
- the way the user interacts with the scene (e.g. changing the viewpoint) and the individual objects in a scene (e.g. clicking on an object to enquire about its features)
- the projection of the AV scene so composed on the desired viewing/hearing point.

Audiovisual scenes may be composed of several AVOs, organized in a hierarchical fashion. Primitive AVOs might include:

- a 2-dimensional fixed background
- the image of a talking person (without the background)
- the voice associated with that person
- a synthetic object (a desk, a chair etc.)
- a synthetic sound (background music)
- etc.

MPEG-4 standardises a number of types of such primitive AVOs, capable of representing both natural and synthetic content types, which can be either 2 or 3 dimensional. In addition to the AVOs mentioned above, MPEG-4 defines the coded representation of objects such as:

- text and graphics
- talking heads and associated text to be used at the receiver's end to synthesize the speech and animate the head
- animated human bodies.

In their coded form, these objects are represented as efficiently as possible. This means that the bits used for coding these objects are no more than necessary for supporting of desired functionalities. Examples of such functionalities are error robustness, allowing extraction and editing of an object, or having an object available in a scaleable form. It is important to note that in their coded form, objects (aural or visual) can be represented independently of their surroundings or background.

8.5.2 Applicability of MPEG-4 to Military Avionics

As can be seen from the above overview, although MPEG-4 provides for image and video coding in ways similar to MPEG-1 and 2, it also provides radically new approaches different to any of the other compression techniques considered here. The applicability of the AVO approach to the type of subject matter processed in military avionic applications is uncertain, although in most images there are relatively static or uninteresting areas which could be coded as separate objects at lower bit rate. It may be appropriate for artificially generated symbology etc., but it is difficult to see how well or otherwise it would lend itself to say views of terrain.

8.5.3 Maturity

MPEG-4 was released in late 2000 and is hence a new standard and can only be described as immature. The high power support that the development of the standard has received suggests that it will, in time be widely be adopted. However, it is (as of January 2001) too soon to be sure how successful the standard will be.

8.6 MPEG-7

MPEG-7 is known as “Multimedia Content Description Interface”. The objective of MPEG-7 is to assist implementation of systems that search for audio-visual content in much the same way as current widely used search engines seek textual information. Since MPEG-7 is not a standard for compression it is beyond the scope of this report. Potential applications of MPEG-7 in avionics have yet to be investigated by ASSC.

8.7 JPEG and its Variants

8.7.1 Overview

The JPEG (Joint Photographic Experts Group) introduced a still image graphics format in 1987, known as “ISO/IEC 10918 Information technology - Digital compression and coding of continuous-tone images” (more often called simply JPEG).

Since then a variety of standards including JPEG in their titles have appeared including :

- JPEG-L
- JPEG-LS
- MJPEG
- JPEG 2000
- MJPEG 2000.

These are briefly are discussed in 8.7.2 to 8.7.7 below, with consideration of the various issues already discussed.

8.7.2 JPEG

There are several modes defined for JPEG including baseline, lossless, progressive and hierarchical. The baseline mode is the most popular one and supports lossy coding only.

In the baseline mode, the image is divided in 8x8 blocks and each of these is transformed with the Differential Cosine Transform (DCT). The transformed blocks are quantized with a uniform scalar quantizer, zig-zag scanned and entropy coded with Huffman coding. The quantization step size for each of the 64 DCT coefficients is specified in a quantization table, which remains the same for all blocks. The DC coefficients of all blocks are coded separately, using a predictive scheme.

The progressive and hierarchical modes of JPEG are both lossy and differ only in the way the DCT coefficients are coded or computed when compared to the baseline mode. They allow a reconstruction of a lower quality or lower resolution version of the image by partial decoding of the compressed bitstream. Progressive mode encodes the quantized coefficients by a mixture of spectral selection and successive approximation, while hierarchical mode uses a pyramidal approach to computing the DCT coefficients.

8.7.2.1 Image Format

The JPEG standard only defines a set of compression algorithms and not to a specific image format. Hence the image format is constrained by the format of the file used to convey the compressed image.

JFIF (JPEG File Interchange Format), which was developed by C-Cube Microsystems, has been used as the de-facto standard within the Internet community for some time. However, many systems still use proprietary formats and as a result some compatibility problems exist - especially across different platforms. Ref. 12 provides a specification for JFIF as its Annex B.

The ISO JPEG committee has defined a format and known as the Still Picture Interchange Format (SPIFF) Ref. 13. It is claimed that it is upward compatible with JFIF and that there should be no compatibility problems with the predominant JFIF format.

8.7.2.2 Frame Rate

Since JPEG is for still images, frame rate is not defined.

8.7.2.3 Compression Ratio

Ref 12 tabulates compression ratio versus image quality for the same image (a standard image called 'Bike') for four different JPEG software implementations at various quality settings.

The lowest compression ratio is 3.6 for a quality setting between two levels described as 'no discernible difference from original' and one described as 'some slight artefacts at some points in the image'. The best result is a compression ratio of 6.9 for a quality setting described as 'no discernible difference from original'.

The highest compression ratio is 90 for a different implementation and for a quality setting between two levels described as 'image badly degraded' and one described as 'image visibly degraded across entire picture'. The worst compression at low quality is 35.6 for a quality between two levels described 'image visibly degraded across entire picture' and one described as 'some slight artefacts at some points in the image'.

These results show that there are significant variations in compression ratios and image qualities for similar quality settings for the different JPEG implementations, thus demonstrating the different levels of performance that may be achieved by different JPEG implementations.

8.7.2.4 Lossless Compression

See 8.7.3 re L-JPEG and 8.7.4 re JPEG-LS.

8.7.2.5 Lossy Compression

Achievable compression ratios vary considerably depending on subject matter and image quality requirements.

Ref. 12 includes a table listing compression ratios for a variety of subjects and image qualities. This indicates compression ratios from 2 to 45 for 'visually lossless' image qualities with the majority of images compressing from 8 to 15 times.

The highest compression indicated is 45:1 for a 'low quality' image, subject girl's face. The lowest is 1.2:1 for a 'maximum quality' image, subject 256 colour resolution chart.

8.7.2.6 Latency

Latency is very dependent upon image subject matter, compression ratio and implementation. Ref. 7 includes a table listing decoding times for a variety of images at 2 bpp lossy compression. This shows JPEG in progressive mode having decoding times ranging from 0.07 to 1.79 seconds. Decoding was performed in software on a PC computer with a 550 MHz Pentium™III processor, 512kB of cache and 512MB of RAM

under Linux 2.2.12. Note that the times are for decoding only, it is not clear whether the JPEG implementation tested was symmetrical, i.e. whether coding and decoding times are the same.

8.7.2.7 Image Quality

Ref. 7 includes a graph of PSNR v bpp. For JPEG in lossy progressive mode this indicates PSNRs ranging from 23dB for 0.25 bpp to 43dB for 2 bpp.

8.7.2.8 Data Rate

All other things being equal, data rate will be implementation dependent.

8.7.2.9 Error Resilience

Ref. 7 includes a table of PSNR versus BER for various bpp for JPEG and two modes of JPEG 2000. This shows JPEG having PSNR values ranging from 12dB at 2 bpp and BER of $1e^{-4}$ and 37dB at 2 bpp and BER of 0.

The following is adapted from Ref. 9 and is included to give an indication of the scope that may exist for enhancing error resilience:

JPEG is inherently one of the most robust video encoding standards. Since intraframe coding is used, bursty errors are not propagated to following images. The JPEG image quality is maintained with a RBER up to 10^{-5} , after this threshold the degradation rapidly increases (see graph in Figure 1) and the resulting artefacts are 'blocky' making the image unusable. However, with additional protection the threshold may be moved to around 10^{-2} .

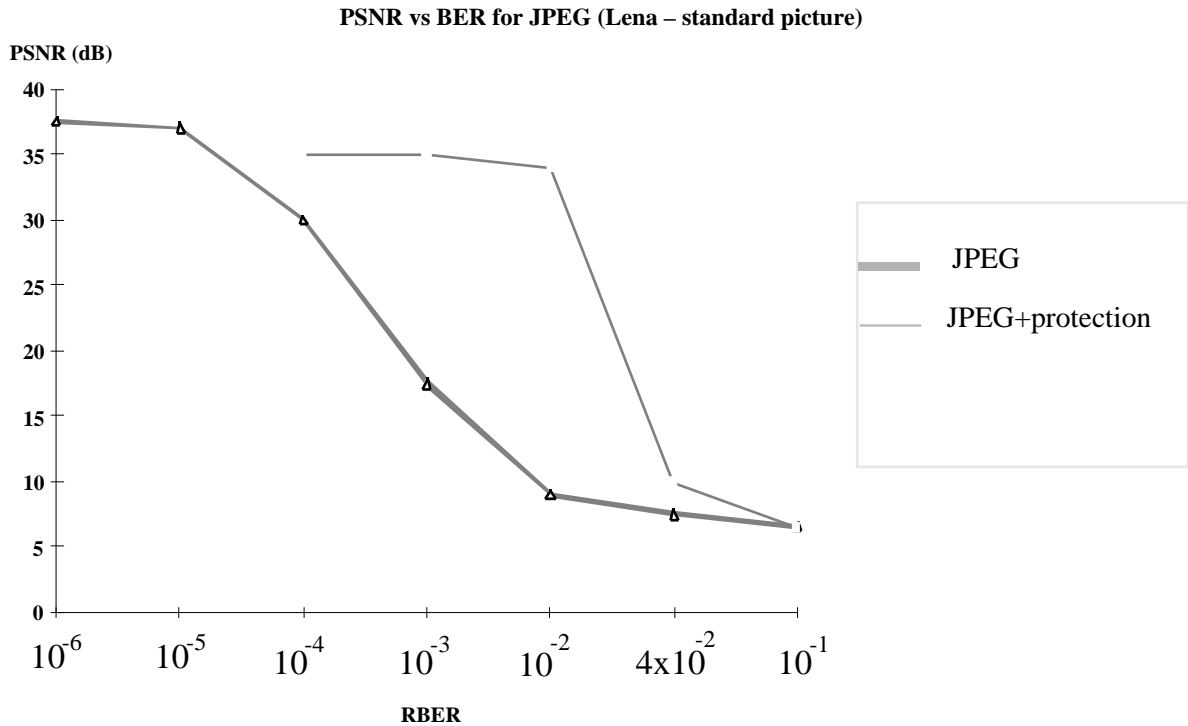


Figure 1: Comparison of JPEG CODEC image degradation against Bit Error Rate with and without RS error protection

When any error correcting code is decoded, there will be occasional decoding errors. In wireless communication these errors will inevitably be bursty in nature, regardless of the code used. Concatenation is a method of putting a second layer of code on top of the first to correct the majority of the decoding errors which occur after the first decoding. The outer code must be bursty error correcting and for practical purposes and efficiency a code known as Reed Solomon (RS) code is used as a good bursty error hunter. RS is also claimed to work at the limit of the Shannon's Law and therefore must be close to the most efficient error correction method.

Unlike some other coding techniques, RS is based on symbol error correction instead of bit error correction. RS uses a block code error correcting scheme and generates 'r' redundant symbols for each block, up to $r/2$ symbols can be corrected. The JPEG frame may be divided into blocks of data of 28 symbols with 8 bits per symbol (224 bits per block), in which $r=4$ symbols are redundant giving a protection scheme overhead of 14%.

8.7.2.10 Functionality

Refer to Ref. 16 which includes a functionality matrix for various standards including JPEG.

8.7.2.11 Maturity

Since JPEG (more properly ISO/IEC 10918) was introduced in 1987 and has been very widely adopted, it may be described as mature. The main question under this heading would be whether, to what extent and when it might be supplanted by a newer standard such as JPEG-2000?

8.7.2.12 Hardware/software availability

JPEG has been very widely implemented in both hardware and software for over a decade.

8.7.3 L-JPEG

8.7.3.1 Overview

Lossless JPEG, sometimes referred to as L-JPEG, is said to be not popular, it provides for lossless coding and does not support lossy coding.

L-JPEG is based on a completely different algorithm, which uses a predictive scheme. The prediction is based on the nearest three causal neighbours and seven different predictors are defined (the same one is used for all samples). The prediction error is entropy coded with Huffman coding.

8.7.3.2 Image Format

Image format is constrained by the file format. See 6.2.1.4 above.

8.7.3.3 Frame Rate

Since L-JPEG is for still images, frame rate is not defined.

8.7.3.4 Lossless Compression

Ref. 7 includes a table comparing lossless compression ratios for various JPEG variants. This indicates an average compression ratio for a variety of images for L-JPEG of 2.09, compared with 2.5 for JPEG 2000 and 2.98 for JPEG-LS.

8.7.3.5 Lossy Compression

L-JPEG does not provide lossy compression.

8.7.3.6 Latency

Ref. 7 includes a table listing decoding times for a variety of images with lossless compression using a variety of algorithms. This may be interpreted as showing L-JPEG with decoding times ranging from 0.05 to 1.6 seconds. See 8.7.2.6 for implementation

details. Note that the times are for decoding only, it is noted that L-JPEG was used in an asymmetrical mode, i.e. coding and decoding times different. Although no indication is given as to the degree of asymmetry.

8.7.3.7 Image Quality

With true lossless compression image quality following compression/decompression is, by definition, the same.

8.7.3.8 Data Rate

Data rate will be implementation dependent.

8.7.3.9 Error Resilience

No information specific to error resilience of L-JPEG has currently been identified.

8.7.3.10 Functionality

No information specific to the functionality of L-JPEG has currently been identified.

8.7.3.11 Maturity

Since L-JPEG is described as not popular there must be a suspicion that it has not been adopted to an extent that would permit it to be described as mature.

8.7.3.12 Hardware/software availability

Since L-JPEG is described as not popular there must be a suspicion that it has not been widely implemented.

8.7.4 JPEG-LS

8.7.4.1 Overview

JPEG-LS officially ISO/IEC 14495-1:1999: Information technology - Lossless and near-lossless compression of continuous-tone still images: Baseline, Dec. 1999.

JPEG-LS provides for lossless and “near-lossless” compression of still images. Ref. 7 states that Part-I, the baseline system, is based on adaptive prediction, context modeling and Golomb coding. In addition, it features a flat region detector to encode these in run-lengths. Near-lossless compression is achieved by allowing a fixed maximum sample error. Part-II will introduce extensions such as an arithmetic coder, but is still under preparation (as of July 2000). This algorithm was designed for low-complexity while providing high lossless compression ratios. However, it does not provide support for scalability, error resilience or any such functionality.

8.7.4.2 Image Format

If, like JPEG, JPEG-LS is restricted to defining coding/decoding algorithms then Image format is constrained by the file format. See 6.2.1.4 above.

8.7.4.3 Frame Rate

Since JPEG-LS is for still images, frame rate is not defined.

8.7.4.4 Lossless Compression

JPEG-LS is aimed specifically at lossless compression and according to Ref. 7 provides compression ratios ranging from 1.51 (for a standard aerial photographic image) to 3.77 (for cpmnd1 a standard compound document with text, photographs and computer graphics).

8.7.4.5 Lossy Compression

JPEG-LS only provides near lossless compression.

8.7.4.6 Latency

Ref. 7 includes a table listing decoding times for a variety of images with lossless compression using a variety of algorithms. This shows JPEG-LS with decoding times ranging from 0.07 to 2.25 seconds. See 8.7.2.6 for implementation details. Note that while the times are for decoding only, it is noted that JPEG-LS is symmetrical, i.e. coding and decoding times similar.

8.7.4.7 Image Quality

With true lossless compression image quality following compression/decompression is, by definition, the same. The extent to which image quality is degraded in near lossless mode is not clear. However, it would be expected that degradation would be virtually imperceptible in most cases.

8.7.4.8 Data Rate

Data rate will be implementation and application dependent.

8.7.4.9 Error Resilience

No information specific to error resilience of L-JPEG has currently been identified.

8.7.4.10 Functionality

Refer to Ref. 7 which includes a functionality matrix for various standards including JPEG-LS.

8.7.4.11 Maturity

JPEG-LS is a relatively recent standard (December 1999) described by Ref 7 (dated July 2000) as 'starting to appear in various applications'. Hence it probably does not meet our definition of mature.

8.7.4.12 Hardware/software availability

It is not clear to what extent hardware or software implementations of JPEG-LS have been produced.

8.7.5 MJPEG

8.7.5.1 Overview

Since a moving image can be (and often is) represented by a sequence of still images JPEG can be used to compress and decompress video sequences. In this case intra-frame coding is employed for each frame in the sequence independently, and the sequence is sent in abbreviated format (see 8.2.3 above).

This use of JPEG is known as Motion JPEG or MJPEG, it is not covered by an official JPEG or ISO/IEC standard, but is the subject of proprietary or de-facto standards. Ref. 12 p43 states that the most formal definition is in a Microsoft Multimedia technical note defining the JPEG DIB format (Ref. 14).

8.7.5.2 Image Format

MJPEG is restricted to defining coding/decoding algorithms and Image format is therefore constrained by the file format, see Ref 14.

8.7.5.3 Frame Rate

Since MJPEG consists of transmitting multiple JPEG images and JPEG is for still images, frame rate is not defined. Frame rate does not appear to be constrained by Ref 14.

Achievable frame rate will clearly be implementation dependent.

8.7.5.4 Lossless Compression

Ref. 14 is intended for implementation of the full scope of JPEG Baseline Sequential DCT process and hence does not appear to cater for lossless compression.

8.7.5.5 Lossy Compression

It would be expected that compression ratios would be similar to those achieved by JPEG. However, for a given processing capability compression ratio would be one

element in a trade off that would include (among other issues) image quality and frame rate.

Ref. 15 describes video editing cards such as Fast's AV Master or Miro's DC50, or the much less expensive Matrox Marvel product series, reducing the resulting data stream of a standard television signal from approximately 30 Mb/s to 6 Mb/s (MJPEG file). This corresponds to a compression ratio of 5:1.

However the above reference makes no mention of image quality.

8.7.5.6 Latency

Since JPEG uses only intraframe coding latency would have to be less than the reciprocal of frame rate.

8.7.5.7 Image Quality

Since MJPEG simply consists of JPEG applied to a series of images it is to be expected that MJPEG would produce a similar image quality to JPEG for similar quality settings, compression ratio, image content etc. However, image quality would be the subject of a trade off with compression ratio, frame rate etc. etc.

8.7.5.8 Data Rate

Data rate would be application/implementation dependent.

See also 8.7.5.5 above.

8.7.5.9 Error Resilience

It would be expected that MJPEG would exhibit similar error resilience to JPEG discussed in 8.7.2.9 above. Furthermore, since frames are displayed for only a fraction of a second it is unlikely that minor corruption of a single image would be noticeable. However, consideration would have to be given to the possibility of corruption of the file header etc.

8.7.5.10 Maturity

JPEG/ ISO/IEC 10918 was introduced in 1987 and the Microsoft Multimedia Technical Note: JPEG DIB Format (Ref 14) was published in May 1993, hence this particular de-facto standard for use of JPEG for video sequences has been around for about eight years. On this basis, and taking into account its widespread use, MJPEG can be described as mature. However, it is not possible to predict how soon MJPEG may be supplanted by another standard, most likely Motion JPEG 2000.

8.7.5.11 Hardware/software availability

Given the maturity and widespread use of MJPEG it would seem reasonable to assume that hardware and software are readily available.

8.7.6 JPEG 2000

8.7.6.1 Overview

JPEG 2000 uses wavelet compression as discussed in 8.2.2.2 above.

Ref. 16 reviews the procedures that divide the original image as follows:

- The image is decomposed into components.
- The image and its components are decomposed into rectangular tiles. The tile-component is the basic unit of the original or reconstructed image.
- Performing the wavelet transform on a tile-component creates decomposition levels. These decomposition levels can create components with different resolutions.
- These decomposition levels are made up of sub-bands of coefficients that describe the frequency characteristics of local areas (rather than across the entire tile-component) of the tile-component.
- The sub-bands of coefficients are quantized and collected into rectangular arrays of code-blocks.
- The bit-planes of the coefficients in a code-block are entropy coded in three coding passes.
- Some of the coefficients can be coded first to provide a region of interest.

At this point the data is fully decomposed and coded.

Ref 16 reviews the procedures that that reassemble these bit stream units into the codestream as follows:

- The coding passes from the code-blocks are collected in layers.
- Packets are composed of one partition of a single layer of a single decomposition level of a single tile-component.

The packets are the basic unit of the compressed data.

- All the packets from a tile are interleaved in one of several orders and placed in one, or more, tile-parts.

The tile-parts have a descriptive tile-part header and can be interleaved in any order.

- The codestream has a main header at the beginning that describes the original image and the various decomposition and coding styles that shall be used to locate, extract, decode, and reconstruct the image with the desired resolution, fidelity, region of interest, and other characteristics.
- The optional file format describes the meaning of the image and its components in the context of the application.

Annex I to Ref. 17 defines an optional file format “that applications may choose to use to contain JPEG 2000 compressed image data”. This file format is known as JP2. Clearly this is an improvement on the MJPEG situation where no file format was defined other than in proprietary and de-facto standards.

8.7.6.2 Image Format

The JP2 file Image Header Box specification includes 4 byte unsigned integer fields for image height and image width. This implies that image height and width may be regarded as virtually unrestricted.

8.7.6.3 Frame Rate

Since JPEG 2000 is intended for still images frame rate is not relevant.

8.7.6.4 Lossless Compression

Ref. 17 states that “J2K will offer scalable lossy-to-lossless decoding options”.

An example is shown with three images decoded at 0.15, 0.5 and 15.6 bpp. The latter is said to be an example of totally reversible encoding with 3:2 lossless compression

8.7.6.5 Lossy Compression

The 0.5 bpp example referred to in 8.7.6.4 above implies a compression ratio of 47:1.

See also See 8.7.6.7 below.

8.7.6.6 Latency

No information on latency has so far been identified. Latency will in any case be very implementation dependent.

8.7.6.7 Image Quality

Ref. 17 shows a comparison of JPEG and JPEG-2000 images. This shows original, JPEG processed and JPEG-2000 processed images, both of the processed images having been compressed at 0.27 bpp. While the JPEG image displays significant blocking, the JPEG-2000 image is virtually indistinguishable from the original.

This would clearly imply that if compression ratios were set to give similar image qualities then JPEG 2000 would succeed in equalling JPEG's quality at a much higher compression ratio.

8.7.6.8 Data Rate

JPEG 2000 is said to provide any desired bit rate (up to the rate used in compression).

8.7.6.9 Error Resilience

No information on Error Resilience of JPEG 2000 has been identified.

8.7.6.10 Functionality

Ref. 17 lists some of JPEG 2000's features as:

- Superior lossy performance at low bit rates (below ~0.25 bpp)
- Integrated lossy and lossless operation with scalable lossy-to-lossless decoding options
- Progressive transmission i.e. progressive with respect to spatial resolution and/or pixel precision
- Broader range of image types and content than JPEG-1
 - Image dimensions larger than 64K pixels
 - Pixel precision anywhere from 1 to 16 bits or more, vs. just 8 or 12
 - Multiple-component imagery (up to 256 components)
 - Mixed-content (e.g., continuous-tone and bi-level); computer-generated
- Random codestream access for region-of-interest (ROI) coding
- Better robustness to codestream bit errors
- Content-based and IPR metadata; image authentication features

8.7.6.11 Maturity

JPEG 2000 Part 1 Final Draft International Standard was released in August 2000. Hence JPEG 2000 certainly cannot be described as mature. However, the high powered support from over 60 different companies and government organisations for the JPEG committee (HP, Mitsubishi, Motorola, IBM, Sony, Kodak et al) suggests that the standard is likely to be widely adopted.

8.7.6.12 Hardware/software availability

No information on current hardware/software availability has been identified so far.

8.7.7 Motion JPEG 2000

8.7.7.1 Overview

Motion JPEG 2000 as consists of transmitting a sequence of pictures coded as per JPEG 2000. Motion JPEG 2000 will form Part 3 of the JPEG 2000 standard. It is described as low-complexity video coder based on Part 1. So, unlike MJPEG, MJPEG 2000 will be defined in an international standard. Ref.18.

Motion JPEG 2000 is said to use the same file format as JPEG 2000, see 8.7.6.1 above.

See 8.7.6 above for comments on image format, lossless compression, lossy compression, latency etc. etc.

8.8 **H261 and H263**

8.8.1 Overview

H.261 has primarily been defined for video telephony applications employing communications links having a capacity that is a multiple of 64kb/s per second (nominally an ISDN line). H.263 is a recent revision of the H.261 standard which includes a number of enhancements aimed at improving fidelity and levels of compression, and allowing the transfer of image sequences via communications links with a capacity less than 64kb/s. The standards are defined in the following documents:

- | | |
|-------|---|
| H.261 | ITU-T Recommendation H.261, Video codec for audio-visual services at $p \times 64$ kbit/s where p is the number of channels |
| H.263 | ITU-T Recommendation H.263, Video coding for narrow telecommunication channels at < 64 kbit/s. |

Raw images are constrained to have resolutions which are the various standard multiples and divisors of CIF.

Like MPEG, H261 and H263 use a combination of intraframe and interframe coding.

For video telephony H261 and H263 are used in combination with various other “H.324 standards” such as G.723 for audio and H.245 for control protocol.

8.8.2 H261

8.8.2.1 Image Format

H261 handles video CIF (352x288) and QCIF (176x144) image formats with YcbCr 4:2:0 colour space coding.

H261 also allows transmission of still images at four times the currently selected video format. i.e. if currently selected the video format is QCIF then a CIF still image may be transmitted, if the video format is CIF then a still image of 704x576 may be transmitted.

8.8.2.2 Frame Rate

Frame Rate is restricted to 29.97 FPS.

8.8.2.3 Lossless Compression

H261 does not appear to cater for lossless compression.

8.8.2.4 Lossy Compression

Compression ratio will be very dependent upon implementation/application and image subject matter, no definitive indications of typical values have been identified.

8.8.2.5 Latency

Latency will be dependent upon implementation/application and image subject matter, no definitive indications of typical values have been identified.

8.8.2.6 Image Quality

Ref. 19 includes example video clips comparing H261 performance with MPEG, plus graphs of PSNR v bit rate for H261 with various standard images. There is also a graph of PSNR plotted against the compression ratio.

8.8.2.7 Data Rate

H261 handles video bit rates between 64kb/s and 2Mb/s.

8.8.2.8 Error Resilience

511/493 BCH forward error correction code is used.

Prediction for interframe coding uses only the closest previous picture and is thus less prone to propagation of errors.

8.8.2.9 Functionality

No definitive information on functionality has been identified.

8.8.2.10 Maturity

H261 is a mature standard, but will probably be supplanted by H263.

8.8.2.11 Hardware/software availability

Hardware and software is probably widely available, but will probably be supplanted by H263.

8.8.3 H263

8.8.3.1 Image Format

H263 handles the following image formats:

- 16CIF 1408x1152
- 4CIF 704x576
- CIF 352x288
- QCIF 176x144
- SQCIF 128x96

Colour space coding is 4:2:0.

8.8.3.2 Frame Rate

H263 handles only one frame rate: 29.97 FPS.

8.8.3.3 Lossless Compression

H263 does not appear to provide lossless compression.

8.8.3.4 Lossy Compression

H263 is said to be about twice as efficient as H261.

8.8.3.5 Latency

Latency will be dependent upon implementation/application and image subject matter etc., no definitive indications of typical values have been identified.

8.8.3.6 Image Quality

Ref. 19 provides example video clips comparing H263 with MPEG, plus graphs of PSNR v bit rate for H263 for various standard images. A graph of PSNR versus compression ratio for H263 is provided together with graphs comparing the performance of H261 and H263.

8.8.3.7 Data Rate

H.263 was designed for very low bit rate coding applications. However it is expected that the standard will be used for a wide range of bitrates, not just low bitrate applications. It has been suggested that H.263 will replace H.261 in many applications.

8.8.3.8 Error Resilience

No definitive indications of Error Resilience have been identified.

8.8.3.9 Functionality

No definitive indications of functionality have been identified.

8.8.3.10 Maturity

Since H263 was published in 1996 and has numerous implementations, it may thus be considered mature.

8.8.3.11 Hardware/software availability

H263 hardware is available from numerous suppliers including Apple, Microsoft and Intel.

9 CONCLUSIONS

This investigation set out to compare the capabilities of a number of video compression standards with the requirements of military avionic systems.

It has been found that insufficient definition of those requirements currently exists to permit a comprehensive comparison. However, comparisons have been undertaken insofar as is possible given the limited requirement definition. It is important to appreciate that comparisons are based on constraints imposed by the standards, and this does not necessarily mean that implementations are available to meet all the requirements.

The comparisons have been supported by discussions of the capabilities of various standards, compared with the likely requirements for avionic applications, and with indications of the additional work needed to improve the definition of requirements and the manner in which comparisons might be presented. These discussions should serve to highlight the issues to be considered in selecting standards for video compression in military avionic applications. A list of some of the issues that need to be considered in the development of a more comprehensive definition of requirements for digital video compression in military avionics has been provided in Table 5.

Some important, if generalised, conclusions that may be drawn from the investigation into the characteristics of various compression techniques and standards relates to the relative merits of interframe and intraframe coding.

The combination of intraframe and interframe coding, as used in MPEG-1, 2 & 4 etc., is likely to provide higher compression ratios in many situations than intraframe coding only (although this may not apply to certain military avionic applications). However, interframe coding may suffer from disadvantages in terms of latency, loss of picture quality when image content changes rapidly and possible greater inherent vulnerability to the effects of errors.

Intraframe only coding (as used in various JPEG standards) may be less subject to latency, be more inherently error resilient and not prone to loss of picture quality when image content changes rapidly, but is likely to give less compression than interframe coding.

It is clear that development of digital video techniques continues apace in the commercial digital television and multimedia industries and that there are significant potential benefits for military avionics in adoption of such technology. There are, however, significant problems and risks associated with exploitation of COTS digital video, especially in the areas of environmental requirements and obsolescence.

REFERENCES

1 ASSC/130/2/163 Issue 1 March 2002, Study of the Application of the MPEG-2 Digital Video Compression Standard for use in Uninhabited Aerial Vehicle Systems.

2 Riley, M.J. and Richardson I.E.G, Digital Video Communications, Artech House, 1997 ISBN 0-89006-890-9

3. Phil's Home Page, University of Bath www.bath.ac.uk/~en8pjh/page6.htm

4. Keith J

Video Demystified: a handbook for the digital engineer

HighText , 2nd ed.,1996, ISBN 1-878703-23-1

5. Editor: Koenen R

MPEG-4 Overview - (V.16 - La BauleVersion)

International Organisation for Standardisation ISO/IEC JTC1/SC29/WG11

Coding of Moving Pictures and Audio ISO/IEC JTC1/SC29/WG11 N3747

La Baule - October 2000

6. Homepage of Michael Clemens and Michael J. Thul, University of Kaiserslautern www.student.uni-kl.de/~miclemen/web/psnr.html

7 Santa Cruz D, Ebrahimi T et al.

An analytical study of JPEG 2000 functionalities N1816

To be published in the Proceedings of SPIE, vol 4115, of the 45th annual SPIE meeting, Applications of Digital Image Processing XXIII., July 2000.

8. Simon K

An Introduction to JPEG

IAM Research Group Department of Electronics & Computing Science, University of Southampton.

www.ecs.soton.ac.uk/~srk98r/jpeg/JPEG.html

-
9. Amin R, Gill D, Hobday J, Stone F, Violante S, Whitehead C
Understanding Distributed Multimedia Technologies and Applications
ERA Report 97-0735R, Leatherhead, ERA Technology Ltd., 1997, 134pp.
10. Chen, C.-T. Error detection and concealment with an unsupervised MPEG2 video decoder, Journal of Visual Communications and Image Representation, 1995 Vol3, No. 2 pages 265-279
11. MPEG and multimedia communications Chiariglione L, CSELT - Torino - Italy
www.cselt.it/ufv/leonardo/paper/isce96.htm
12. Clark R, IEE IST/37 committee
Technical Guide to JPEG Digital Compression of Photographic Images
DISC PD 0006
ISBN 0 580 249107
13. ITU ISO/IEC
Digital Compression and Coding of Continuous-tone Still Images – Extensions
Still Picture Interchange File Format (SPIFF)
Annex F of of ITU-T Recommendation T.84 | ISO/IEC IS 10918-3
14. Microsoft
Multimedia Technical Note: JPEG DIB Format
Microsoft Development Library, 1994
15. Tom's Hardware Guide www4.tomshardware.com/video/99q3/990924/video-3-03.html
16. Boliek M, Christopoulos C, Majani E
Coding of Still Pictures
ISO/IEC JTC1/SC29 WG1, N1646R, JPEG 2000 Final Committee Draft Version 1.0, 16 March 2000
17. Brislawn C
A Preview of the Upcoming JPEG 2000 Image Coding Standard
Los Alamos National Laboratory April 2000
www.c3.lanl.gov/~brislawn/JPEG_demo/sld005.htm

18 Grant J

Motion JPEG 2000

Nine Tiles Networks Ltd, Cambridge England, ASSC/130/2/248, 41 slides.

19. H.263 Video Coding www-mobile.ecs.soton.ac.uk/peter/h263/h263.html