NATO STANDARDIZATION AGENCY
AGENCE OTAN DE NORMALISATION

13 October 2009 NSA/1117(2009)-JAIS/4609

STANAG 4609 JAIS (EDITION 3) – NATO DIGITAL MOTION IMAGERY STANDARD

References:
a. NSA/0554(2007)-AIR/4609 dated 15 June 2007 (Edition 2)
b. AC/224-D(2009)0011 dated 25 June 2009(Edition 3) Ratification Draft
c. AC/224-D(2009)0011(AS) dated 25 September 2009

1. The enclosed NATO Standardization Agreement, which has been ratified by nations as reflected in the NATO Standardization Documents Database (NSDD), is promulgated herewith.

2. The reference listed above is to be destroyed in accordance with local document destruction procedures.

ACTION BY NATIONAL STAFFS

3. National staffs are requested to examine their ratification status of the STANAG and, if they have not already done so, advise the Defence Investment Division through their national delegation as appropriate of their intention regarding its ratification and implementation.

Juan A. MORENO
Vice-Admiral, ESP(N)
Director, NATO Standardization Agency

Enclosure:
STANAG 4609 (Edition 3)

NATO Standardization Agency – Agence OTAN de normalisation
B-1110 Brussels, Belgium Internet site: http://nsa.nato.int
E-mail: lazar.liviu@hq.nato.int – Tel 32.2.707. 7914 – Fax 32.2.707.4103

STANAG 4609
(Edition 3)

NORTH ATLANTIC TREATY ORGANIZATION
(NATO)

NATO STANDARDISATION AGENCY
(NSA)

STANDARDIZATION AGREEMENT
(STANAG)

SUBJECT: NATO DIGITAL MOTION IMAGERY STANDARD

Promulgated on 13 October 2009

Juan A. Moreno
Vice Admiral, ESP (N)
Director, NATO Standardization Agency

RECORD OF AMENDMENTS

No.	Reference/date of Amendment	Date Entered	Signature

EXPLANATORY NOTES

AGREEMENT

1. This NATO Standardization Agreement (STANAG) is promulgated by the Director NATO Standardization Agency under the authority vested in him by the NATO Standardization Organisation Charter.

2. No departure may be made from the agreement without consultation with the Custodian. Nations may propose changes at any time to the Custodian where they will be processed in the same manner as the original agreement.

3. Ratifying nations have agreed that national orders, manuals and instructions implementing this STANAG will include a reference to the STANAG number for purposes of identification.

RATIFICATION, IMPLEMENTATION, AND RESERVATIONS

4. Ratification, implementation and reservation details are available on request or through the NSA websites (internet http://nsa.nato.int; NATO Secure WAN http://nsa.hq.nato.int).

FEEDBACK

5. Any comments concerning this publication should be directed to NATO/NSA – Bvd Leopold III – 1110 Brussels – BE.

NATO STANDARDIZATION AGREEMENT (STANAG)

Motion Imagery

Annexes:
A. Terms and Definitions
B. Motion Imagery System (MIS)
C. Standards

The following Standardization Agreements (STANAGs), Military Standards (MIL-STDs), International Telecommunication Union (ITU) Recommendations and International Standards Organization (ISO) standards contain provisions, which, through references in this text, constitute provisions of this STANAG. At the time of publication, the editions indicated were valid.

REFERENCE DOCUMENTS:

Normative References

[1] ISO/IEC 7498-1:1994 Open Systems Interconnection - Basic Reference Model: The Basic Model

[2] AEDP-8 Motion Imagery Allied Engineering Document

[3] ISO/IEC 13818-1:2007, Information technology - Generic coding of moving pictures and associated audio information: Systems (also known as MPEG-2 Systems)

[4] ISO/IEC 13818-2:2000, Information technology - Generic coding of moving pictures and associated audio information: Video (also known as MPEG-2 Video)

[5] ISO/IEC 13818-3:1998, Information technology - Generic coding of moving pictures and associated audio information: Audio (also known as MPEG-2 Audio)

[6] ISO/IEC 13818-4:2004, Information technology - Generic coding of moving pictures and associated audio information: Compliance Testing (also known as MPEG-2 Compliance)

[7] ITU-R BT.601-5: Studio encoding parameters for digital television for standard 4:3 and wide-screen 16:9 aspect ratios, 1995

[8] ITU-R BT.1358: Studio parameters of 625 and 525 line progressive scan television systems

[9] SMPTE 292M-2006, Television - 5 Gb/s Signal/Data Serial Interface

[10] SMPTE 274M-2005, Television - 1920 x 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequences for Multiple Picture Rates (progressive origination only)

[11] ITU-T Rec. H.264, Advanced Video Coding for Generic Audio Visual Services, 2007

ISO/IEC 14496-10:2008, Coding of audio-visual objects, Part 10: Advanced Video Coding (also known as H.264)

[12] ITU-T Rec. H.264, Advanced video coding for generic audiovisual Services Amendment 1 (2006): Support of additional colour spaces and removal of the High 4:4:4 Profile

[13] SMPTE 259M-2006, Television - 10-Bit 4:2:2 Composite and 4 fsc Composite Digital Signals – Serial Digital Interface

[14] SMPTE 349M-2001, Transport of Alternate Source Image Formats though SMPTE 292M

[15] SMPTE 291M-2006, Television - Ancillary Data Packet and Space Formatting

[16] STANAG 4545 NATO Secondary Imagery Format

[17] ITU-T Rec. T.800 ISO/IEC 15444-1:2004, Information Technology - JPEG 2000 Image Coding System: Core Coding System

[18] ISO/IEC 10918-1:1994, Digital compression and coding of continuous-tone still images: Requirements and guidelines

[19] ISO/IEC BIIF Profile BPJ2K01.00 Amd 1, BIIF Profile for JPEG 2000 Version 01.00 Amd 1, 2007

[20] MISB RP 0705.2 Version 1.1, LVSD Compression Profile, 2008

[21] STANAG 7023, Air Reconnaissance Imagery Data Architecture

[22] SMPTE 335M-2001, Metadata Dictionary Structure

[23] SMPTE 336M-2007, Data Encoding Protocol Using Key-Length-Value

[24] MISB STANDARD 0807, KLV Metadata Dictionary

[25] SMPTE RP210.10-2007, SMPTE Metadata Dictionary Contents

[26] SMPTE 12M-1999, Television, Audio and Film - Time and Control Code

[27] SMPTE 309M-1999, Transmission of Date and Time Zone Information in Binary Groups of Time and Control Code

[28] SMPTE RP 214-2002, Packing KLV Encoded Metadata and Data Essence into SMPTE 291M Ancillary Data Packets

[29] SMPTE RP 217-2001, Nonsynchronized Mapping of KLV Packets into MPEG-2 System Streams

[30] MISB EG 0801, Profile 1: Photogrammetry Metadata Set for Digital Motion Imagery, 2008

[31] MISB STANDARD 0601.2,"UAV Datalink Local Data Set," 2008

[32] IEEE 1003.1, Information Technology-Portable Operating System Interface (POSIX), 2004

[33] STANAG 4586, UAV Control System (UCS) Architecture

[34] ISO 3166-1, Codes for the representation of names of countries and their subdivisions: Country Codes, 1 October 1997 and updated by the ISO 3166 Management Authority (MA) at: http://www.din.de/gremien/nas/nabd/iso3166ma/codlstp1/index.html

[35] STANAG 1059 Ed 1 rev 1, Letters Codes for Geographical Entities

[36] MISB RP 0101, Use of MPEG-2 Systems Streams in Digital Motion Imagery Systems

[37] MISB RP 0102.4, Security Metadata Universal and Local Sets for Digital Motion Imagery, 2007

[38] MISB STANDARD 0604, Time Stamping Compressed Motion Imagery, 2008

[39] SMPTE RP 188 - 1999, Transmission of Time Code and Control Code in the Ancillary Data Space if a Digital Television Data Stream

[40] MISB RP 0603, Common Time Reference for Digital Motion Imagery Using Coordinated Universal Time (UTC), 2006

[41] SMPTE 328M-2000, MPEG-2 Video Elementary Stream Editing Information

[42] SMPTE 377M-2004, Material Exchange Format (MXF) File Format Specification (Standard)

[43] SMPTE 378M-2004, Material Exchange Format (MXF) Operational pattern 1A (Single Item, Single Package)

[44] SMPTE 391M, Material Exchange Format (MXF) Operational Pattern 1b (Single Item, Ganged Packages) Advanced Authoring Format Object Specification, V 1.0, AAF Association, 9 June 2000

[45] SMPTE 379M-2004 Material Exchange Format (MXF) MXF Generic Container

[46] SMPTE 381M-2005, Material Exchange Format (MXF) Mapping MPEG streams into the MXF Generic Container (Dynamic)

[47] SMPTE 380M-2004, Material Exchange Format (MXF) Descriptive Metadata Scheme (Standard, Dynamic)

[48] EIA-608, Recommended Practice for Line 21 Data Service, September 1994

[49] SMPTE 296M-2001, Television - 1280x720 Progressive Image Sample Structure - Analog and Digital Representation and Analog Interface

[50] MISB EG 0104.5, Predator UAV Basic Universal Metadata Set, 15 June 2004

[51] SMPTE 330M-2003, Unique Material Identifier (UMID)

[52] SMPTE 295M-1997, Television - 1920x1080 50-Hz Scanning and Interface

[53] SMPTE 294M-2001, Television - 720x483 Active Line at 59.94-Hz Progressive Scan Production - Bit-Serial Digital Interfaces

[54] SMPTE EG37-2001, Node Structure for the SMPTE Metadata Dictionary

[55] Director of Central Intelligence, Community Management Staff, Controlled Access Program Coordination Office (CAPCO), Intelligence Community Classification and Control Markings Implementation Manual, 10 Sep 1999, amended 12 Oct 2000

[56] CAPCO Authorized Classification and Control Markings Register, 12 Oct 2000

[57] Federal Information Processing Standards (FIPS) Publication 10-4, Countries, Dependencies

Informative References

Areas of Special Sovereignty, and Their Principal Administrative Divisions, National Institute of Standards and Technology, April 1995 (through Change Notice 6, 28 January 2001)

ASI-00209 Rev D, Exploitation Support Data (ESD) External Interface Control Document, 04 December, 2002

Director of Central Intelligence Directive 1/7, 30 Jun 1998

Director of Central Intelligence Directive (DCID) 6/3, Security Requirements for Interconnected Information Systems, 4 Feb 2000

DOD Directive (ASD (NII)) “Data Sharing in a Net-Centric Department of Defense,” Number 8320.2, 2 December 2004.

DOD Directive 5100.55: U. S. Security Authority for NATO Affairs reissued. 27 February 2006.

DOD Directive 5200.1 (ASD (C3I)), 13 December 2001, certified current 24 November 2003.

DOD Instruction, Number 5210.52, Security Classification of Airborne Sensor Imagery and Imaging Systems, 18 May 1989

DOD Net-centric Data Strategy, 9 May 2003, Classified National Security Information

DOD 5220.22-M (USD(I)) National Industrial Security Program Operating Manual (NISPOM), 28 February 2006

ETS 300 421, “Digital broadcasting systems for television, sound and data services; framing structure, channel coding and modulation for 11/12 GHz satellite services” (DVB-S)

ETS 300 744, “Digital Video Broadcasting; framing structure, channel coding and modulation for digital Terrestrial television” (DVB-T)

Executive Order 12958, Jun 1995

Executive Order 13292, 25 March 2003, Further Amendment to EO 12958, as amended,

IEEE STD 1394-1995, Standard for a High Performance Serial Bus

Imagery Policy Series, Particular Section 6,”National Airborne Reconnaissance Imagery”

ISO 1000:1992(E), SI units and recommendations for the use of their multiples and of certain other units, 11 January, 1992

ISO/IEC 12087-5:1998, Information technology - Computer graphics and image processing - Image Processing and Interchange (IPI) - Functional specification - Part 5: Basic Image Interchange Format (BIIF)

ISO/IEC 13818-6:1998, Information technology - Generic coding of moving pictures and associated audio information: Extension for Digital Storage Media Command and Control (also known as MPEG-2 DSM-CC)

ISO/IEC 13818-9:1996, Information technology - Generic coding of moving pictures and associated audio information: Real-time Interface Specification (also known as MPEG-2 RTI)

ISO/IEC 15444-9:2005, Information technology - JPEG 2000 image coding system: Interactivity tools, APIs and protocols

ITU-T Rec. H.222, Amendment 3, 2004: Transport of AVC data over ISO/IEC 13818-1/ H.222.0 for MPEG2 TS containment for MPEG4 AVC

ITU-T Rec. T.801 ISO/IEC 15444-2:2004, Information Technology - JPEG 2000 Image Coding System: Extensions.

ITU-T Rec. T.808 ISO/IEC 15444-9:2004, Information Technology - JPEG 2000 Image Coding System: - Part 3: Interactive protocols and APIs

MIL-STD-2500B, National Imagery Transmission Format Version 2.1 for the National Imagery

MIL-STD-2500C V2.1, National Imagery Transmission format Standard, 2006

MISB RP 0103.1, Timing Reconciliation Metadata Set for Digital Motion Imagery, October 2001

NATO ISRI (Intelligence, Surveillance, and Reconnaissance) Integration Working Group Terms and Definitions, Draft 2001

NIIA Document 26 March 2002

STANAG 3678 Guide to Security Classification of Air Reconnaissance Imagery, 2005

STANAG 4559 Image Product Library Interface Standard (NSILI)

STANAG 4575 Imagery Air Reconnaissance (Digital Storage)

STANAG 7024 Imagery Air Reconnaissance Tape Recorder Standards

STANAG 7085 Interoperable Data Links for Imaging Systems

SMPTE 170M-2004, Television - Composite Analog Video Signal - NTSC for Studio Applications

SMPTE EG 41-2004 Material Exchange Format (MXF) Engineering Guideline (Informative)

SMPTE EG 42-2004 Material Exchange Format (MXF) MXF Descriptive Metadata

Transmission Format Standard, 22 August 1997

AIM

1. The aim of this agreement is to promote interoperability of present and future motion imagery systems in a NATO Combined/Joint Service Environment. Interoperability is required because it will significantly enhance the warfighting capability of the forces and increase flexibility and efficiency to meet mission objectives through sharing of assets and common utilization of information generated from motion imagery systems.

AGREEMENT

2. Participating nations agree to implement the standards presented herein in whole or in part within their respective Motion Imagery systems to achieve interoperability.

DEFINITIONS

3. The terms and definitions used in this document are listed in Annex A.

GENERAL SECTION

4. The outline of this STANAG follows the following format:

- Annex A contains the Terms and Definitions used in the STANAG.
- Annex B contains the description of the Motion Imagery System (MIS)
- Annex C contains the Standards mandated by this STANAG

DETAILS OF AGREEMENT

5. The Motion Imagery Architecture STANAG defines the architectures, interfaces, communication protocols, data elements, message formats and identifies related STANAGs, which compliance with is required.

IMPLEMENTATION OF THE AGREEMENT

6. This STANAG is implemented by a nation when it has issued instructions that all such equipment procured for its forces will be manufactured in accordance with the characteristics detailed in this agreement.

TERMS AND DEFINITIONS

1 Acronyms and Abbreviations. The following acronyms are used for the purpose of this agreement. Note: There will only be words associated with this STANAG that are not already included in the ISRIWG Dictionary.

A	
AEDP	Allied Engineering Documentation Publication
AES3	Audio Engineering Society 3
ANSI	American National Standards Institute
AAF	Advanced Authoring Format
ATM	Asynchronous Transfer Mode
ATV	Advanced Television
B	
C	
C2	Command and Control
C3I	Command Control Communication, and Intelligence
C4I	Command, Control, Communications, Computers and Intelligence
CCI	Command and Control Interface
CDL	Common Data Link
CGS	Common Ground Segment, Common Ground Station
CIF	Common Image Format (352x288)
COTS	Commercial Off-The-Shelf
D	
DVB-T	Digital Video Broadcast - Terrestrial
DVB-S	Digital Video Broadcast - Satellite
DCGS	Distributed Common Ground Station
DoD	Department of Defense
DLI	Data Link Interface
DTED	Digital Terrain Elevation Data
DV	Digital Video
DVD	Digital Versatile Disk; Digital Video Disk
D-VHS	Digital VHS
D-VITC	Digital VITC
E	
EBU	European Broadcast Union
ED	Enhanced Definition
EG	Engineering Guideline
EIA	Electronic Industries Association
ETR	European Telecommunications Report
F	
FCC	Federal Communications Commission

FLIR	Forward Looking Infrared
FOV	Field Of View
FPS	Frames Per Second
FTP	File Transfer Protocol
G	
GB	Gigabyte
Gb	Gigabits
GBS	Global Broadcast Service
GOP	Group Of Pictures
GOTS	Government Off-The-Shelf
GPS	Global Positioning System
H	
HD	High Definition
HDTV	High Definition Television
HL	High level
Hz	Hertz
I	
IC	Intelligence Community
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronic Engineers
IMINT	Imagery Intelligence
IP	Internet Protocol/Intellectual Property
IPL	Image Product Library
IR	Infrared
ISDN	Integrated Services Digital Network
ISO	International Standards Organization
ISR	Intelligence, surveillance, reconnaissance
ITU	International Telecommunication Union
J	
JFC	Joint Forces Commanders
JPEG	Joint Photographic Experts Group
JPIP	JPEG 2000 Interactive Protocol
JTA	Joint Technical Architecture
JTF	Joint Task Force
JWICS	Joint Worldwide Intelligence Communications System
K	
Kb/s	Kilobits per second
KB/s	Kilobytes per second
Kilo	1,000
KLV	Key-Length-Value
L	
LVSD	Large Volume Streaming Data
M	
Mb/s	Megabits per second
MB/s	Megabytes per second

MIL	Military
MIL-STD	Military Standard
MISM	Motion Imagery Systems Matrix
MISM-L	Motion Imagery Systems Matrix - Level
MJD	Modified Julian Date
ML	Main Level
MP	Mission Planning; Main Profile
MPEG	Moving Pictures Experts Group
N	
N/A	Not Applicable
NATO	North Atlantic Treaty Organization
NCIS	NATO Common Interoperability Standards
NITFS	National Imagery Transmission Format Standard
NRT	Non Real-Time, Near Real Time
NSIF	NATO Secondary Imagery Format
NSIL	NATO Standard Image Library
NSILI	NATO Standard Image Library Interface
NTIS	NATO Technical Interoperability Standards
NTSC	National Television Standards Committee
O	
OC-3	Fiber Optic Communications Standard (155 Mbps)
OC-12	Fiber Optic Communications Standard (655 Mbps)
P	
PAL	Phase Alternate Line
p	Progressive
ps	progressive scan
PS	Program Stream
Q	
QoS	Quality of Service
QSIF	Quarter SIF (176 x 120 Pixels)
R	
RF	Radio Frequency
RP	Recommended Practice
RSTA	Reconnaissance Surveillance and Target Acquisition
Rx	Receive
S	
s	seconds
SATCOM	Satellite Communications
SD	Standard Definition
SDI	Serial Digital Interface
SDTI	Serial Data Transport Interface
SECAM	System Electronique Couleur Avec Mémoire
SIF	Standard Image Format (352x240 pixels)
SMPTE	Society of Motion Picture and Television Engineers
SNR	Signal to Noise Ratio

STANAG	(NATO) Standardization Agreement
S-VHS	Super Vertical Helical Scan
T	
TBD	To Be Defined
TS	MPEG-2 Transport Stream
TST	Technical Support Team
TUAV	Tactical UAV
TV	Television
Tx	Transmit
U	
UAV	Unmanned/Uninhabited Aerial Vehicle
UCAV	Unmanned/ Uninhabited Combat Aerial Vehicle
US	United States
UTC	Universal Time Code Coordinated
V	
VANC	Vertical Ancillary Interval
VCR	Video Cassette Recorder
VHS	Vertical Helical Scan
VITC	Vertical Interval Time Code
W	
X	
XML	eXtensible Markup Language
Y	
Z	

2. Terms and Definitions. The following terms and definitions are used for the purpose of this agreement.

Analysis	In intelligence usage, a step in the processing phase of the intelligence cycle in which information is subjected to review in order to identify significant facts for subsequent interpretation
Byte	Eight binary bits
Engineering Guidelines	Engineering Guidelines represent well-defined, informative engineering principals. Engineering Guidelines are not mandated.
Image	A two-dimensional rectangular array of pixels indexed by row and column

Imagery	A likeness or representation of any natural or man-made feature or related object or activity. Collectively, the representations of objects reproduced electronically or by optical means on film, electronic display devices, or other media.
Interface	(1) A concept involving the definition of the interconnection between two equipment items or systems. The definition includes the type, quantity, and function of the interconnecting circuits and the type, form, and content of signals to be interchanged via those circuits. Mechanical details of plugs, sockets, and pin numbers, etc., may be included within the context of the definition. (2) A shared boundary, e.g., the boundary between two subsystems or two devices. (3) A boundary or point common to two or more similar or dissimilar command and control systems, subsystems, or other entities against which or at which necessary information flow takes place. (4) A boundary or point common to two or more systems or other entities across which useful information flow takes place. (It is implied that useful information flow requires the definition of the interconnection of the systems, which enables them to interoperate.) (5) The process of interrelating two or more dissimilar circuits or systems. (6) The point of interconnection between user terminal equipment and commercial communication-service facilities.
Intelligence	The product resulting from the collection, processing, integration, analysis, evaluation and interpretation of available information concerning foreign countries or areas
Interlace Scan	Interlace scanning scans from left to right for one line then skips every other line to form a field of the image. The second field is made up of the lines that were skipped in the first field. The combination of two fields constitutes a frame. It should be noted that motion between fields in the frame causes interlace artifacts in the frame and the loss of vertical and temporal resolution.
Interoperability	Interoperability is the ability of systems, units or forces to provide services to and accept services from other systems, units of forces and to use the services so exchanged to enable them to operate effectively together
Motion Imagery	A likeness or representation of any natural or man-made feature or related object or activity utilizing sequential or continuous streams of images that enable observation of the dynamic behavior of objects within the scene. Motion Imagery temporal rates, nominally expressed in frames per second must be sufficient to characterize the desired dynamic phenomenon. Motion Imagery is defined as including metadata and nominally beginning at frame rates of 1 Hz (1 frame per second) or higher within a common field of regard. Full Motion Video (FMV) falls within the context of these standards.
Near-Real-Time	Delay caused by automated processing and display between the occurrence of an event and reception of the data at some other location

Non-Real Time Processing	Non-flight critical processing accomplished within the host system software including interface to C4I system(s). Pertaining to the timeliness of data or information that has been delayed by the time required for electronic communication and automatic data processing. This implies that there are no significant delays.
Open Systems Interconnect Model	This model is defined in [1]
Profile	A PROFILE documents a mandated, unique and fully defined configuration of standards and specifications for an application or system under the STANAG 4609
Progressive Scan	The image is continuously scanned from left to right and from top to bottom using all pixels in the capture. This is opposed to interlace scanning used in conventional television, which scans from left to right for one line then skips every other line to form a field of the image. Then, the second field is made up of the lines that were skipped in the first field. The combination of the two fields constitutes a complete frame. It should be noted that progressive scan systems do not suffer the motion artifacts caused by interlace scanning, and the loss of vertical and temporal resolution caused by motion occurring between the scanned fields of an interlaced system.
Protocol	(1) [In general], A set of semantic and syntactic rules that determine the behavior of functional units in achieving communication. For example, a data link protocol is the specification of methods whereby data communication over a data link is performed in terms of the particular transmission mode, control procedures, and recovery procedures. (2) In layered communication system architecture, a formal set of procedures that are adopted to facilitate functional interoperation within the layered hierarchy. Note: Protocols may govern portions of a network, types of service, or administrative procedures.
Real-time Processing	AV command and control information including antenna positioning and AV video receipt and processing. Pertaining to the timeliness of data or information that has been delayed only by the time required for electronic communication. This implies that there are no noticeable delays.
Recommended Practice	Where the term A RECOMMENDED PRACTICE is used, the item documents a practice that further clarifies the implementation of a STANDARD or PROFILE in order to enforce interoperability across NATO systems
Reconnaissance	A mission undertaken to obtain, by visual observation or other detection methods, information about the activities and resources of an enemy or potential enemy; or to secure data concerning the meteorological, hydrographic characteristics of a particular area.
Resolution	A measurement of the smallest detail, which can be distinguished by a sensor system under specific conditions
Secondary Imagery	Secondary Imagery is digital imagery and/or digital imagery products derived from primary imagery or from the further processing of secondary imagery
Sensor	Equipment, which detects, and may indicate, and/or record objects and activities by means of energy or particles emitted, reflected, or modified by objects

Situational Awareness	Situational Awareness is the human perception of the elements of the operational environment in the context of forces, space and time, the comprehension of their meaning, and the projection of their status in the near future. A Situational Awareness Product is a concise, transportable summary of the state of friendly and enemy elements conveyed through information such as full-motion video (FMV), imagery, or other data that can contribute to the development of Situational Awareness either locally or at some distant node.
Software	A set of computer programs, procedures and associated documentation concerned with the operation of a data processing system, e.g. compilers, library routines, manuals, and circuit diagrams.
Standards	The Standardization Agreements (STANAGs), Military Standards (MIL-STDs), International Standards Organization (ISO), International Telecommunications Union (ITU) Recommendations and other International Standards contain provisions which, through references in this text, constitute provisions of this STANAG.
Storage	A) The retention of data in any form, usually for the purpose of orderly retrieval and documentation. B) A device consisting of electronic, electrostatic or electrical hardware or other elements into which data may be entered, and from which data may be obtained.
Surveillance	The systematic observation of aerospace, surface or subsurface areas, places, persons, or things, by visual, aural, electronic, photographic, or other means
System specification (a spec)	The document which accurately describes the essential equipment requirements for items, materials or services, including the procedures by which it will be determined that the requirements have been met.
Technical Architecture	A minimal set of rules governing the arrangement, interaction, and interdependence of the parts or elements whose purpose is to ensure that a conformant system satisfies a specific set of requirements. It identifies system services, interfaces, standards, and their relationships. It provides the framework, upon which engineering specifications can be derived, guiding the implementation of systems. Simply put, it is the “building codes and zoning laws” defining interface and interoperability standards, information technology, security, etc.
Television Imagery	Imagery acquired by a television camera and recorded or transmitted electronically
Unmanned Aerial Vehicle	A powered, aerial vehicle that does not carry a human operator; uses aerodynamic forces to provide vehicle lift, can fly autonomously or be piloted remotely; can be expendable or recoverable, and can carry a lethal or non-lethal payload. Also called a UAV.
Video Imagery	Images, with metadata collected as a timed sequence in standard motion imagery format, which is managed as a discrete object and displayed in sequence. Video imagery is a subset of the class of motion imagery.

MOTION IMAGERY SYSTEMS

TABLE OF CONTENTS

1 General

Motion Imagery (MI) is a valuable asset for commanders that enable them to meet a variety of theatre, operational and tactical objectives for intelligence, reconnaissance and surveillance. STANAG 4609 is intended to provide common methods for exchange of MI across systems within and among NATO nations. STANAG 4609 is intended to give users a consolidated, clear and concise view of the standards they will need to build and operate motion imagery systems. The STANAG includes guidance on uncompressed, compressed, and related motion imagery sampling structures; motion imagery time standards, motion imagery metadata standards, interconnections, and common language descriptions of motion imagery system parameters.

STANAG 4609 mandates that all visible light MI systems used by participating nations shall be able to decode all MPEG-2 transport streams with MPEG-2 compressed data types (Standard Definition, Enhanced Definition, High Definition) up to and including MISM Level 9M and all H.264 compressed data types up to and including MISM Level 9H, but each Nation may choose to ORIGINATE one, two or all data types. Levels 9M and 9H are defined in the Motion Imagery System Matrix (MISM) as found in AEDP-8 [2].

Likewise, STANAG 4609 mandates that all Infrared MI systems used by participating nations shall be able to decode all MPEG-2 transport streams with MPEG-2 compressed data types up to and including MISM Level 8M and all H.264 compressed data types up to and including MISM Level 8H, but each Nation may choose to ORIGINATE either compression type at whatever level it chooses. The levels of the IR System Matrix are found in Edition 3 of AEDP-8. The objective of STANAG 4609 is to provide governance so as to allow participating nations to share MI to meet intelligence, reconnaissance, surveillance and other operational objectives with interoperable MI systems.

2 Relation with Other Standards

The technology outlined in STANAG 4609 is based on commercial systems and components designed to defined open standards. No single commercial motion imagery standard provides all of the guidance necessary to build interoperable systems for use across the diverse missions of NATO; therefore STANAG 4609 is a profile of standards and practices on how component systems based on commercial standards can interconnect and provide interoperable service to NATO users.

STANAG 4609 and associated AEDP identify commercial standards that support interoperability for motion imagery environments and systems (such as common control vans, interconnections nodes, and NATO command centres), spanning high bandwidth transmission of uncompressed to lower bandwidth transmission of compressed motion imagery (video) signals. STANAG 4609 and associated AEDP also identify approaches for interoperability between high bandwidth and low bandwidth systems.

The core attributes of STANAG 4609 for motion imagery can be expressed in a "Simplified Motion Imagery System Matrix," as shown in Table 1. The cornerstone of this matrix is MPEG-2 or [3].

	Image / Structure	Serial Interface (reference only)	Compression	Stream	Simple File	Moderate File	Rich File
High Definition	SMPTE 296M[49] 274M [10] , 295M[52]	SMPTE 292M[9]	MPEG-2 MP@HL H.264 HP@L4.0 H.264 MP@L4.0 H.264 MP@L3.2	MPEG-2 TS	MPEG-2 TS or PS	MXF[1]	MXF[1]
Enhanced Definition	ITU Rec.1358 SMPTE 294M [53]	SMPTE 349M[14]	MPEG-2 MP@HL H.264 MP@L3 (L3.1 > 30 FPS)	MPEG-2 TS	MPEG-2 TS or PS	MXF[1]	MXF[1]
Standard Definition	ITU-R BT.601-5[7]	SMPTE 259M[13]	MPEG-2 MP@HL H.264 MP@L3	MPEG-2 TS	MPEG-2 TS or PS	MXF[1]	MXF[1]
Metadata	SMPTE 335M[22], 336M[23], RP210[25], EG37[54]		N/A	MPEG-2 TS	MPEG-2 TS or PS	MXF[1]	MXF[1]

Table 1: Simplified Motion Imagery System Matrix

3 Motion Imagery Operations Concept

3.1 Motion Imagery

MOTION Imagery is defined in the preceding Terms and Definitions section.

3.2 Other Video Systems

Video teleconference, telemedicine and support systems are not considered for this STANAG. If the applicability of the standards and recommended practices given in the STANAG 4609 are deemed practical for such systems, implementation with STANAG 4609 is encouraged to foster broader interoperability across NATO.

[1] Not mandated as of this Edition of STANAG 4609

Terms of Reference

STANDARDS

The Standardization Agreements (STANAGs), Military Standards (MIL-STDs), International Standards Organization (ISO), International Telecommunications Union (ITU) Recommendations and other International Standards contain provisions which, through references in this text, constitute provisions of this STANAG. Upgrading of referenced standards is conditional on satisfactory analysis of the impact to the STANAG.

PROFILES

A PROFILE documents a mandated, unique and fully defined configuration of standards and specifications for an application or system under the STANAG 4609.

RECOMMENDED PRACTICES

The term RECOMMENDED PRACTICE documents a practice that further clarifies the implementation of a STANDARD or PROFILE in order to enhance interoperability across NATO systems. Recommended practices are found in the Allied Engineering Documentation Publication [2].

ENGINEERING GUIDELINES

Engineering Guidelines represent well-defined, informative engineering principals. Engineering Guidelines are not mandated and are found in [2].

4 Frame Rate Annotation

STANAG 4609 uses the following scanning format nomenclature:

60p = 60 Frames per Second (FPS), Progressively Scanned
60p/1.001 = 59.94 FPS (NTSC compatible frame rate), Progressively Scanned

50p = 50 FPS, Progressively Scanned

30p = 30 FPS, Progressively Scanned
30p/1.001 = 29.97 FPS (NTSC compatible frame rate), Progressively Scanned

25p = 25 FPS, Progressively Scanned

24p = 24 FPS, Progressively Scanned
24p/1.001 = 23.98 FPS (NTSC compatible frame rate), Progressively Scanned

30i = 30 FPS, Interlace Scanned, yielding 60 fields-per-second
Note that many commercial documents use the term 60i to mean 30i
30i/1.001 = 29.97 FPS (NTSC frame rate), Interlace Scanned
Note this is the frame rate associated with "television" in the United States

25i = 25 FPS, Interlace Scanned, yielding 50 fields-per-second
Note this is the frame rate associated with "television" in Europe

24i	=	24 FPS, Interlace Scanned, yielding 48 fields-per-second
24i/1.001	=	23.98 FPS (NTSC compatible frame rate), Interlace Scanned

5 Standard, Enhanced, and High Definition

STANAG 4609 uses the following scanning format definitions, defined by the commercial world, consistent throughout all of the specified profiles (see Motion Imagery System Matrix for detailed technical specifications for each profile):

High Definition (HD) is defined as spatial resolution at or greater than 1280x720 pixels, progressively scanned, at temporal rates at or greater than 24 Hz.

Enhanced Definition (ED) is defined as spatial resolution of at least 720x480 pixels at 60 Hz or 720x576 pixels at 50 Hz, progressively scanned. Enhanced definition provides twice the scanning lines of standard definition.

Standard Definition (SD) is defined as interlaced scanned format at 720x576 at 50 Hz or 720x480 at 60 Hz.

6 Motion Imagery Roadmap

NATO user communities have diverse mission requirements and will select motion imagery systems across both a range of capabilities and a wide spectrum of bandwidth and system performance. Not all users will require a migration to the highest possible spatial and temporal resolution, but all users should be aware of a frame of reference that includes a spectrum of capabilities from standard definition to advanced high definition. In a digital Motion Imagery architectural construct, the specific pixel density of an origination system does not directly relate to the end-to-end required bandwidth: variables include desired image quality/performance, pixel density, pixel bit depth, imagery type and context, compression ratio, latency, error robustness, and other engineering trade spaces. Therefore, the frame of reference describes a continuum of capabilities that each Nation must consider to meet their specific needs.

The fundamental direction for NATO motion imagery systems is towards industry-adopted standards for digital motion and image processes that feature progressively scanned square-pixel images, and greater spatial, temporal, and spectral resolutions as technology affords. Interlaced scanning systems are to be treated as legacy systems and shall be replaced with progressive systems at the end of their service lives.

Standard Definition analogue interlace systems, which marked the legacy initial state, are formally considered to be obsolete systems within NATO. New systems may not be replaced with such legacy analogue systems. Within analogue families, component signal processes (R:G:B, Y:R-Y:B-Y, Y:C) are always preferred over composite signal processes (such as NTSC or PAL).

Standard Definition digital interlace ([7] component processing), systems with Serial Digital Interfaces (SDI, SMPTE 259M[13]/291M [15]) are a logical and economical upgrade from analogue interlace systems. However, as the cost differential between standard-definition, digital-interlace systems and enhanced-definition digital-progressive systems continues to decrease a migration to enhanced definition is strongly advised.

Enhanced Definition, digital progressive systems, such as 720 x 480 x 60p (480p) and 720 x 576 x 50p (576p), yield an optimal combination of improved spatial and temporal resolution at minimal differential costs as compared to current broadcast quality digital interlace [7] systems. However, 480p and 576p systems do not utilize square pixels, and thus have insufficient horizontal pixel resolution to deliver 16:9 aspect-ratio imagery. Therefore, enhanced definition may offer a suitable objective end-state for imagery systems that have no requirements to move to higher definition spatial or temporal resolutions or require a wider (16:9) aspect ratio.

High Definition (HD) progressive-scan imagery (SMPTE 296M [49]) is the near-term desired end-state for NATO motion imagery systems. 1280 x 720 x (50p) 60p is the target HD format for all existing and currently planned motion imagery collection systems that will be fielded over the next five to ten years. 1920 x 1080 x (50p) 60p is anticipated to become the revised end-objective in approximately five years once the technology matures. User communities that do not require high temporal resolution may consider use of 1920 x 1080 x 24p/25p/30p systems in special limited applications with controlled environments, such as studio production, training, etc. The dynamic geo-political landscape and military battle space environment will necessitate an application-specific optimization in spatial and temporal resolution; however, 1280 x 720 x (50) 60p will remain an architectural end-goal.

STANDARDS

TABLE OF CONTENTS

1 Sampling Structures

1.1 STANDARD 0202 - Standard Definition Digital Motion Imagery Sampling Structure

Component (4:2:2) digital video [7] shall be the NATO STANDARD sampling structure for baseband (uncompressed) standard definition motion imagery signals.

Furthermore, while both 10 bit and 8 bit (per component) implementations are allowed under the standard, 10 bit implementations are recommended.

Note 1: Once Motion Imagery has been originated in digital format or converted from legacy analog to standardized digital formats, it must remain in its digital format. Dual standard (525/30i /625/25i) analog display devices may be used as termination elements of an otherwise all-digital motion imagery system.

Note 2: It is recommended that in the event of transitional sampling, compression conversion, format conversion or processing, [7] shall be used as the intermediate sampling structure (within bit-serial interface input/output signal processing equipment) for subsequent use in further processing nodes.

1.2 STANDARD 0219 - Analog Video Migration

All NATO motion imagery production systems that currently use analog video waveforms (to include legacy STANAG 3350 systems) shall convert to Component (4:2:2) digital sampling structure [7] as soon as practical in the image processing chain.

Furthermore, all new digital baseband motion imagery system production sampling structures shall conform to Component (4:2:2) sampling structures [7].

Furthermore, unique mission systems with legacy analog video waveforms should convert such analog video waveforms to Component (4:2:2) sampling structures [7] as soon as possible in the signal processing chain, with no processing node backwards conversions to analog waveforms allowed.

1.3 STANDARD 0211 - Progressively Scanned Enhanced Definition Digital Motion Imagery

The NATO STANDARD motion imagery sampling structure for progressively scanned, digital, enhanced definition motion imagery systems shall be defined by [8].

Furthermore, while both 10 bit and 8 bit (per pixel) implementations are allowed under the standard, 10 bit implementations are recommended.

Note 1: It is mandated that once Motion Imagery has been originated in digital format or converted from legacy analog to standardized digital formats, it must remain in its digital format. Analog display devices may be used as termination elements of an otherwise all-digital motion imagery system.

Note 2: It is recommended that in the event of transitional sampling, compression conversion or processing, [8] shall be used as the intermediate sampling structure (within bit-serial interface input/output signal processing equipment) for subsequent use in further processing nodes

1.4 STANDARD 0210 - High Definition Television Systems (HDTV)

The NATO STANDARD motion imagery sampling structure for progressively scanned digital high definition systems based on 720 vertical scanning lines shall be defined by [9]. The parallel connector interface defined for [9] shall not be used.

Furthermore, while both 10 bit and 8 bit (per pixel) implementations are allowed under the standard, 10 bit implementations are recommended.

Note 1: It is mandated that once Motion Imagery has been originated in digital format or converted from legacy analog to standardized digital formats, it must remain in its digital format. Analog display devices may be used as termination elements of an otherwise all-digital motion imagery system.

Note 2: It is recommended that in the event of transitional sampling, compression conversion or processing, [9] shall be used as the intermediate sampling structure (within bit-serial interface input/output signal processing equipment) for subsequent use in further processing nodes.

The NATO STANDARD motion imagery sampling structures for progressively scanned digital high definition systems based on 1080 vertical scanning lines shall be defined by [10] (progressive only).

Note 1: It is mandated that once Motion Imagery has been originated in digital format or converted from legacy analog to standardized digital formats, it must remain in its digital format. Analog display devices may be used as termination elements of an otherwise all-digital motion imagery system.

Note 2: It is recommended that in the event of transitional sampling, compression conversion or processing, [10] (Progressive Mode Only) shall be used as the intermediate sampling structure (within bit-serial interface input/output signal processing equipment) for subsequent use in further processing nodes.

1.5 STANDARD 0203 - Digital Motion Imagery, Uncompressed Baseband Signal Transport and Processing

If Nations require interchange of motion imagery in baseband format, the following standards should be used.

SMPTE 259M [13], Level C (4:2:2) standard definition (270Mb/s Serial Digital Interface - SDI) is the STANDARD for uncompressed baseband signal transport and processing for Standard Definition digital motion imagery, audio and metadata origination, system interface, production/analysis center processing and manipulation.

SMPTE 349M [14] *Transport of Alternate Source Image Formats through SMPTE 292M* [9] is the STANDARD for uncompressed baseband signal transport and processing for Enhanced Definition digital motion imagery, audio and metadata origination, system interface, production/analysis center processing and manipulation.

SMPTE 292M [9] (1.5 Gb/s Bit-Serial Interface) is the STANDARD for uncompressed baseband signal transport and processing for High Definition digital motion imagery, audio and metadata origination, system interface, production/analysis center processing and manipulation.

2 Compression Systems

MPEG-2 [3] and H.264 [11] are the compression standards to be used for motion imagery.

2.1 STANDARD 0201 - Digital Motion Imagery Compression Systems

[3,4,5,6] (commonly known as MPEG-2) are the established NATO STANDARDs for all standard definition, enhanced definition and high definition compressed motion imagery, with the following PROFILE specifications:

For Standard Definition, the "MPEG-2, Main Profile @ Main Level" (MP@ML) shall be the standard definition motion imagery compression PROFILE.

For Enhanced Definition and High Definition, the "MPEG-2, Main Profile @High Level" (MP@HL) shall be the Enhanced Definition and High Definition motion imagery compression PROFILE for NATO origination, acquisition, production, manipulation, exploitation, distribution, archiving and end-user motion imagery product distribution, including real-time wide area transmissions.

Note 1: See Motion Imagery AEDP Recommended Practice 0220 for guidelines concerning applications constrained by low bandwidth channels and low motion imagery data rates

Note 2: See Motion Imagery AEDP Recommended Practice 0200 for guidelines concerning other digital motion imagery compression formats.

Note 3: Latency concerns for some applications (example UAV flight control and targeting) which require a low-latency compression mode, a low-latency compression mode is recommended, using a lower buffer-size and no B frames.

2.2 Advanced Digital Motion Imagery Compression Systems

H.264 [11] (Baseline, Main, Extended, and High Profiles – to be defined) is an advanced compression standard beneficial for applications constrained by bandwidth, which may not be adequately supported by MPEG-2. H.264 shall be carried over the MPEG-2 transport streams using [3]. See [2] for recommended practices.

2.3 STANDARD 0204 - Use of MPEG-2 System Streams

For streaming applications, MPEG-2 Transport Streams will be used for NATO applications.

2.4 STANDARD 0223 - Compressed High Definition Advanced Television (ATV) and Associated Motion Imagery Systems

Systems [3] and Video [4] (commonly known as MPEG-2) "High Level", which defines a broad family of high definition video compression capabilities, shall be the NATO STANDARD for compressed high definition advanced television and motion imagery, with the following PROFILE specifications:

The MPEG-2, Main Profile (4:2:0) @ High Level (MP@HL), shall be the high definition motion imagery compression PROFILE for NATO end-user motion imagery product distribution, including real-time wide area transmissions.

Furthermore, to promote universal interoperability, NATO high definition advanced television and motion imagery systems must be able to decode, process and display all of the diverse sampling structures and temporal rates within the MPEG-2 High Level profiles specified above, where the systems may either display the received signal in its native format or the signal may be re-formatted to the highest common progressive format supported by the system. The following specific motion imagery sampling formats and temporal rates are noted as a mandatory sub-set under the broader MPEG-2 High Level receiver umbrella:

Horizontal Resolution (pixels)	Vertical Resolution (pixels)	Frame Rate (Hz)	Aspect Ratio (H to V)
1920	1080	30p, 30p/1.001 30i, 30i/1.001 25p, 25i 24p	16:9
1280	720	60p, 60p/1.001 50p 30p, 30p/1.001 25p 24p	16:9
720	576	50p 25p, 25i 24p	16:9 4:3
720	480 483	60p, 60p/1.001 30p, 30p/1.001 30i, 30i/1.001 24p, 24p/1.001	16:9 4:3
640	480	60p, 60p/1.001 30p, 30p/1.001 24p, 24p/1.001	4:3

Note 1: For future enhancement and migration options, the following additional formats should be decoded by NATO MP@HL receiving systems, where the systems may either display the received signal in its native format or the signals may be re-formatted to the highest common progressive format supported by the display see [10]: 1920x1080, frame rates 60p, 60p/1.001, 50p; 16:9 Aspect Ratios

Furthermore, NATO high definition advanced television and motion imagery

ORIGINATION, ACQUISITION, PRODUCTION, MANIPULATION, and or **PROCESSING** systems must generate at least one of the following sampling formats and its associated temporal rates:

For High Definition applications:

1280 x 720, frame rates 60p, 50p, 30p, 25p, 24p; 16:9 Aspect Ratios
1920 x 1080, frame rates 30p, 25p, 24p; 16:9 Aspect Ratios

Note 2: For future enhancement and migration options, 1080 progressive scan formats (50p/60p) are included as future objectives for high definition motion imagery applications, but the MI TST notes that 1080 50p/60p systems are not yet commercially available. Therefore, 1080 50p/60p systems are not mandated under this STANAG. The MI TST will continue to periodically evaluate the availability of 1080 progressive scan format systems for future consideration.

Note 3: Dual mode interlaced and progressive scan systems are authorized under this STANAG profile, provided that for NATO applications, 1) only the progressive scan mode shall be used and 2) provided that the progressive scan mode is derived from a native progressive capture and is not derived from an interlaced image capture.

For Standard Definition applications **ORIGINATION, ACQUISITION, PRODUCTION, MANIPULATION**, and or **PROCESSING** systems must generate at least one of the following sampling formats and its associated temporal rates:

720 x 576, frame rates 50p, 25p, 25i, 24p; 16:9 or 4:3 Aspect Ratios
720 x 480 (483), frame rates 60p, 30p, 30i, 30i/1.001, 24p; 16:9 or 4:3 Aspect Ratios
640 x 480, frame rates 60p, 50p, 30p, 25p, 24p; 4:3 Aspect Ratios

Note 4: 720 horizontal pixels are the standard width for NATO standard and enhanced definition program origination and processing. NATO systems shall not originate imagery content using 704 horizontal pixels.

2.5 STANDARD 0206 - Motion Imagery Still Frames

STANAG 4545 [16] (NSIF 1.0) shall be the NATO STANDARD for digital still images that have been extracted from motion imagery sequences. Once an image has been captured for individual still image processing, exploitation and dissemination; the image is no longer considered to be motion imagery, and is therefore not subject to this STANAG (but must meet STANAG 4545 image standards).

Furthermore, still images should be extracted from full resolution bit-serial interface video streams, with direct conversion and storage into STANAG 4545 image formats (using no transitional analog processing steps).

Furthermore, still images may be directly extracted from MPEG-2 digital files provided there are no transitional analog processing steps.

2.6 STANDARD 0802 - Advanced Motion Imagery

MPEG-2 and H.264 compression methods may be inadequate for handling very large frame sizes (those falling into the category of Advanced Motion Imagery in this Document). One possible solution to the above stated problem is JPEG 2000. Until NATO develops a standardized way of dealing with very large frame sizes, the use of JPEG 2000 compression is recommended as a solution.

JPEG 2000 [17] is a wavelet based compression method and with high versatility and scalability. The JPEG 2000 standard allows for region-of-interest encoding and feature scalability, and is an emerging commercial technology used in digital cinema and other large image applications.

JPEG 2000 offers extensive features that accommodate large frame sizes, large numbers of spectral components and high bit-depth data. JPEG 2000 is recognized as the option of choice to accommodate advanced motion imagery sensors that are characterized by their very large frame sizes (10^8 pixels and larger).

Studies are underway to determine the most appropriate way to standardize this emerging technology (separate standard, best profiles, consistency with STANAG 7023...)

It is recommended that interested parties, when they cannot use the provisions of the main body of this STANAG, rely on Motion Imagery AEDP Recommended Practice 0708 for guidelines concerning compression of advanced motion imagery data using JPEG 2000.

2.7 STANDARD 0803 - Infrared Motion Imagery

Infrared (IR) motion imagery is defined as being in the spectral wavelengths from 1 to 14 µm. Standards and Recommended Practices for IR are similar to those in the motion imagery standards levels (MISL) for the visible spectrum. This section enumerates the standards, recommended practices, interoperability profiles, and engineering guidelines specifically designed for IR. Collectively this range of standards shall also be referred herein as “infrared” or “IR”. It is beneficial for IR to use motion imagery standards whenever possible to achieve the advantage of the higher volume, lower cost EO motion imagery product availability, utilize the same or similar modules for IR and EO motion imagery, and aid in fused products.

For Infrared motion imagery, frame rates of 25, 30, 50, and 60 are preferred, but lower and higher frame rates are allowed and tolerance in the system should allow for 1/1.001 of 30 Hz and 1/1.001 of 60 Hz. The resolution classes of IR are 160x120, 320x240, 640x480 (including 640x512, 720x480, 720x512, and 720x576), 1024x720 (including 1280x720 and 1024x1024), 1920x1080, and 2048x2048 progressively scanned. See Recommended Practice 0706 in Edition 3 of AEDP-8 for further details. Interlaced scanning IR systems are to be treated as legacy systems and shall be replaced with progressive systems at the end of their service lives. Infrared motion imagery typically has higher bit depths such as 12 and 14 bits, which are preferred.

If compression is needed for Infrared Motion Imagery, Systems [3] and Video [4] (commonly known as MPEG-2) and H.264 [11] shall be the NATO STANDARDs for compressed infrared motion imagery, with the following PROFILE specification. The MPEG-2, Main Profile Main Level (MP@ML) shall be the compression PROFILE for infrared motion imagery 720x480/30Hz and 720x576/25Hz for NATO origination, acquisition, production, manipulation, exploitation, distribution and archiving. The MPEG-2, Main Profile @ High Level (MP@HL) shall be the compression PROFILE for infrared motion imagery 1280x720/60Hz for NATO origination, acquisition, production, manipulation, exploitation, distribution and archiving. [12], called High Profile in [11] is recommended over MPEG-2 for providing higher bit depth, monochrome operation, and superior compression performance. The new High 4:4:4 Profile operated in monochrome mode is preferred because it provides 14-bit depth magnitude monochrome operation, and provides H.264 compression performance.

STANAG 4609 mandates that all IR motion imagery systems used by participating nations shall be able to decode all MPEG-2 transport streams with MPEG-2 compressed data types up to and including MISM Level 8M as defined in RP 0706 of Edition 3 of AEDP-8 and all H.264 compressed data types up to and including MISM Level 8H. However, each Nation may choose to ORIGINATE any level it chooses. The objective of STANAG 4609 is to provide governance so as to allow participating nations to share IR motion imagery to meet intelligence, reconnaissance, surveillance and other operational objectives with interoperable MI systems.

3 Metadata

All STANAG 4609 compliant systems will be designed to exploit, as a minimum, the set of metadata commonly agreed between the participating nations as required for interoperability; but, they will also accept, and pass-through without any system performance degradation, whatever syntax-compliant metadata is encountered.

3.1 STANDARD 0212 - Motion Imagery Metadata Dictionary Structure

SMPTE 335M [22], *Metadata Dictionary Structure*, is the NATO STANDARD for the interchange and structure definition of metadata dictionaries used by digital motion imagery systems/products.

3.2 STANDARD 0213 - Data Encoding using Key-Length-Value

SMPTE 336M [23], *Data Encoding Protocol Using Key-Length-Value,* is the NATO STANDARD protocol for encoding data essence and metadata into Motion Imagery streams, files, and associated systems. Universal sets are mandated for NATO use.

3.3 STANDARD 0207 - Metadata Dictionary

The *KLV Metadata Dictionary* [24], found at www.gwg.nga.mil/MISB, is the NATO standard for KLV keys. SMPTE RP210 [25], *SMPTE Metadata Dictionary Contents is the* NATO STANDARD for the identification of metadata elements encoded in digital motion imagery products not found in [24].

3.4 STANDARD 0208 - Embedded Time Reference for Motion Imagery Systems

SMPTE 12M [26], commonly known as SMPTE time code, shall be the NATO STANDARD for time annotation and embedded time references for motion imagery systems.

Furthermore, within SMPTE 12M, Drop Frame Time Code shall be used for 60/1.001, 30/1.001, 24/1.001 frames per second (FPS) systems. Non-Drop Frame Time Code shall be used for 60, 50, 30, 25, and 24 FPS systems.

SMPTE 309M [27] shall be the NATO STANDARD for precision time and date embedding into SMPTE 12M time code data streams.

Furthermore, within SMPTE 309M, NATO users will use the Modified Julian Date (MJD) (Y2K compliant) date encoding format and Universal Coordinated Time (UTC) as the time zone format.

Note: If Motion Imagery time code data is used as a data element for transference to other NATO systems (example NSIF still imagery), then the MJD / Time Code data will need to be translated to an appropriate date/time format for the application.

3.5 STANDARD 0214 - Time Code Embedding

If KLV Metadata is not available, and traditional time code (see [26, 27]) is used for date/time information, the following standards apply:

Digital Vertical Interval Time Code (D-VITC) shall be embedded on digital video line 9 of all [7] component (4:2:2) and bit-serial interface systems. Users may implement LTC for internal processing (such as in tape recorders) provided D-VITC is always forwarded to the next processing element on digital video line 9.

Furthermore, SMPTE Ancillary Time Code (embedded in the bit-serial interface Ancillary data space) may be used instead of D-VITC, provided such time code data is part of other metadata delivered by the ancillary data stream.

3.6 STANDARD 0215 - Time Reference Synchronization

Universal coordinated time (UTC, also known as “Zulu”), clock signals shall be used as the universal time reference for NATO SMPTE 12M [26] time code systems, allowing systems using time code to accurately depict the actual Zulu time of day of motion imagery acquisition / collection / operations.

Furthermore, when NATO “original video acquisition” motion imagery sequences are used as sources for editing onto new “edit master” sequences, the “edit master” sequence may have a new, continuous time code track. The time code for the new sequence should reflect the “document date” of the new motion imagery product.

Furthermore, Global Positioning System time, corrected to UTC, is the standard for the source of time data.

3.7 STANDARD 0218 - Timing Reconciliation Universal Metadata Set for Digital Motion Imagery

This standard (Appendix 1) defines a timing reconciliation metadata set to correct (reconcile) the original capture time of metadata with a User Defined Time Stamp usually associated with the capture time of the digital motion imagery or audio essence. Timing reconciliation metadata is not required if the application using the metadata does not depend on the amount of timing error or uncertainty between the metadata capture and the video or audio essence capture.

3.8 STANDARD 0216 - Packing KLV Packets into SMPTE 291 Ancillary Data Packets

If a Serial Digital Interface (see STANDARD 0203) is used, SMPTE RP 214 [28], *Packing KLV Encoded Metadata and Data Essence into SMPTE 291M* [15] *Ancillary Data Packets* is the NATO STANDARD for the encoding of metadata elements into Serial Digital Interface (SDI) SMPTE 291M [15] ancillary data packets.

3.9 STANDARD 0217 - Packing KLV Packets into MPEG-2 Systems Streams

If MPEG-2 is used with Nonsynchronized metadata, SMPTE RP 217 [29], *Nonsynchronized Mapping of KLV Packets into MPEG-2 System Streams,* is the NATO STANDARD for the non-synchronous encoding of metadata elements into MPEG-2 Systems Streams.

Note: To be STANAG compliant, KLV metadata in BOTH the Transport Stream and Program Stream must be identified by the registered format_identifier 0x4B4C5641 (“KLVA”). RP 217 [29] states that 0x4B4C5641 is the format_identifier to be used for the Transport Stream, but 0x4B4C5641 or “some other descriptor” may be used for the Program Stream.

If MPEG-2 is used with Synchronized metadata, [3] is mandated for the synchronous encoding of metadata for exchange of motion imagery and metadata files for collaboration of production work in progress among analysts; storage of work in progress for access by multiple users; and permanent archive of all contributions to a finished work.

3.10 STANDARD 0224 - Bit and Byte Order for Metadata in Motion Imagery Files and Streams

KLV Metadata in NATO Motion Imagery systems shall use Big-Endian in Byte order and Big-Endian in bit order.

Note: This is consistent with STANAG 4545 [16] and STANAG 7023 [21].

3.11 STANDARD 0209 - Use of Closed Captioning for Core Metadata Legacy Analog Video Encoding

EIA-608 [48] (Data Services), commonly known as closed captioning, shall be the NATO STANDARD for legacy system analog video vertical interval metadata encoding using video line 21.

Note: Analog video system data encoding is to be considered for legacy analog systems. New systems shall NOT use Closed Captioning, but must conform to all applicable digital motion imagery, audio, and metadata protocols specified in the STANAG. Furthermore, unique mission systems with legacy closed caption shall convert such closed caption data into KLV Metadata as soon as possible in the signal processing chain, with no processing node backwards conversions to closed captioning allowed.

3.12 STANDARD 0801 - Unmanned Aerial System (UAS) Datalink Local Metadata Set

3.12.1 Scope

Motion Imagery Standards Board MISB Standard 0801 [30] shall be the standard for all new and upgraded UAS systems. This standard references MISB Standard 0601 [38] found at www.gwg.nga.mil/misb as the document defining this standard. The Engineering Guideline found in Annex F of AEDP-8 is deprecated and is discouraged from use except in existing legacy systems.

The following information repeats the basic information for the Unmanned Air System (UAS) Datalink Local Data Set (LDS) for UAS platforms found in MISB Standard 0601 [31]. The UAS Datalink LDS is an extensible SMPTE (Society of Motion Picture Television Engineers) Key-Length-Value (KLV) Local Metadata Set designed for transmission through a wireless communications link (Datalink). In addition, this standard encourages the use of Standard 0601 in other platforms in addition to UAS.

This standard provides direction on the creation of a standard Local Data Set for a reliable, bandwidth-efficient exchange of metadata among digital motion imagery systems on UAV platforms. This standard also provides a mapping to Predator Exploitation Support Data (ESD) for continued support of existing metadata systems. The UAS Local Data Set metadata is intended to be produced locally within a UAS airborne platform and included in an MPEG2 Transport Stream (or equivalent transport mechanism). The MPEG2 Transport Stream (or equivalent) also contains compressed motion imagery from sensors such as Visual / Infrared video capture device.

Synchronization between the metadata and the appropriate video packet is also required for ensuring the validity of the metadata. The MPEG2 Transport Stream (or equivalent) embedded with UAS LDS metadata is then transmitted over a medium bandwidth (e.g. 1 to 5Mb/s) wireless Datalink and then disseminated.

The scope of this document is to provide a framework for an extensible bandwidth efficient Local Data Set, which enhances sensor, captured imagery with relevant metadata. This Standard also provides a mapping between UAS Datalink Local Data Set items, ESD items, and Universal Data Set (UDS) items defined in the latest SMPTE KLV dictionary (RP210 [25]) and in the MISB-managed Department of Defense (DoD) key-space [24].

3.12.2 Introduction

A SMPTE 336M [23] Universal Data Set (UDS) provides access to a range of KLV formatted metadata items. Transmitting the 16-byte key, basic encoding rules (BER) formatted length, and data value is appropriate for applications where bandwidth isn’t a concern. However, transmitting the 16-byte universal key quickly uses up the available bandwidth in bandwidth-challenged environments.

The Motion Imagery Standards Board (MISB) Engineering Guideline MISB EG 0104.5 [50] entitled “Predator UAV Basic Universal Metadata Set” shows a translation between basic ESD and Universal Data Set (UDS) metadata items that exist in the most current version of the SMPTE KLV dictionary. The UDS items in the MISB EG 0104.5 document are more appropriate for higher bandwidth interfaces (e.g. > 10Mb/s) like for dissemination, whereas this document targets low to medium bandwidth interfaces (e.g. 1 to 5Mb/s).

UAS airborne platforms typically use a wireless communications channel that allots a limited amount of bandwidth for metadata. Because of the bandwidth disadvantages of using a Universal Data Set, it is more desirable to use a Local Data Set for transmission over a UAS Datalink. As discussed in SMPTE 336M, a Local Data Set can use a 1, 2 or 4-byte key with a 1, 2, 4-byte, or BER encoded length. This UAS Local Data Set uses a BER encoded key and BER encoded length to minimize bandwidth requirements while still allowing the LDS ample room for growth.

This standard identifies a way to encode metadata locally in the airborne platform into a standard KLV Local Data Set. This standardized method is intended to be extensible to include future relevant metadata with mappings between new LDS, UDS, and ESD metadata items (where appropriate). When a new metadata LDS item is added or required, action must be taken to add an equivalent (i.e. identical in data format) Universal Data Set metadata item to the proper metadata dictionary (public or private) if the UDS metadata item does not already exist. This method also provides a mapping between Local Data Set items and currently implemented Universal Data Set items defined in the SMPTE KLV dictionary [25].

3.12.3 Local Data Set Changes and Updates

This document defines the UAS Datalink Local Metadata Set and is under configuration management. Any changes to this document must be accompanied by a document revision and date change and coordinated with the managing organization.

Software applications that implement this interface should allow for metadata items in the UAS Local Data Set that are unknown so that they are forward compatible with future versions of the interface.

3.12.4 UAS Datalink Local Data Set

This section defines the UAS Datalink Local Data Set (LDS). The keys that are supported in this LDS are defined and mapped to metadata items in the SMPTE KLV Dictionary [25] as well as the Exploitation Support Data (ESD) specification where appropriate. The UAS Datalink Local Metadata Set is SMPTE 336M [23] KLV compliant. The following section defines the metadata items contained in the LDS. The subsections that follow discuss the topics listed below:

- LDS Packet Structure
- Data Collection and Dissemination
- Time stamping
- Error Detection

The 16-byte Universal Key for this UAS Local Data Set is listed below:

Key: 06 0E 2B 34 - 02 0B 01 01 - 0E 01 03 01 - 01 00 00 00
Date Released: May 2006
Description: Released key defined in the MISB DoD Keyspace for the UAS LDS

A key history is provided below as a way to track the keys used in engineering and development. Note that the information below is informative only. DO NOT use the below keys in any future development.

Key: 06 0E 2B 34 - 01 01 01 01 - 0F 00 00 00 - 00 00 00 00
Date Released: November 2005
Description: Experimental node key used in software development efforts at General Atomics prior to the assignment of a defined key.

Key: 06 0E 2B 34 - 02 03 01 01 - 01 79 01 01 - 01 xx xx xx
Date Released: October 25, 2005
Description: This key was released as a placeholder within early versions of this document. Much development has been based around draft versions of this document, which has used this key in some software implementations.

3.12.5 LDS Packet Structure

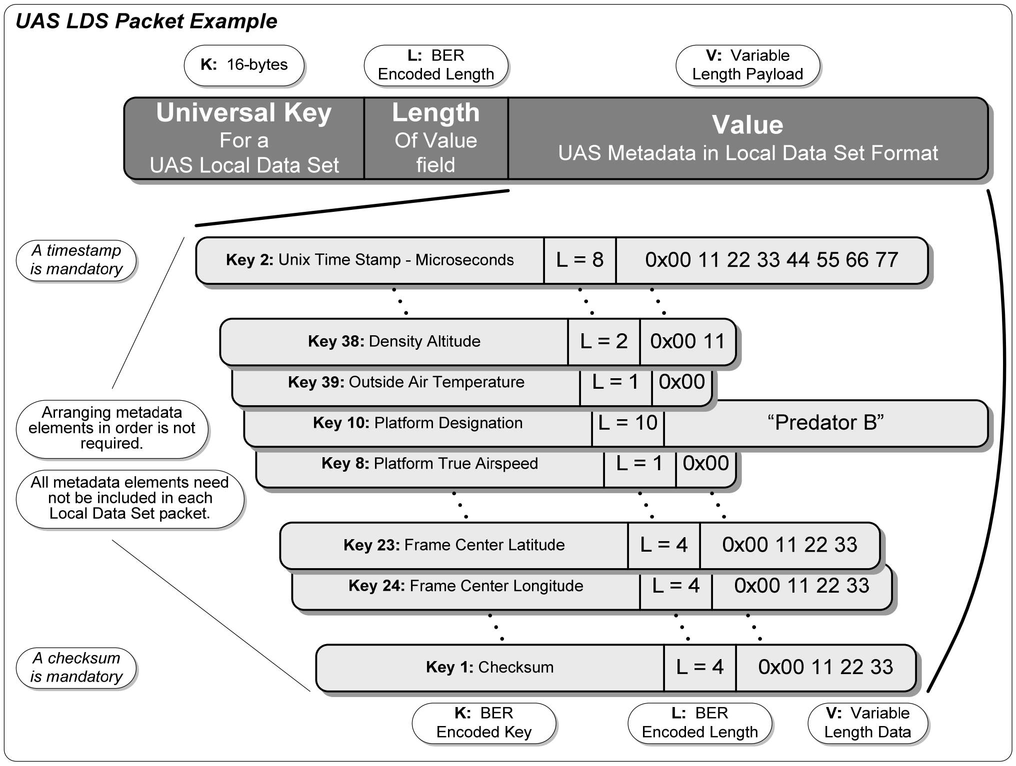

Figure 1: Example of a UAV Local Data Set Packet

Figure 1 shows the general format of how the LDS is configured. It is required that each LDS packet contain a Unix-based timestamp that represents the time of birth of the metadata within the LDS packet. Time stamping of metadata is discussed in a later section. A checksum metadata item is also required to be included in each LDS packet. Checksums are also discussed in a later section.

Any combination of metadata items can be included in a UAS Local Data Set packet. Also the items within the UAV LDS can be arranged in any order. However, the timestamp is always positioned at the beginning of an LDS packet, and similarly the checksum always appears as the last metadata item to support algorithms surrounding its computation and creation.

3.12.6 Bit and Byte ordering

All metadata is represented using big-endian (Most Significant Byte (MSB) first) encoding. Bytes are big-endian bit encoding (most significant bit (msb) first).

3.12.7 Key and Length Field Encoding

Both the LDS metadata item keys and length fields are encoded using basic encoding rules (BER) for either short or long form encoding of octets. This length encoding method provides the greatest level of flexibility for variable length data contained within a KLV packet.

In practice, the majority of metadata items in a LDS packet will use the short form of key and length encoding which requires only a single byte to represent the length. The length of the entire LDS packet, however, is often represented using the long form of length encoding since the majority of packets have a payload larger than 127 bytes. The key for the entire LDS packet is always 16 bytes. The length of a single packet is represented by 2 bytes whenever the payload portion of the LDS packet is less than 256 bytes. Both short and long form encoding is discussed in the subsections that follow. See SMPTE 336M [23] section 3.2 for further details.

3.12.8 BER Short Form Length Encoding Example

For UAS LDS packets and data elements shorter than 128 bytes, the length field is encoded using the BER short form (Figure 2). Length fields using the short form are represented using a single byte (8 bits). The most significant bit in this byte signals that the long form is being used. The last seven bits depict the number of bytes that follow the BER encoded length. An example LDS packet using a short form encoded length is shown below:

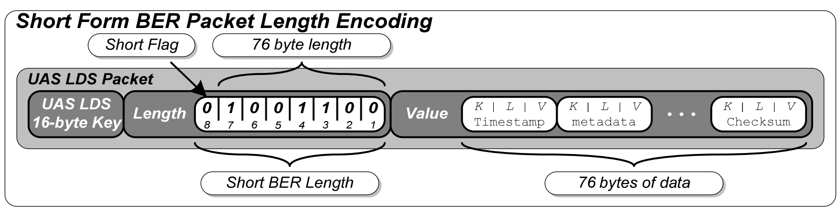

Figure 2: Example Short Form Length Encoding

Although this example illustrates the length representing the entire LDS packet, short form BER encoding also applies to the keys and lengths within the LDS packet.

3.12.9 BER Long Form Length Encoding

For LDS packets and data elements longer than 127 bytes, the length field is encoded using the BER long form. The long form encodes length fields using multiple bytes. The first byte indicates long form encoding as well as the number of subsequent bytes that represent the length. The bytes that follow the leading byte are the encoding of an unsigned binary integer equal to the number of bytes in the packet. An example LDS packet using a long form encoded length is shown in Figure 3:

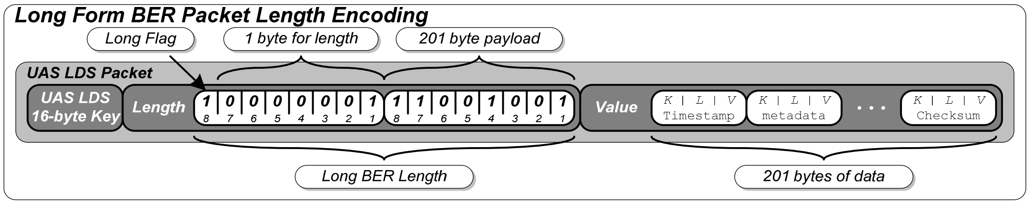

Figure 3: Example Long Form Length Encoding

Although this example depicts long form BER encoding on the length field of the entire LDS packet, long form BER encoding also applies to the keys and lengths within the LDS packet.

3.12.10 Data Collection and Dissemination

Within the air vehicle, metadata is collected, processed, and then distributed by the flight computer (or equivalent) through the most appropriate interface (SMPTE Serial Digital Interface (SDI), RS-422, 1553, Ethernet, Firewire, etc.). See Figure 4 below:

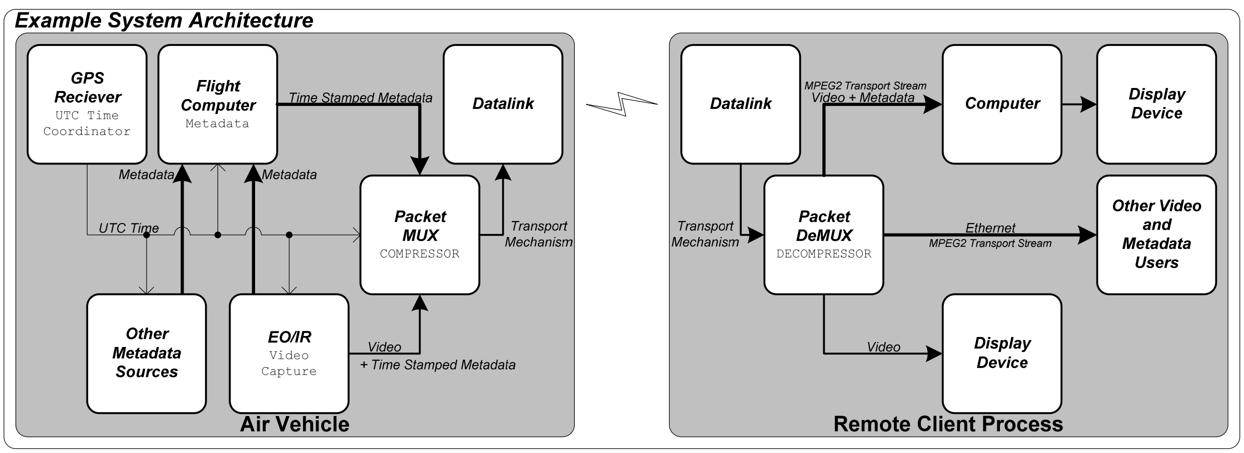

Figure 4: System Architecture

Sensors and other metadata sources pass metadata to the flight computer. The flight computer (or equivalent) places a timestamp into the UAS LDS packet prior to the Video Encoder / Packet Multiplexer (see the next section for more information about using timestamps in the LDS metadata packet.) The flight computer merges all metadata items, the timestamp, and the checksum into a LDS packet, which is then sent to the video encoder and Packet Multiplexer. The Packet Multiplexer merges the encoded video with the LDS metadata packet onto a transport mechanism for communication over a link to a remote client process. Subsequently, the client demultiplexes the encoded video and metadata after removal from the transport mechanism, decodes the video and processes the metadata. The motion imagery and metadata can then be displayed and distributed as appropriate.

3.12.11 Time Stamping

Every LDS KLV packet is required to include a Unix-based timestamp to relate the metadata to some standardized timing reference, which aides in associating imagery frames for subsequent downstream analysis. This section describes how to include a timestamp within a UAS Local Data Set packet.

Metadata sensing sources and the flight computer (or equivalent) are synchronized to operate on the same coordinated time, which is GPS derived. Either the source of metadata, or the flight computer, can thus provide a timestamp for inclusion in a LDS packet. The mandatory timestamp is named “Unix Timestamp”. The LDS timestamp (Key 2) is an 8-byte unsigned integer that represents the number of microseconds that have elapsed since midnight (00:00:00), January 1, 1970. This date is known as the UNIX epoch and is discussed in the IEEE POSIX standard [32].

A LDS timestamp is inserted at the beginning of the “value” portion of a LDS packet. The timestamp represented by Key 2 (Unix Timestamp) applies to *all* metadata within the LDS packet, and corresponds to the time of birth of all the data within the LDS packet. This timestamp can be used to associate the metadata with a particular video frame and can be displayed or monitored. An example LDS packet containing a timestamp is shown in Figure 5:

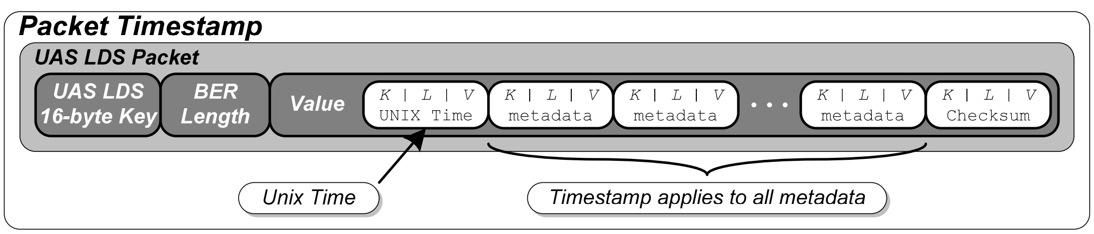

Figure 5: Packet Timestamp Example

3.12.12 Error Detection

To help guard against errors in the metadata after collection a 16-bit checksum is included in every Local Data Set packet. The checksum may be located anywhere within the packet, but is recommended to be placed at the end to facilitate ease in processing. The checksum represents a running 16-byte sum over a complete LDS packet beginning with the 16-byte Local Data Set key and ending with the length field of the checksum data item. Figure 6 shows the data that the checksum is performed over:

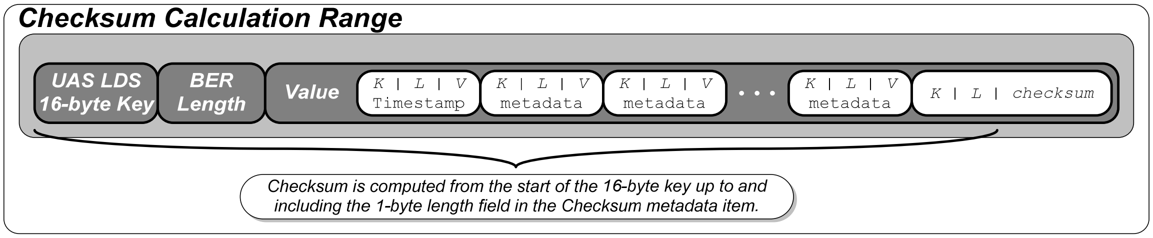

Figure 6: Checksum Computation Range

An example algorithm for calculating the checksum is as follows:

```
Unsigned short bcc_16 (
        Unsigned char * buff,     // Pointer to the first byte in the 16-byte UAS LDS key.
        Unsigned short len)       // Length from 16-byte UDS key up to 1-byte checksum length.
        {
         // Initialize Checksum and counter variables.
          Unsigned short bcc = 0, i;

         // Sum each 16-bit chunk within the buffer into a checksum
        for (i = 0; i < len; i++)
                bcc += buff[i] << (8 * ((i + 1) % 2));
                 return bcc;
        }       // end of bcc_16 ()
```

If the calculated checksum of the received LDS packet does not match the checksum stored within the packet, the packet is discarded. A lost LDS packet may have little impact since another packet will be available within reasonable proximity (in both data and time). In any event, the data cannot be trusted so it must not be used.

3.12.13 UAS Local Data Set Tables

See [31] at www.gwg.nga.mil/misb for the definitions of the content of the UAS Local Data Set as well as translations between the local data set and the universal data set (Predator Exploitation Support Data found in Annex F of AEDP-8.)

STANAG 4586 and STANAG 4609 Minimum KLV Metadata Elements (Subset of UAS Local Metadata Set)

The following paragraphs and table reflect the KLV metadata implementation that was agreed to by STANAG 4586 [33] on UAS, STANAG 4609, and by the US Motion Imagery Standards Board. This section contains information regarding common metadata parameters, which should be used by a STANAG 4586 compliant Unmanned Air Vehicle Control System (UCS). Table 2 provides a comprehensive list of metadata elements from Standard 0601 [31] UAS Datalink Local Metadata Set, which has been adopted by many existing UAV systems.

"X" in the first column indicates that the particular element should be implemented in a STANAG 4586 compliant UCS in order to enhance imagery exploitation for that system and is required for STANAG 4609 compliance. If the particular element is implemented in a STANAG 4586 compliant UCS, then it shall be applicable to the UCS interface specified in the second column of the table - either the Command Control Interface (CCI) only, or both the CCI and Data Link Interface (DLI) as defined in STANAG 4586. Refer to STANAG 4586 for actual mapping of these elements to the DLI.

Table 2: Standard 0601 [31] KLV Metadata Elements

Mandatory Elements [ii]	*DLI / CCI* [iii]	*UAS LDSKey* [i]	*Name* [i]
X	**Co**	**1**	Checksum
X	**D&C**	**2**	UNIX Time Stamp
X	**Co**	**3**	Mission ID
	Co	**4**	Platform Tail Number
X	**D&C**	**5**	Platform Heading Angle
X	**D&C**	**6**	Platform Pitch Angle
X	**D&C**	**7**	Platform Roll Angle
	Co	**8**	Platform True Airspeed
	Co	**9**	Platform Indicated Airspeed
X	**Co**	**10**	Platform Designation
X	**D&C**	**11**	Image Source Sensor
X	**Co**	**12**	Image Coordinate System
X	**D&C**	**13**	Sensor Latitude
X	**D&C**	**14**	Sensor Longitude
X	**D&C**	**15**	Sensor True Altitude
X	**D&C**	**16**	Sensor Horizontal Field of View
X	**D&C**	**17**	Sensor Vertical Field of View
X	**D&C**	**18**	Sensor Relative Azimuth Angle
X	**D&C**	**19**	Sensor Relative Elevation Angle
X	**D&C**	**20**	Sensor Relative Roll Angle
X	**Co**	**21**	Slant Range
X	**Co**	**22**	Target Width
X	**Co**	**23**	Frame Center Latitude

X	**Co**	**24**	Frame Center Longitude
X	**Co**	**25**	Frame Center Elevation
	Co	**26**	Offset Corner Latitude Point 1
	Co	**27**	Offset Corner Longitude Point 1
	Co	**28**	Offset Corner Latitude Point 2
	Co	**29**	Offset Corner Longitude Point 2
	Co	**30**	Offset Corner Latitude Point 3
	Co	**31**	Offset Corner Longitude Point 3
	Co	**32**	Offset Corner Latitude Point 4
	Co	**33**	Offset Corner Longitude Point 4
	D&C	**34**	Icing Detected
	Co	**35**	Wind Direction
	Co	**36**	Wind Speed
	D&C	**37**	Static Pressure
	D&C	**38**	Density Altitude
	D&C	**39**	Outside Air Temperature
	Co	**40**	Target Location Latitude
	Co	**41**	Target Location Longitude
	Co	**42**	Target Location Elevation
	Co	**43**	Target Track Gate Width
	Co	**44**	Target Track Gate Height
	Co	**45**	Target Error Estimate - CE90
	Co	**46**	Target Error Estimate - LE90
	Co	**47**	Generic Flag Data 01
X	**Co**	**48**	Security Local Metadata Set
	D&C	**49**	Differential Pressure
	D&C	**50**	Platform Angle of Attack
	D&C	**51**	Platform Vertical Speed
	D&C	**52**	Platform Sideslip Angle
	Co	**53**	Airfield Barometric Pressure
	Co	**54**	Airfield Elevation
	Co	**55**	Relative Humidity
	D&C	**56**	Platform Ground Speed
	Co	**57**	Ground Range
	D&C	**58**	Platform Fuel Remaining

	Co	**59**	Platform Call Sign
	Co	**60**	Weapon Load
	Co	**61**	Weapon Fired
	Co	**62**	Laser PRF Code
	Co	**63**	Sensor Field of View Name
	D&C	**64**	Platform Magnetic Heading
X	**D&C**	**65**	UAS LDS Version Number
	Co	**66**	Target Location Covariance Matrix
	D&C	**67**	Alternate Platform Latitude
	D&C	**68**	Alternate Platform Longitude
	D&C	**69**	Alternate Platform Altitude
	D&C	**70**	Alternate Platform Name
	D&C	**71**	Alternate Platform Heading
	Co	**72**	Event Start Time - UTC
	Co	**73**	Remote Video Terminal LDS Conversion

Table notes:

i. The element name and tag refers to MISB STANDARD 0601 [31] UAS Datalink Local Metadata Set.

ii. Elements marked with an ”X” to be included in a STANAG 4586 UCS as an extended list of elements, oriented for image exploitation.

iii. (**Co**): The element shall be available at the CCI only.
(**D&C**): The element shall be available at the DLI and the CCI.

3.13 STANDARD 0901 - Security Metadata Universal Set for Digital Motion Imagery

3.13.1 Scope

This standard describes the use of security metadata in MPEG-2 digital motion imagery applications. For applications involving national security it is mandatory that each part of a motion imagery file be marked correctly and consistently with security classification and other security administration information. The standard shall be applied to all MPEG-2 motion imagery implementations and shall be used to link security metadata to essence (video, audio, or data) and/or other metadata. The standard describes the security metadata structure and format, and not where it is done in the processing chain.

This standard defines the format of embedding security metadata in MPEG-2 files only; in particular, it addresses the MPEG-2 transport protocol and is independent of the compression used for the video essence. The methods used to gather security information, create files and insert security-metadata into files are the responsibility of application system developers. Similarly, the proper display of security information on screens, computer displays, printed output, etc. is the responsibility of system application developers. Originators and application users are responsible for the proper handling, and ultimately, for the use and disposition of classified information.

3.13.2 Security Metadata Set for Digital Motion Imagery

This standard defines the contents and the application of a Security Metadata Set in digital motion imagery. The first section explains the individual elements that are normative in the SMPTE Metadata Dictionary [25] and the MISB Metadata Registry [24]. The construction of a Security Metadata Set from these elements follows SMPTE Standard 335M [22] and uses the KLV metadata encoding protocol. Finally, this standard defines how the Security Metadata Set shall be used for tagging essence and other metadata sets in MPEG-2 Transport Streams (TS), Program Streams (PS), and files.

The sections of this standard are applicable only to MPEG-2 bitstreams. The standard shall be followed to ensure that all parts of an MPEG-2 TS or PS are tagged correctly with security information for use by applications. All metadata shall be represented using big-endian (most significant byte – MSB – first) encoding. Bytes shall be big-endian bit encoding (most significant bit – msb – first).

3.13.2.1 Security Metadata Elements

The following Security metadata elements comprise information needed to comply with CAPCO and other referenced security directives. These normative documents govern when certain fields are mandatory and when fields are optional. Security requirements may dictate that some or all entries are mandatory. In all applications the presence or absence of certain metadata will depend on the context of the application and its unique security requirements. Whenever there is conflict between this standard and directions of Security officials on the required presence or absence of entries the direction of Security officials takes precedence. Table 2 presents a summary of metadata elements within the Security Metadata Universal Set.

3.13.2.2 Security Classification

This metadata element contains a value representing the entire security classification of the file in accordance with U.S. and NATO classification guidance. Values allowed are: TOP SECRET, SECRET, CONFIDENTIAL, RESTRICTED, and UNCLASSIFIED (all caps) followed by a double forward slash “//”. This is a mandatory entry whenever the Security Metadata Sets are used.

3.13.2.3 Classifying Country Releasing Instructions Country Coding Method

This metadata element identifies the country coding method for the Classifying Country and Releasing Instructions metadata. The Country Coding Method shall use FIPS 10-4 [57] two-letter or four-letter alphabetic country code; ISO-3166 [34] two-letter, three-letter, or 3-digit numeric; or STANAG 1059 [35] two-letter, three-letter, or 3-digit numeric codes.

Example of Country Coding Method: ISO-3166 Two Letter

3.13.2.4 Classifying Country

This metadata element contains a value for the classifying country code preceded by a double slash "//." The default is the FIPS 10-4 two-letter code.

Example of classifying country: //**DEU** (Example of ISO-3166 code)
//**UK** (Example of default FIPS 10-4 code)

3.13.2.5 Sensitive Compartmented Information (SCI) / Special Handling Instructions (SHI)

If the classification of any material in the transport stream or file is Top Secret, Secret, or Confidential and requires special handling, then SCI/SHI digraphs, trigraphs, or compartment names must be added identifying a single or a combination of special handling instructions. A single entry shall be ended with a double forward slash "//". Multiple digraphs, trigraphs, or compartment names shall be separated by a single forward slash "/" and the last entry shall be ended with a double forward slash "//". Multiple SCI/SHI digraphs, trigraphs, or compartment names shall be concatenated in one metadata element free-text entry and shall not be encoded as individual metadata elements in the Sets.

3.13.2.6 Caveats

This metadata element set contains a value representing all pertinent caveats/codewords from each category of the CAPCO register. These caveats form a field in the classification line marking. Entries in this field may be abbreviated or spelled out. This field shall be used to indicate FOR OFFICIAL USE ONLY, which may be abbreviated as FOUO. The caveat FOUO shall always be preceded by the classification element containing the string UNCLASSIFIED// and shall not stand alone.

Examples of Caveats: **NOFORN**
REL TO
RELEASABLE TO
FOR OFFICIAL USE ONLY
FOUO

3.13.2.7 Releasing Instructions

This metadata element contains a list of country codes to indicate the countries to which information in a digital motion imagery file is releasable. Multiple country codes shall be separated by a blank (space; NOT underscore). Multiple country codes shall be concatenated in one Releasing Instructions metadata element entry and shall not be encoded as individual metadata elements in the Sets. The use of blank spaces to separate country codes, instead of semi-colons or other characters, is to comply with security guidelines and to allow parsing of fields by automated security screening systems. The country code of the originating country shall appear first, then the country codes of other countries to which the data are releasable shall appear in alphabetical order, and, finally, the codes of any non-state organizations (such as NATO) to which the data are releasable shall appear in alphabetical order.

Example of Releasing Instructions: **USA DEU**

3.13.2.8 Classified By

This metadata element identifies the name and type of authority used to classify the file. The metadata element is free text and can contain either the original classification authority name and position or personal identifier, or the title of the document or security classification guide used to classify the material.

3.13.2.9 Derived From

This metadata element contains information about the original source file or document from which the classification was derived. The metadata element is free-text.

3.13.2.10 Classification Marking System

This metadata element identifies the classification or marking system used in this Security Metadata Set as determined by the appropriate security entity for the country originating the data. The entry shall be a free text field.

Example of Classification or Marking System: **XYZ Marking System**

3.13.2.11 Classification Reason

This metadata element contains the reason for classification or a citation from a document (see below). The metadata element is free-text.

3.13.2.12 Declassification Date

This metadata element provides either a date when the classified material may be automatically declassified or if it is subject to Manual Review (MR) and is exempt from automatic declassification. The declassification date format shall be YYYYMMDD or the letters “MR” shall be used.

3.13.2.13 Object Country Coding Method

This metadata element identifies the coding method for the Object Country Code metadata. The Object Country Coding Method shall use FIPS 10-4 two-letter or four-letter alphabetic country code; ISO-3166 two-letter, three-letter, or 3-digit numeric; or STANAG 1059 two-letter, three-letter, or 3-digit numeric codes. Use of this element is optional; its absence shall indicate that the default FIPS 10-4 two-letter code is be used in the Object Country Code element. Use of "Other" for Object Country Coding Method shall indicate that the entry for Object Country Code is free-text and should not be parsed.

3.13.2.14 Object Country Code

This metadata element contains a value identifying the country (or countries) that is the object of the video or metadata in the transport stream or file. Multiple country codes shall be separated by a semi-colon ";" (no spaces). Multiple country codes shall be concatenated in one Object Country Code metadata element entry and shall not be encoded as individual metadata elements in the Sets. Note: The use of the semi-colon to separate country codes, instead of blanks or other characters, is to allow processing by current, automated imagery processing and management tools.

3.13.2.15 Comments

This metadata element allows for security related comments and format changes that may be necessary in the future. This field may be used in addition to those required by appropriate security entity and is optional.

3.13.3 Security Metadata Universal Set

The individual metadata elements that comprise information needed to identify the security classification of MPEG-2 streams and files and other metadata are defined as SMPTE KLV metadata elements in [25] (and updated versions) and the MISB Metadata Registry [24].

The Security Metadata Universal Set 16-byte Universal Label Key shall be:

06 0E 2B 34 02 01 01 01 02 08 02 00 00 00 00 00

Required security and linking information shall be contained entirely within a Security Metadata Universal Set that conforms to SMPTE 336M [23] KLV Universal Set encoding rules. The Security Metadata Set shall be a compliant Universal Set as determined by the metadata originator. While it is possible that Security metadata could be expressed as a Global Set, a pack or even as a label, the decision was made to use the Universal Set to reduce ambiguity or chances for misinterpretation.

3.13.4 Security Metadata Local Set

The individual metadata elements that comprise information needed to identify the security classification of MPEG-2 streams and files and other metadata are defined as SMPTE KLV metadata elements in SMPTE RP210 [25] and the MISB Metadata Registry.

The Security Metadata Local Set 16-byte Universal Label Key shall be:

06 0E 2B 34 03 01 01 01 0E 01 03 03 02 00 00 00

Required security and linking information shall be contained entirely within a Security Metadata Local Set that conforms to SMPTE 336M KLV Local Set encoding rules.

3.13.5 Security Metadata Universal and Local Set Application in MPEG-2 Streams

Security Metadata Universal and Local Sets shall be associated with the information, which they describe by containing a link to some essence or metadata in the transport stream or file. The following metadata elements shall be used to associate Security Metadata Sets with essence (video, audio, and data) or metadata within MPEG-2 streams or files, which may contain multiple material types:

3.13.5.1 Metadata Links within MPEG-2 Streams

Any KLV metadata that conforms to SMPTE 336M [23] (whether individual metadata, sets, or packs) may be linked to MPEG-2 ES within TS or PS using the following unique MPEG-2 stream identifiers:

3.13.5.2 Unique Material Identifier (UMID)

If used, the 32-byte UMID defined by SMPTE 330M [51] shall be used to identify the essence to which security metadata is linked.

3.13.5.3 Stream ID

In MPEG-2 Program Streams the 8-bit stream_id specifies the type and number of the Elementary Stream. In MPEG-2 Transport Streams the stream_id may be set by the user to any valid value which correctly describes the Elementary Stream type. ([3], par 2.4.3.7 and Table 2-18.) The stream_id shall be the Value for the Stream ID metadata element.

3.13.5.4 Transport Stream ID

When multiple Transport Streams are present in a network environment the 16-bit transport_stream_id uniquely identifies a specific Transport Stream from any other Transport Stream to remove any ambiguity. Its value is defined by the originator ([3] Sec 2.4.4.5). The transport_stream_id shall be the Value for the Transport Stream ID.

3.13.5.5 Universal Label Key ID

The 16-byte Universal Label Key for the element, set or pack to which the Security Metadata Set is linked shall be the Value of the Universal Label Key ID.

3.13.5.6 Linking Security Metadata to MPEG-2 Streams

To indicate the security classification of individual MPEG-2 streams the appropriate link metadata elements shall be contained within a Security Metadata Set as follows:

Elementary Streams – Use of stand-alone ES formats is discouraged for the reasons cited in the MISB RP 0101, *Use of MPEG-2 Systems Streams in Digital Motion Imagery Systems* [36]. However, each Elementary Stream within a Transport Stream or Program Stream shall be associated with a valid Metadata Security Set by containing the one or more UMID or Stream ID metadata elements for the streams to which they apply. If the same Metadata Security Set applies to multiple Elementary Streams then the Metadata Security Set shall contain each of the UMIDs or Stream IDs separately in the Set.

Transport Streams – Each Transport Stream shall be associated with a valid Metadata Security Set by containing the UMID or Transport Stream ID metadata element for that Transport Stream. The Security Metadata Set for the Transport Stream shall convey all the security information for the highest classification Elementary Stream or metadata contained in the Transport Stream.

Program Streams – The UMID shall be used for directly linking Security metadata to identified Program Streams in their entirety. The Security Metadata Set for the Program Stream shall convey all the security information for the highest classification Transport Stream, Elementary Stream or metadata contained in the Program Stream.

3.13.5.7 Linking Security Metadata to Other Metadata

When a single metadata element is associated with a Security Metadata Set the Security Metadata Set shall contain Universal Label Key ID whose Value is the 16-byte Universal label Key for the single metadata element.

When some but not all metadata elements within a set or pack must be linked to a Security Metadata Set the Security Metadata Set shall contain each individual Universal Label Key ID for the metadata to which it is linked.

When all metadata in a set or pack is associated with a Security Metadata Set then the set or pack shall contain the Security Metadata Set with a Universal Label Key ID whose value is the Universal Label Key for the set or pack. If all metadata in an Elementary Stream is associated with the same Security Metadata Set then the two shall be associated using the method above for Elementary Streams.

3.13.5.8 Security Metadata without Links

Security Metadata Sets that do not contain a Stream ID link or a Transport Stream ID link to MPEG-2 streams or a Universal Label Key ID link to other metadata are non-compliant and prohibited. The presence of a stand-alone Security Metadata Set without links is ambiguous and presents a potential security hazard.

3.13.5.9 Classification of Metadata Security Sets

Every effort shall be made to keep the contents (values) within a Security Metadata Set Unclassified. When one or more elements in a Security Metadata Set must be classified they must be linked to another (or the same) Security Metadata Set by a Universal Label Key ID for the classified element(s).

If an entire Security Metadata Set must be classified it shall be linked to another (or the same) Security Metadata Set by the Universal Label Key ID for itself.

3.13.5.10 Security Metadata Set Repetition Rate

Security Metadata Sets shall be repeated at regular, short intervals such as every 5, 10, 15, 30, or 60 seconds. The maximum repetition interval shall be 60 seconds. Applications that produce very short motion imagery clips or segments of a few seconds in duration may need to repeat Security Metadata Sets as often as every frame.

3.13.5.11 Unclassified Essence and Metadata

When essence and/or metadata are unclassified the Security Metadata Set shall consist of the value “UNCLASSIFIED//” for Security Classification. Other entries in the Set that limit or clarify the classification are optional.

3.13.5.12 Partial Security Metadata Sets

For some classifications (e.g. unclassified, collateral), or other circumstances, not all metadata elements may be required. It is the responsibility of the originator and their cognizant security office to ensure that all appropriate security entries are used.

3.13.5.13 Absence of Security Metadata Sets in MPEG-2 Streams

The absence of one or more Security Metadata Sets cannot and shall not be construed as rendering an MPEG-2 stream or metadata as Unclassified. The proper insertion of Security Metadata Sets into MPEG-2 streams and the extraction of Security information is the responsibility of system developers. It is the responsibility of bitstream originators and system developers to incorporate continual checks for Security Metadata Sets in their applications.

3.13.5.14 Version Number

The version number of the Security Metadata Universal and Local Set for Digital Motion Imagery is indicated via the Version Key. For MISB RP 0102 [37] this key shall be required. In the absence of this key, the version RP 0102.3 shall be assumed.

Table 2 - Security Metadata Universal Set Elements (Normative)

16-byte UL	Name	Data Type or References	Allowed Values or References	Maximum or Default Length (Bytes)	Required/ Optional/ Context
06 0E 2B 34 01 01 01 03 02 08 02 01 00 00 00 00	Security Classification	ISO 7 bit Enumerated Text	TOP SECRET SECRET CONFIDENTIAL RESTRICTED UNCLASSIFIED	14	Required
06 0E 2B 34 01 01 01 03 07 01 20 01 02 07 00 00	Classifying Country and Releasing Instructions Country Coding Method	ISO 7 bit Enumerated Text	ISO-3166 [34] Two Letter ISO-3166 Three Letter ISO-3166 Numeric FIPS 10-4 [57] Two Letter FIPS 10-4 Four Letter 1059 Two Letter 1059 Three Letter 1059 Numeric	21 (40 max)	Required
06 0E 2B 34 01 01 01 03 07 01 20 01 02 08 00 00	Classifying Country	Enumerated Text preceded by '//'	FIPS 10-4 ISO-3166 STANAG 1059	6	Required
06 0E 2B 34 01 01 01 01 0E 01 02 03 02 00 00 00	Security-SCI/SHI Information	ISO 7 bit	Security Ref [55]	40	Context
06 0E 2B 34 01 01 01 03 02 08 02 02 00 00 00 00	Caveats	Free Text	Security Ref [56]	20 (32 max)	Context
06 0E 2B 34 01 01 01 03 07 01 20 01 02 09 00 00	Releasing Instructions	ISO 7 bit Free Text	Security Refs [31,35 56,58]	40	Context
06 0E 2B 34 01 01 01 03 02 08 02 03 00 00 00 00	Classified By	ISO 7 bit Free Text	Security Refs [57,58]	40	Context
06 0E 2B 34 01 01 01 03 02 08 02 06 00 00 00 00	Derived From	ISO 7 bit Free Text	Security Refs [57,58]	40	Context
06 0E 2B 34 01 01 01 03 02 08 02 04 00 00 00 00	Classification Reason	ISO 7 bit Free Text	Security Refs [57,58]	40	Context
06 0E 2B 34 01 01 01 03 02 08 02 05 00 00 00 00	Declassification Date	ISO 7 bit Free Text	YYYYMMDD or MR	8 (32 max)	Context
06 0E 2B 34 01 01 01 03 02 08 02 08 00 00 00 00	Classification and Marking System	ISO 7 bit Free Text	N/A	4400	Context

06 0E 2B 34 01 01 01 03 07 01 20 01 02 06 00 00	Object Country Coding Method	ISO 7 bit Enumerated Text	ISO-3166 Two Letter ISO-3166 Three Letter ISO-3166 Numeric FIPS 10-4 Two Letter FIPS 10-4 Four Letter 1059 Two Letter 1059 Three Letter 1059 Numeric Other	21 (40 max)	Optional
06 0E 2B 34 01 01 01 03 07 01 20 01 02 01 01 00	Object Country Codes	16-bit UNICODE string Free Text	Refs [31,58]	40	Optional
06 0E 2B 34 01 01 01 03 02 08 02 07 00 00 00 00	Classification Comments	ISO 7 bit Free Text	N/A	480	Optional
06 0A 2B 34 01 01 01 01 01 01 01 XY 00 00 00 00	UMID Video	SMPTE RP210 [25]	SMPTE 330M [51]	32	Context
06 0A 2B 34 01 01 01 01 01 01 02 XY 00 00 00 00	UMID Audio	SMPTE RP210	SMPTE 330M	32	Context
06 0A 2B 34 01 01 01 01 01 01 03 XY 00 00 00 00	UMID Data	SMPTE RP210	SMPTE 330M	32	Context
06 0A 2B 34 01 01 01 01 01 01 04 XY 00 00 00 00	UMID System	SMPTE RP210	SMPTE 330M	32	Context
06 0E 2B 34 01 01 01 03 01 03 04 02 00 00 00 00	Stream ID	Integer	[3]	1	Context
06 0E 2B 34 01 01 01 03 01 03 04 03 00 00 00 00	Transport Stream ID	Integer	[3]	2	Context
06 0E 2B 34 01 01 01 03 01 03 06 01 00 00 00 00	Item Designator ID (16 byte)	SMPTE 336M [23]	SMPTE 336M	16	Context
06 0E 2B 34 01 01 01 01 0E 01 02 05 04 00 00 00	Version	UInt16	Value is version number of this document; e. g. for RP 0102 [37], this value is 1024	2	Required

Table 3 - Security Metadata Local Set Elements (Normative)

Tag	Name	Data Type or References	Allowed Values or References	Maximum or Default Length (Bytes)	Required/ Optional/ Context
1	Security Classification	Unsigned Integer	UNCLASSIFIED (0x01) RESTRICTED (0x02) CONFIDENTIAL (0x03) SECRET (0x04) TOP SECRET (0x05)	1	Required
2	Classifying Country and Releasing Instructions Country Coding Method	Unsigned Integer	ISO-3166 [34] Two Letter (0x01) ISO-3166 Three Letter (0x02) FIPS 10-4 Two Letter (0x03) FIPS 10-4 Four Letter (0x04) ISO-3166 Numeric (0x05) 1059 Two Letter (0x06) 1059 Three Letter (0x07) 1059 Numeric (0x08) Other (0x09)	1	Required
3	Classifying Country	Enumerated Text	FIPS 10-4 ISO-3166 STANAG 1059	6	Required
4	Security-SCI/SHI information	ISO7	Security Ref [55]	40	Context
5	Caveats	Free Text	Security Ref [56]	20 (32 max)	Context
6	Releasing Instructions	Free Text	Security Refs [31,35,56,58]	40	Context
7	Classified By	Free Text	Security Refs [57,58]	40	Context
8	Derived From	Free Text	Security Refs [57,58]	40	Context
9	Classification Reason	Free Text	Security Refs [57,58]	40	Context
10	Declassification Date	Free Text	YYYYMMDD or MR	8	Context
11	Classification and Marking System	Free Text	N/A	40	Context
12	Object Country Coding Method	Unsigned Integer	FIPS-2 (Default) (0x01) ISO-3166 Two Letter (0x01) ISO-3166 Three Letter (0x02) ISO-3166 Numeric (0x03) FIPS 10-4 Two Letter (0x04) FIPS 10-4 Four Letter (0x05) 1059 Two Letter (0x06) 1059 Three Letter (0x07) 1059 Numeric (0x08) Other (0x09)	1	Optional
13	Object Country Codes	Free Text	Refs [31,58]	40	Optional

14	Classification Comments	Free Text	N/A	480	Optional
15	UMID Video	SMPTE RP210 [25]	SMPTE 330M [51]	32	Context
16	UMID Audio	SMPTE RP210	SMPTE 330M	32	Context
17	UMID Data	SMPTE RP210	SMPTE 330M	32	Context
18	UMID System	SMPTE RP210	SMPTE 330M	32	Context
19	Stream ID	Integer	[3]	1	Context
20	Transport Stream ID	Integer	[3]	2	Context
21	Item Designator ID (16 byte)	SMPTE 336M [23]	SMPTE 336M	16	Context
22	Version	UInt16	Value is version number of this document; e. g. for RP 0102 [37], this value is 0d04	2	Required

3.13.6 Conversion of Security Metadata Elements between Universal and Local Sets

For bandwidth efficiency, some elements in the local set are formatted differently than the Universal set equivalent. This section provides conversion information for the differing items.

Security Classification

From Universal Set to Local Set: Convert string to unsigned integer

From Local Set to Universal Set: Convert unsigned integer to all uppercase string

Classifying Country and Releasing Instructions Country Code

From Universal Set to Local Set: Convert string to unsigned integer

From Local Set to Universal Set: Convert unsigned integer to all uppercase string

Object Country Coding Method

From Universal Set to Local Set: Convert string to unsigned integer

From Local Set to Universal Set: Convert unsigned integer to all uppercase string

3.14 STANDARD 0802 - Time Stamping Compressed Motion Imagery

3.14.1 Scope

This standard defines methods to time stamp compressed video streams and to transport video and metadata asynchronously or synchronously in compressed motion imagery streams. Implementation methods are defined that leverage the transport layer of MPEG-2 for carriage of motion imagery streams of varying types and bit rates as defined in the Motion Imagery Standards Profile concept of “X on 2”. Specific compressed video formats covered include MPEG-2 and H.264.

3.14.2 Introduction

The MPEG-2 transport layer [3] provides an infrastructure for the carriage of video, audio and metadata in a single motion imagery stream as shown in the following diagram.

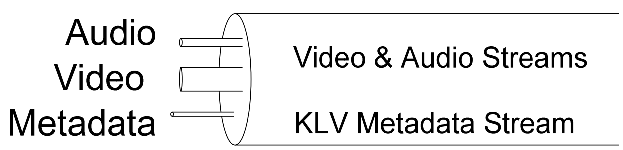

Figure 7: MPEG-2 Transport Stream (example)

The Motion Imagery Standards Profile (MISP) endorses the use of MPEG-2 Transport Streams for this purpose. The MISB has been researching the use of MPEG-2 Transport Streams for the carriage of other motion imagery formats in a study known as “Xon2”. Recent recommendations extend the use of MPEG-2 Transport Streams as a means for carriage of H.264 video in the compressed domain as defined in [3].

The advantages of using Universal Coordinated Time (UTC) as the master clock reference for video and metadata are outlined in MISB RP 0603 [40], Common Time Reference for Digital Motion Imagery using Coordinated Universal Time (UTC), which discusses several time formats and the relationships between them. MISB STANDARD 0604 [38], Time Stamping Compressed Motion Imagery defines how the UTC time can be used to stamp MPEG-2 and H.264 video streams, and how the video and metadata can be synchronously transported in motion imagery streams. A sample architecture showing how such a system can be configured is shown in Figure 8.

Motion imagery analysis and processing applications require various levels of temporal accuracy when referencing metadata elements and the video frames associated with those elements. Compressed imagery generated from standard definition analog video sensors has traditionally utilized asynchronous methods for carriage of metadata in a private data stream. This was adequate for metadata that was not time sensitive, or for metadata, which only needed to be associated to within a few seconds of the correct video frame. Asynchronous transport could not be used reliably for systems, which required metadata to be frame or sub-frame accurate.

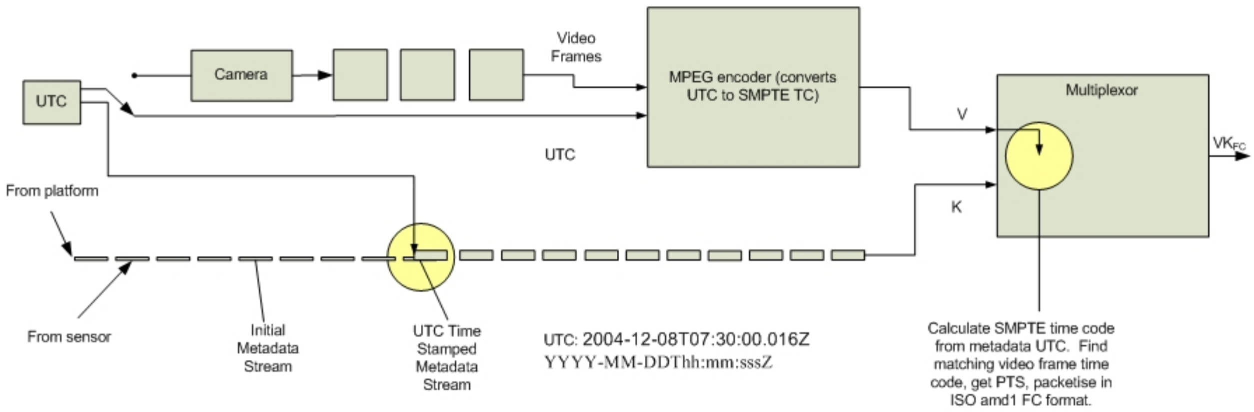

Figure 8: Metadata Synchronization

Synchronous multiplexing of metadata with video ensures that the proximity between a metadata item and the associated video is well defined. This in turn reduces the latency in the system and helps prevent the metadata from being separated from the associated video when the video is processed.

MISB STANDARD 0604 [38] provides guidance on methods to synchronously transport video frames and associated metadata elements with varying levels of precision as determined by the user's requirements.

3.14.3 Time Stamping Video

System designers should be aware of the accuracy requirements for the time stamps in their system. The use of UTC as a deterministic common time reference for the correlation of motion imagery frames and metadata is defined in MISB RP 0603 [40], which also describes several types of systems and the relative accuracies of each.

Time stamps may be introduced into a compressed video stream in one of two ways. If the uncompressed video signal contains a time stamp in the Vertical Interval Time Code (VITC) or the Vertical Ancillary Data Space (VANC), it is recommended that the encoder extract the time stamp from the VITC or VANC of the incoming video signal and insert it into the Video Elementary Stream as indicated in the following sections.

If the uncompressed video signal does not contain a time stamp, the encoder should be enabled to read the time stamp from the system time clock or an external source and insert it into the Video Elementary Stream.

The following sections describe how to insert the time stamp into MPEG-2 and H.264 video streams.

3.14.3.1 MPEG-2

Universal coordinated time (UTC, also known as “Zulu”), clock signals shall be used as the universal time reference for SMPTE 12M [26] time code systems, allowing systems using time code to accurately depict the actual Zulu time of day of motion imagery acquisition/collection/operations.

The following sections describe how to use the GOP (Group of Pictures) time code to time stamp MPEG-2 compressed video, and how a time stamp in the video elementary stream User Data field or a time stamp in the MPEG-2 video elementary stream editing information may be used in systems which require a more persistent time stamp or one with a higher level of precision.

GOP Time Code

The MPEG-2 video layer includes the definition of a time code within the Group of Pictures (GOP) header. This time code is of the form HH:MM:SS:FF, in a format specified by [26].

It is strongly recommended that the SMPTE time code in the GOP header be filled in with a time stamp, which represents UTC time for MPEG-2 video streams for all motion imagery systems.

The accuracy of the SMPTE 12M time code as it is inserted into the video signal for systems with integer and non-integer frame rates is indicated in MISB RP 0603 [40], Common Time Reference for Digital Motion Imagery using Coordinated Universal Time (UTC), and for cameras which are or are not phase locked to the master time reference.

For systems which process signals with integer frame rates, and for video sources that are genlocked to a UTC time reference, the accuracy of the time stamp in the GOP header can be quite accurate (sub-frame accuracy). The accuracy decreases for systems with non-integer frame rates.

The usefulness of the GOP Time Code has some limitations:

- The GOP Time Code is generally not persistent, and not considered by the MPEG-2 standards to be an absolute time. When a video is edited, the editor will often re-stamp the GOP Time Code.
- The GOP Time Code includes a time, but not a date. The date information, if needed, must be extracted from the KLV metadata in the stream.
- The accuracy of the GOP Time Code is limited, particularly in motion imagery with non-integer frame rates.

Some of these limitations can be addressed by also populating a time code in the elementary stream user data or MPEG-2 video elementary stream editing information as described in the following sections.

3.14.3.2 Elementary Stream User Data

The MPEG-2 format allows user defined data to be inserted into the video elementary stream in a user data field (start code = B2). The 13818 [4] specification allows the user data field to be placed in several different places in the video bitstream. The user data field containing the time stamp must be placed between the picture header and the picture data so that it relates to a frame of video.

The elementary stream user data time stamp may be used in systems which are required to associate a highly accurate, microsecond resolution time stamp with the video frame. This UTC time stamp shall be derived from GPS as described in section 4 of MISB RP 0603 [40] and will be formatted as defined in Annex A of same. The user data message consists of an identification string and a time stamp as defined below:

Identification String: 16 bytes that shall be set to the value:

Bytes 1-8: 0x4D, 0x49, 0x53, 0x50, 0x6D, 0x69, 0x63, 0x72,

Bytes 9-16: 0x6F, 0x73, 0x65, 0x63, 0x74, 0x69, 0x6D, 0x65

This represents the ASCII string: “MISPmicrosectime”

Time Stamp: 12 additional bytes defined as follows:

Byte 17: Status

Bit 7 0= GPS Locked (internal clock locked to GPS)
1= GPS Flywheel (internal clock not locked to GPS, so it is running on an internal oscillator)

Bit 6 0= Normal (time incremented normally since last message)
1= Discontinuity (time has not incremented normally since last message)

Bit 5 0= Forward (If Bit 6=1, this indicates that the time jumped forward)
1= Reverse (If Bit 6=1, this indicates that the time jumped backwards)

Bits 4-0: Reserved (=1)

Bytes 18, 19: Two MS bytes of Microseconds

Byte 20: Start Code Emulation Prevention Byte (0xFF)

Bytes 21, 22: Two next MS bytes of Microseconds

Byte 23: Start Code Emulation Prevention Byte (0xFF)

Bytes 24, 25: Two next LS bytes of Microseconds

Byte 26: Start Code Emulation Prevention Byte (0xFF)

Bytes 27, 28: Two LS bytes of Microseconds

This represents the 64 bit microsecond UTC time where byte 18=MSB, bytes 19,21,22,24,25,27 are intermediate bytes and byte 28=LSB. In both fields, Byte 1 is transmitted first.

3.14.3.3 MPEG-2 Video Elementary Stream Editing Information

Additional information that may be carried in the user data area of a video elementary stream is described in SMPTE 328M [41]. One of the additional metadata elements is a 64 bit time code, which complies with SMPTE 12M [26], and SMPTE 309M

[27]. The time code represents the time that the frame was captured (HH:MM:SS:FF), and it contains a date as defined in SMPTE 309M.

3.14.3.4 H.264

As with MPEG-2, the H.264 Compression Format provides places to include a time stamp in the video stream. Both of the time stamps described below are placed in the Supplemental Enhancement Information (SEI) Message.

3.14.3.5 Pic_Timing Time Stamp

The H.264 format, specified in [11] provides for an optional time stamp to be defined in the Supplemental Enhancement Information (SEI) Message. The "picture timing SEI message" (pic_timing) specifies the time as HH:MM:SS:FF. It is a persistent time stamp that reflects the time of frame capture, and it also contains flags to specify whether the video is drop-frame, and whether there is a discontinuity in the video time line.

For H.264 compression systems, it is strongly recommended that the pic_timing field in the SEI Message be filled in with a time stamp that represents UTC time for H.264 video streams for all motion imagery systems.

3.14.3.6 User Data

The H.264 format also allows user defined data to be associated with a particular video frame using the user data unregistered SEI Message. The user data unregistered SEI Message may be used in systems which are required to associate a highly accurate, microsecond resolution time stamp with the video frame. This UTC time stamp shall be derived from GPS as described in section 4 of MISB RP 0603 [40] and will be formatted as defined in Annex A of the same. The user data unregistered message consists of two fields as defined below:

Uuid_iso_iec_11578 is a 16 byte field that shall be set to the value:

Bytes 1-8: 0x4D, 0x49, 0x53, 0x50, 0x6D, 0x69, 0x63, 0x72,

Bytes 9-16: 0x6F, 0x73, 0x65, 0x63, 0x74, 0x69, 0x6D, 0x65

This represents the ASCII string: "MISPmicrosectime"

User_data_payload_bytes is a variable length field. For this application, 12 bytes will be used as follows:

Byte 1: Status

Bit 7 0= GPS Locked (internal clock locked to GPS)
1= GPS Flywheel (internal clock not locked to GPS, so it is running on an internal oscillator)

Bit 6 0= Normal (time incremented normally since last message)
1= Discontinuity (time has not incremented normally since last message)

Bit 5 0= Forward (If Bit 6=1, this indicates that the time jumped forward)

1= Reverse (If Bit 6=1, this indicates that the time jumped backwards)

Bits 4-0: Reserved (=1)

Bytes 2, 3: Two MS bytes of Microseconds

Byte 4: Start Code Emulation Prevention Byte (0xFF)

Bytes 5, 6: Two next MS bytes of Microseconds

Byte 7: Start Code Emulation Prevention Byte (0xFF)

Bytes 8, 9: Two next LS bytes of Microseconds

Byte 10: Start Code Emulation Prevention Byte (0xFF)

Bytes 11, 12: Two LS bytes of Microseconds

This represents the 64 bit microsecond UTC time where byte 2=MSB, bytes 3,5,6,8,9,11 are intermediate bytes and byte 12=LSB. In both fields, Byte 1 is transmitted first.

3.14.4 Time Stamping Metadata

Systems that are capable of time stamping both the video stream and the metadata stream have all of the information necessary to multiplex this information together in a synchronized motion imagery stream. The structure for KLV metadata is defined in [23]. The KLV element "User Defined Time Stamp (microseconds since 1970)" is typically used as the time stamp in a KLV stream. The definition and format of this KLV element is defined in MISB RP 0603 [40].

3.14.5 Carriage of Metadata in Transport Stream

If the requirements for a motion imagery system dictate that a metadata element is associated with a particular frame of video, or that the time associated with the metadata element is correlated to the same time line as the video, then [3] shall be used to transport the video and associated metadata in an MPEG-2 Transport Stream.

3.14.5.1 Asynchronous Carriage of Metadata

The transport of KLV metadata over MPEG-2 transport streams in an asynchronous manner[2] is defined in SMPTE RP 217 [29]. As shown in Figure 9, the metadata PES packets do not use Presentation Time Stamps (PTS) or Metadata Access Unit Wrappers. The relationship between the metadata and the video frames is typically established by their proximity in the video stream. This type of metadata carriage may be used to transport static metadata, or metadata which is not tied closely in time to the video.

---

[2] Although Section 2.12.3 of MPEG-2 Systems [3] defines three tools for asynchronous delivery of metadata, MISB chooses to support the legacy method defined in RP 217 [29], with the addition of useful descriptors found in [3].

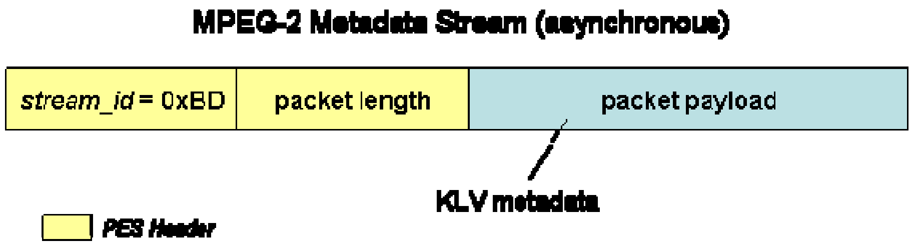

Figure 9: Asynchronous Metadata Stream

Metadata PES Stream

- The *stream_id* shall be 0xBD, indicating "private_stream_1."
- The *data_alignment_indicator* shall be set to one when the PES packet contains the beginning of a KLV item, and shall be set to zero otherwise.
- The delay of any data through the System Target Decoder buffers shall be less than or equal to one second. (This ensures that the metadata is in close proximity to the video frames that it relates to.)

  Note: Careful use of the buffer size and leak rate for metadata defined in the System Target Decoder (STD) Model (and as specified in the *metadata_std_descriptor*) can force a closer proximity of the metadata to the associated frame of video.

Program Map Table (PMT)

- The *stream_type* shall be 0x06, indicating "PES packets containing private data."
- The Metadata Stream shall be defined in the PMT as a separate Stream within the same Program as the Video Elementary Stream. [3] allows for multi-program Transport Streams, and methods for associating metadata in one program to video in another. Multi-program Transport Streams are not covered within the scope of MISB STANDARD 0604 [38].
- For legacy compliance with SMPTE RP 217 [29], the program element loop in the PMT shall contain a *registration_descriptor* as defined in [3], and the *format_identifier* field shall be set to 0x4B4C5641 (KLVA).
- The PMT shall contain a *metadata_descriptor* for each metadata service within the metadata stream. The *metadata_descriptor* shall be within the descriptor loop for the metadata stream. The *metadata_descriptor* contains the *metadata_service_id* for the service it describes. The following values are used to identify metadata types within the *metadata_descriptor*:

  metadata_format = 0xFF (specified by *metadata_format_identifier*)
  metadata_format_identifier = 0x4B4C5641 (KLVA)

Note: Earlier versions of [3] describe the use of the *registration_descriptor* to "uniquely and unambiguously identify formats of private data." The *metadata_descriptor*, however, provides more functionality, and is therefore specified.

- The PMT shall contain a single *metadata_std_descriptor* for the metadata stream.
- The PMT may contain other descriptors such as the *content_labeling_descriptor* and the *metadata_pointer_descriptor*.

The following is a sample *registration_descriptor*, *metadata_descriptor* and *metadata_std_descriptor* for a metadata stream containing asynchronous KLV metadata:

registration_descriptor

descriptor_tag = 0x05 (5)
descriptor_length = 0x04 (4)
format_identifier = 0x4B4C5641 = "KLVA"

metadata_descriptor

descriptor_tag = 0x26 (38)
descriptor_length = 0x09 (9)
metadata_application_format = 0x0100-0x0103 (see Table 4)
metadata_format = 0xFF
metadata format identifier = 0x4B4C5641 = "KLVA"
metadata_service_id = 0x00
decoder_config_flags = '000'
DSM-CC_flag = '0'
reserved = '1111'

metadata_std_descriptor

descriptor_tag: 0x27 (39)
descriptor_length: 0x09 (9)
reserved = '11'
metadata_input_leak_rate: (determined by encoder)
reserved = '11'
metadata_buffer_size: (determined by encoder)
reserved = '11'
metadata_output_leak_rate: (determined by encoder)

Note that the *metadata_output_leak_rate* must be specified for asynchronous metadata.

metadata_application_format (type of KLV metadata)	
0x0100	General
0x0101	Geographic Metadata
0x0102	Annotation Metadata
0x0103	Still Image on Demand

Table 4: KLV metadata type

3.14.5.2 Synchronous Carriage of Metadata

Several ways to carry metadata over MPEG-2 transport streams are detailed in [3]. MISB STANDARD 0604 [38] specifies the method outlined in [3] Section 2.12.4 “Use of PES packets to transport metadata” for transporting metadata that is synchronized with the video stream. This method provides a way to synchronize metadata with video using the Presentation Time Stamp (PTS) found in the Packetized Elementary Stream (PES) header. This time stamp is coded in the MPEG-2 Systems PES layer, and is relevant for H.264 as well as MPEG-2.

The metadata may or may not be sampled at the same time as a video frame depending upon the system design. If it is sampled at the same time as a video frame, the metadata and video frame will have the same PTS. If the metadata is not sampled at the same time as the video frame, it will be stamped with a different PTS, but exist on the same timeline as the video frame.

Figure 10 shows the general structure of a PES packet in the metadata bit stream. In the most common implementation, the packet payload would consist of a single metadata cell that includes a five-byte header followed by KLV metadata.

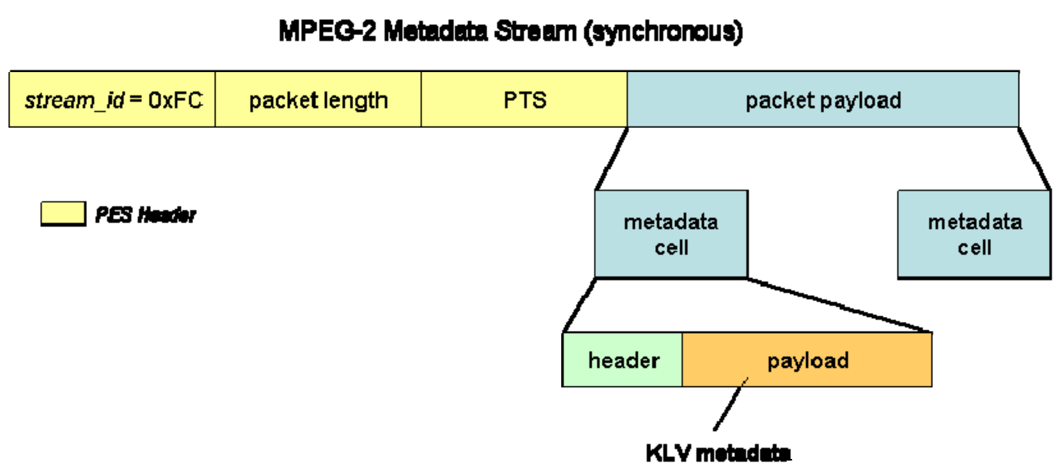

Figure 10: Synchronous Metadata Stream

A metadata service is defined in [3] as "a coherent set of metadata of the same format delivered to a receiver for a specific purpose." When transporting metadata using this service, a unique *metadata_service_id* is assigned to each service. Each metadata service is represented by a collection of metadata access units that are transported in PES packets.

Details of the implementation of this method are given below.

Metadata PES Stream

- The *stream_id* shall be 0xFC, indicating "metadata stream".
- Each PES packet shall have a PTS to be used to synchronize the metadata with the video frames.
- In each PES packet that carries metadata, the first PES packet data byte shall be the first byte of a Metadata Access Unit Cell.
- The PTS in the PES header shall apply to each Access Unit contained in the PES packet.
- The PTS shall signal the time that the metadata Access Unit becomes relevant. It is assumed that the metadata is decoded instantaneously (i.e., no DTS shall be coded). If a video frame and a metadata Access Unit have the same PTS, then they were sampled at the same time.
- Each metadata Access Unit may be carried in one or more Access Unit Cells.
- The delay of any data through the System Target Decoder buffers shall be less than or equal to one second. (This ensures that the metadata is in close proximity to the video frames that it relates to.)

  Note: Careful use of the buffer size and leak rate for metadata defined in the System Target Decoder (STD) Model (and specified in the *metadata_std_descriptor*) can force a closer proximity of the metadata to the associated frame of video.

Program Map Table (PMT)

- The *stream_type* shall be 0x15, indicating "Metadata carried in PES packets."
- The Metadata Stream shall be defined in the PMT as a separate stream within the same Program as the Video Elementary Stream. [3] allows for multi-program Transport Streams, and methods for associating metadata in one program to video in another. Multi-program Transport Streams are not covered within the scope of MISB STANDARD 0604 [38]
- The PMT shall contain a *metadata_descriptor* for each metadata service within the metadata stream. The *metadata_descriptor* shall be within the descriptor loop for the metadata stream. The *metadata_descriptor* contains the *metadata_service_id* for the service it describes. The following values are used to identify metadata types within the *metadata_descriptor*:

metadata_format = 0xFF (specified by metadata format identifier)
metadata_format_identifier = 0x4B4C5641 “KLVA”

- The PMT shall contain a single *metadata_std_descriptor* for the metadata stream.
- The PMT may contain other descriptors such as the *content_labeling_descriptor* and the *metadata_pointer_descriptor*.

The following is a sample *metadata_descriptor*, *metadata_std_descriptor* and *metadata_AU_cell* header for a metadata stream containing synchronous KLV metadata.

metadata_descriptor

descriptor_tag = 0x26 (38)
descriptor_length = 0x09 (9)
metadata_application_format = 0x0100-0x0103 (see Table 4)
metadata_format = 0xFF
metadata format identifier = 0x4B4C5641 = “KLVA”
metadata_service_id = 0x00
decoder_config_flags = ‘000’
DSM-CC_flag = ‘0’
reserved = ‘1111’

metadata_std_descriptor

descriptor_tag: 0x27 (39)
descriptor_length: 0x09 (9)
reserved = ‘11’
metadata_input_leak_rate: (determined by encoder)
reserved = ‘11’
metadata_buffer_size: (determined by encoder)
reserved = ‘11’
metadata_output_leak_rate: (unspecified; recommend setting to 0)

Note that the *metadata_output_leak_rate* is unspecified for synchronous metadata. The recommended value is 0.

Metadata_AU_cell (5-byte header)

metadata_service_id = 0x00
sequence_number = (supplied by encoder; increments each cell)
cell_fragmentation_indication = ‘11’, ‘10’, ‘01’ or ‘00’
decoder_config_flag = ‘0’
random_access_indicator = ‘0’ or ‘1’
reserved = ‘1111’
AU_cell_data_length = (supplied by encoder)

3.14.6 Transition

Many motion imagery systems have been developed based on SMPTE RP 217 [29] for asynchronous carriage of metadata in MPEG-2 Transport Streams. A synchronous method for transporting metadata with the associated video streams is provided in [3]. As systems advance and strive for more accurate metadata, migrating to this new method of transporting metadata is important.

New systems and applications must be capable of handling metadata using both the format in SMPTE RP 217 [29] and [3]. It will be relatively straightforward for motion imagery systems to add support for [3]. Minor changes must be made to the transport layer (multiplexing and demultiplexing) of the motion imagery stream.

4 File Formats

4.1 STANDARD 0205 - Use of MPEG-2 System Streams for Simple File Applications

For simple file applications, MPEG-2 Transport or Program Streams may be used for NATO applications. All NATO systems must be able to receive and decode both Transport and Program Stream files.

4.2 STANDARD 0902 - Advanced File Format

In the other applications, where digital video files need to be exchanged, real-time or not between collection platforms users and data-bases with random access to the motion imagery based on metadata indexing, the Material Exchange Format (MXF) SMPTE 377M [42], can be used. This format makes use of the sampling, compression, and metadata rules as defined in the present Annex, and provides advanced features for access and exchange over communication networks. It is expected that this standard will be mandated in future revisions of this STANAG.

As MXF covers a large number of options and application domains, the present standard restricts as follows the applicable MXF possibilities to a minimum level mandated to achieve interoperability between the implementing nations:

- Only operational patterns 1a (OP-1a) and 1b (OP-1b) as per SMPTE 378M [43] and SMPTE 391M [44], respectively, will be used for file exchange.
- The essence will be wrapped frame by frame using the generic container as per SMPTE 379M [45] and SMPTE 381M [46].
- From the complete list of metadata sets and properties given by SMPTE 380M [47], the participating parties will be required to interpret only a minimum profile (derived from ASPA Profile) listed in AEDP-8 Edition 2. It must be noted that it is a design rule of MXF players to accept dark (unknown) data which obviously will not be interpreted.
- The dynamic metadata will be interleaved within the body.

4.3 STANDARD 0218 - Timing Reconciliation Universal Metadata Set for Digital Motion Imagery

4.3.1 Scope

This standard defines a timing reconciliation metadata set to correct (reconcile) the original capture time of metadata with a User Defined Time Stamp usually associated with the capture time of the digital motion imagery or audio essence. Timing reconciliation metadata is not required if the application using the metadata does not depend on the amount of timing error or uncertainty between the metadata capture and the video or audio essence capture.

4.3.2 Introduction

The time of metadata insertion into an encoded essence stream, file, or frame can be different from the time of its initial capture or sampling by as much as several seconds. In addition, the capture time of the metadata may be different from the capture time of the essence. As a result, the time stamp that associates a stream, a file, or a frame with an element or set (or pack) of metadata will be incorrect. When an application requires more precise information about the time of metadata capture this standard shall be used to convey the metadata capture time as a metadata set that is linked to another set or pack of metadata or to an individual metadata element. All metadata shall be represented using big-endian (most significant byte – MSB - first) encoding. Bytes shall be big-endian bit encoding (most significant bit – msb - first).

4.3.3 Timing Reconciliation Metadata for Digital Motion Imagery

The following time stamp metadata element shall be used to link accurate capture time of metadata to other metadata or essence as described in this section:

06 0E 2B 34 01 01 01 04 07 02 01 01 01 05 01 00	User Defined Time Stamp – Microseconds since 1970 (msb first)

4.3.3.1 Timing Reconciliation Metadata Inside Metadata Sets or Packs

The User Defined Time Stamp metadata element alone may be placed within a metadata set or pack when it unambiguously applies to each and every element of metadata within the set or pack. Its presence in the metadata set or pack shall be the only indication that it is the creation or capture date and time for the contents of that entire set or pack and, if used, it shall always be the first element of metadata within the applicable set or pack. When only a Timing Bias Correction is present in the set it shall be applied to the time to which it is linked or to the time in the set to which it be linked. When both a User Defined Time and Timing Bias Correction are present in the set the Time Bias Correction shall be applied to the User Defined Time in the set.

4.3.3.2 Timing Reconciliation Universal Metadata Set Linked to Other Metadata

The User Defined Time Stamp and a Timing Bias Correction (if needed) may be linked selectively to individual metadata elements or to metadata sets, packs or labels using the Timing Reconciliation Metadata Set (detailed in Table A1).

16-byte Set Designator[3]		Metadata Set or Element Name
Universal Set		
06 0E 2B 34 02 01 01 01	07 02 01 03 01 00 00 00	Timing Reconciliation Metadata Set
06 0E 2B 34 01 01 01 04	07 02 01 01 01 05 01 00	User Defined Time Stamp – Microseconds since 1970 (msb first)
06 0E 2B 34 01 01 01 04	03 01 03 03 03 01 00 00	Timing Bias Correction (microseconds – msb first)
06 0E 2B 34 01 01 01 03	03 01 03 03 04 00 00 00	Description of Timing Bias Correction
06 0E 2B 34 01 01 01 03	01 03 02 00 00 00 00 00	Item Designator ID

Table A1 –Timing Reconciliation Metadata Set

When a single metadata element is linked to a Timing Reconciliation Universal Metadata Set the Timing Reconciliation Universal Metadata Set shall contain an Item Designator ID whose Value is the 16-byte Universal Label Key for the single metadata element to which it is linked. The Timing Reconciliation Universal Metadata Set shall always preceed the metadata element to which it is linked. Figure A1 is an informative example of a Timing Reconciliation Universal Metadata Set linked to one metadata element.

When some but not all metadata elements within a set or pack must be linked to a Timing Reconciliation Unversal Metadata Set the Timing Reconciliation Universal Metadata Set shall contain one individual Item Designator ID for each metadata element to which it is linked. The Timing Reconciliation Universal Metadata Set shall always preceed all of the elements of the metadata set or pack to which it is linked.

When all metadata elements within a set or pack are linked to a Timing Reconciliation Universal Metadata Set and use of the method above may be ambiguous, the Timing Reconciliation Universal Metadata Set shall contain one individual Item Designator ID for the metadata set or pack to which it is linked. The Timing Reconciliation Universal Metadata Set shall always preceed the metadata set or pack to which it is linked.

[3] All Set UL Designators are tentative and may be changed as the SMPTE Sets Registry is developed

4.3.3.3 Timing Reconciliation Universal Metadata Set Placement in Streams

When a Timing Reconciliation Universal Metadata Set is used within an MPEG-2 stream, the metadata linked to it shall always appear in each “I” frame. This does not preclude it also being used in “P” and /or “B” frames but its use in each “I” frame is mandatory.

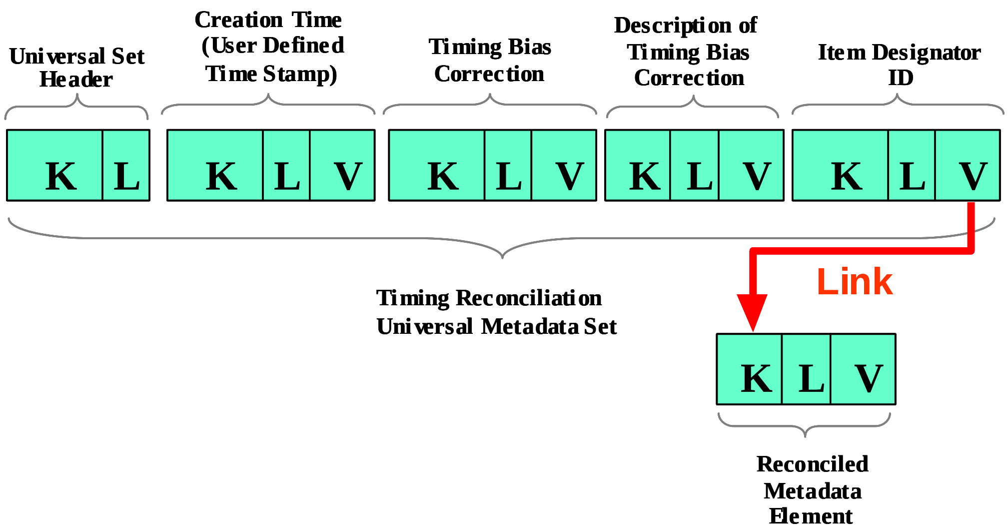

Figure A1: Example of a Timing Reconciliation Universal Metadata Set Linked to a Metadata Element